The mechanisms underlying seasonal timing of breeding

a multi-level approach using bi-directional genomic selection on timing of egg-laying

Irene C. Verhagen

Thesis committee

Promotor

Prof. Dr Marcel E. Visser

Special Professor Ecological Genomics

Wageningen University & Research

Head of department, Department of Animal Ecology

Netherlands Institute of Ecology (NIOO-KNAW), Wageningen

Co-promotor

Dr Veronika N. Laine

Postdoctoral researcher, Department of Animal Ecology

Netherlands Institute of Ecology (NIOO-KNAW), Wageningen

Other members

Prof. Dr Barbara Helm, Rijksuniversiteit Groningen

Prof. Dr Bas J. Zwaan, Wageningen University & Research

Dr Tyler J. Stevenson, University of Glasgow, UK

Dr Koen J. F. Verhoeven, Netherlands Institute of Ecology (NIOO-KNAW)

This research is conducted under the auspices of the Wageningen Institute of Animal Sciences (WIAS) graduate school.

The mechanisms underlying seasonal timing of breeding

a multi-level approach using bi-directional genomic selection on timing of egg-laying

Irene Charlotte Verhagen

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.J.P. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board,
to be defended in public
on Friday 8 November 2019
at 11:00 a.m. in the Aula

Irene C. Verhagen
The mechanisms underlying seasonal timing of breeding: a multi-level approach using bi-directional genomic selection on timing of egg-laying
222 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)
With references, with summaries in English and in Dutch

ISBN 9778-94-6395-039-8
DOI 10.18174/496393

"We snatch in vain at Nature's veil,

She is mysterious in broad daylight,

No screws or levers can compel her to reveal,

The secrets she has hidden from our sight."

(Johann Wolfgang von Goethe, 1808)

Table of contents

CHAPTER 1

General introduction

Evolution and natural selection

Evolution, the process that leads to the formation of and change in biological systems in response to their environment, has shaped life since its beginning and continues to do so. All organisms, simple and complex, evolve over different scales in both time and space. In general, evolution is described and studied as two distinct hierarchical processes; macro-evolution and micro-evolution. Macro-evolution refers to the origin of new morphological forms, species and divisions of the taxonomic hierarchy above the species level, together with the origin of complex adaptations, such as the eye (Reznick & Ricklefs 2009), but on which I will not dwell further. In contrast, micro-evolution refers to the processes of adaptive modifications (i.e. through alteration in allele frequencies in gene pools) within and among populations, propelled by mutation, migration, genetic drift and natural selection. The latter, put first into words by, independently, both Charles Darwin (1859) and Alfred Russel Wallace (1858), is usually defined as the consequence of three organismal properties: (1) variation among organisms of a population, (2) differential reproduction, and (3) traits that are important for survival or reproduction show heritability (Darwin 1859; Wallace 1858).

In general, natural selection is the tendency of beneficial traits to increase in frequency in populations to make them better able to survive and reproduce in its (changing) environment or habitat (Maynard Smith 1989). Basically, when variation in a heritable trait is caused by differential reproduction, the change in the average phenotype of the population is the result from the greater contribution to each generation by the *fittest* individuals (i.e. individuals with traits best adapted to its environment to survive) (Lewontin 1974; Roughgarden 1979), or as Darwin (1859) stated "One general law, leading to the advancement of all organic beings, namely, multiply, vary, let the strongest live and the weakest die." Therefore, natural selection 'prefers' the fittest organism (or population) with respect to a given trait, leading to adaptation of that trait. A textbook example is the peppered moth (*Biston betularia*), where natural selection acted in favour of dark-coloured moths as a consequence of air pollution during the Industrial Revolution in Great Britain and a subsequent increase in predation of white-coloured moths by birds (Kettlewell 1958). As such, the population adapted by increasing the occurrence of dark-coloured moths and persisted.

Adapting to a warming world: novel and intensified selection pressures

However, due to twentieth-century anthropogenic disturbances, most particularly the continuous increase in ambient temperatures (IPCC 2014), many species experience novel and intensified selection pressures (Merilä & Hoffmann 2016). As such, global warming has important ecological consequences and evolutionary impacts on wild populations (Peñuelas & Filella 2001; Walther et al. 2002). Shifts in phenology, i.e. recurring seasonal events, by organisms are among the most observed and studied in relation to climate change (e.g. Parmesan & Yohe 2003; Root et al. 2003; Walther et al. 2002, but see Cohen et al. 2018). Species across taxa have shown mostly advancements in seasonal timing of, among others,

migration (Hüppop & Hüppop 2003), breaking hibernation (Inouye et al. 2000), flowering time (Cayan et al. 2001; Menzel & Dose 2005), spawning date (Beebee 1995) and emergence (Forister & Shapiro 2003; Roy & Sparks 2000). However, these shifts differ in rate between trophic levels, most notably in higher trophic levels (Thackeray et al. 2010, 2016). Therefore, species-specific variation in phenological responses to the environment can disrupt the synchrony of ecological interactions (Harrington et al. 1999; Peñuelas & Filella 2001; Visser & Both 2005). As such, the impacts of climate change may be mediated through these so-called 'mismatches', as they result in major consequences for individual fitness and population persistence (Both et al. 2006; Møller et al. 2008; Platt et al. 2003; Visser & Gienapp 2019; Winder & Schindler 2004; but see Reed et al. 2013).

Before predictions can be made about how future climate change will affect populations, we however need to obtain a far better understanding of the underlying physiological mechanisms that link environmental cues to life-history decisions, like timing of breeding (Pörtner & Farrell 2008) and if so, *how* environmental cues affects these. It proves difficult, at least in wild populations, to determine the actual environmental factor that changes phenotypes and drives selection, which could be numerous (Merilä & Hendry 2014).

Avian seasonal timing of breeding

In birds from temperate zones, continuously rising temperatures have the most predominant effect on the timing of their breeding, i.e. the date at which females initiate egg-laying (e.g. Both et al. 2004; Brown et al. 1999; Crick et al. 1997; Dunn & Winkler 1999; Love et al. 2010; Matthysen et al. 2011). Temperate zone birds rely on food resources to raise their offspring, which usually become available or increases in abundance for a brief period each year (Lack 1968; Perrins 1970; Verhulst & Tinbergen 1991), making them seasonal breeders. This period varies every year due to yearly varying environmental conditions. As such, the onset of breeding needs to be optimally timed, or synchronized, to this increased availability of food resources in order to support successful rearing of offspring (Charmantier et al. 2008; van Noordwijk et al. 1995; Perrins 1965; Sheldon, Kruuk & Merilä 2003).

This synchronization is a delicate matter, because the recrudescence of gonads and early offspring development occurs well before offspring care, which is due to the length of the egg incubation period. As such, synchronization of the breeding season to environmental conditions that favour offspring needs, requires extensive physiological and behavioural 'preparations'. Growth and maturation of the reproductive system can take up 6-8 weeks and is accompanied and followed by finding a mate and nest site, nest building, and ultimately egg-laying and incubation. It follows from this temporal separation of reproductive events that birds need to be able to anticipate or *predict* the onset of the breeding season, for which they use information from environmental cues. These cues are predominantly photoperiod (Dawson et al. 2001; Farner 1985; Follett 1984; Gwinner 1986; Sharp 1996; Silverin et al. 1993; Wingfield 1993) and temperature (Caro & Visser 2009; Lambrechts & Visser 1999; Williams 2012), which provide information about the time of

the day and year, and the local (optimal) conditions (Dawson 2008; Wingfield et al. 1992; Wingfield & Kenagy 1991), respectively.

As such, females often breed earlier after a warm spring than a cold spring (Both & Visser 2001; Brommer et al. 2005; Brommer et al. 2003; McCleery & Perrins 1998; Schaper et al. 2012; Visser et al. 2004), because warmer temperatures predicts insect emergence to be earlier as well (Visser 2008; Visser & Both 2005). This expression of phenotypic plasticity, i.e. the ability of an individual (or genotype) to express different phenotypes in different environments (Woltereck 1909), enables genotypes to produce a better phenotype-environment match in order to track natural changes in their environment and potentially

Box 1: study system

The great tit is a passerine bird, which is widespread and common within the temperate zone throughout Europe, the Middle East and Asia (Perrins 1979). It inhabits any sort of woodland and is in general resident. Great tits are predominantly insectivorous, but broaden their diet during the winter period. They are distinctively feathered with black head, neck and breast stripe, white cheeks and yellow under parts. Males can be recognized from females by a more prominent breast stripe. From January to March great tits form pairs and start nest building followed by egg-laying in April. Despite a wide variety in the onset of egg-laying between populations at different latitudes, great tits are strictly seasonal breeders. Great tits are typical cavity nesters, mostly in tree holes with the entrance only slightly larger than its body, and therefore it nests easily in nest boxes. The nest is built entirely by the female and usually has a foundation of moss and is lined with hair, wool or feathers. In April eggs are laid; with one egg every day resulting in clutches up until 12 eggs. After incubation of 12-15 days the young hatch and fledge after 16-22 days. Parents care for their young until they reach independence. (Hinde 1952)

The individuals studied in this thesis, descent from the wild population in the Hoge Veluwe National Park, The Netherlands that has been continuously monitored since 1955, making it one of few long-term wild study populations. This long-term monitoring demonstrated disrupted synchrony in phenology between oak (*Quercus robur*) and winter moth (*Operophtera brumata*) (van Asch et al. 2007, Visser et al. 2006) and winter moth and great tit (Visser et al. 1998, Visser et al. 2006) due to continuously increasing temperatures in spring. When oak's bud burst in spring, winter moth's eggs hatch, because the winter moth caterpillars feed on the young oak leaves. Great tits, later in spring, feed their chicks with the caterpillars. Optimal timing of great tit hatching is therefore of fundamental importance for their survival and growth. Whereas three winter moth populations in the Netherlands have now restored their match with oak bud burst (van Asch et al. 2013), the great tit has not sufficiently advanced to catch up with the winter moth and selection towards earlier laying has increased (Visser et al. 1998).

'catch up' with trophic levels they interact with (Levins 1968; Parmesan 2006; Stenseth et al. 2002; Visser 2008; Walther et al. 2002). Despite birds being plastic in timing of breeding in response to temperature, climate change can disrupt the association between the environmental variables over time. A females' 'decision' to lay is based on this prediction

informed by cues from the 'environment at the time of decision making' early in the season, e.g. photoperiod and temperature, whereas the fitness benefits of the timing decision are determined by the 'environment at the time of selection', e.g. food abundance (Visser et al. 2004). If environmental variables lose their predictability, due to their disrupted association, a females' plastic response becomes insufficient (Gienapp et al. 2014; van Noordwijk & Muller 1994; Visser et al. 2004). As a result, numerous bird species have been unable to sufficiently advance their breeding (Both et al. 2006; McKinnon et al. 2012; Nielsen & Moller 2006; Shultz et al. 2009; Thomas et al. 2001). For example, the level of plasticity in egg-laying date was insufficient to reach new phenotypic optima, resulting in a mismatch between food abundance and offspring needs and subsequent intensified selection for earlier egg-laying in a wild population of great tits (*Parus major*) (Visser et al. 1998), which is also the study species in this thesis (**Box 1**). We now know *why* birds in general, and great tits in particular, should lay earlier, and it is necessary for (plasticity in) timing of breeding to evolve through micro-evolution.

Mechanisms underlying avian timing of breeding, a complex trait

Phenotypic plasticity does not explain the *between*-individual variation in timing of breeding within a single year. Interestingly, individuals within a population that experience roughly the same environment, show large variation in breeding time. Aviary experiments on genetically related great tits suggest that there is genetic variation in cue sensitivity among individuals (Visser et al. 2011). So *what* explains this variation and *where* can we find it?

For example, the change in photoperiod (i.e. when days get longer), like any other environmental cue, needs to be perceived, assessed, transduced and finally translated into a neuroendocrine response in order to effect a change in physiology and behaviour. This cascade of the photoperiodic pathway (**Box 2**) from perception to a neuroendocrine response is complex, but rather well known on the molecular level for timing of breeding in birds (Nakane & Yoshimura 2014; Ubuka et al. 2013; Yoshimura et al. 2003). Activation of this photoperiodic pathway precedes and is essential for activation of the hypothalamic-pituitary-gonadal-liver (HPGL) axis (**Box 2**), which plays a major role in reproduction, by further regulating gonadal function and ultimately facilitation of timing of breeding (Williams 2012 and references therein). While the function and signalling pathway for photoperiod is clear in timing of breeding (Dawson 2008; Dawson et al. 2001; Nakane & Yoshimura 2014; Ubuka, Bentley & Tsutsui 2013; Yoshimura et al. 2003), this remains largely elusive for temperature (Caro et al. 2013). Temperature, is believed to be the most influential environmental cue to 'fine-tune' seasonal breeding behaviour of birds (Caro et al. 2009; Lambrechts & Visser 1999; Williams 2012), which is most apparent in *female* birds (Caro et al. 2009; Dawson 2008; Williams 2012). Yet, defining which temperature profile, and when a certain profile induces egg-laying poses a challenge (Schaper 2012). From previous research, we know that temperature has an effect on egg-laying (Meijer et al. 1999; Salvante et al. 2007) and that this effect is causal (Visser et al. 2009, but see **Chapter 3**). In addition, the relevant information seems to reside in the periods of increasing temperature

instead of the mean temperature (Schaper et al. 2012).

Box 2: the hypothalamic-pituitary-gonadal-liver (HPGL) axis

The HPGL axis (Figure 1) is a key pathway underlying seasonal reproduction and associated physiological, morphological and behavioural traits (Dawson 2008). In seasonally breeding birds, photoreceptors in the brain receive information about, among others, increasing photoperiod, which is send to the pars tuberalis in the pituitary to induce expression of thyroid-stimulating hormone (TSH). TSH acts on its receptors in the third ventricle where it stimulates expression of type II iodothyronine deiodinase (DIO2). In its turn, DIO2 converts thyroxin (T4) into the biologically active thyroid hormone triiodothyronine (T3). T3 triggers the synthesis and secretion of gonadotropin-releasing hormone (GnRH) in the hypothalamus, which marks the activation of the HPGL-axis (Nakao et al. 2008, Yoshimura 2010). Leaving the hypothalamus, GnRH directly acts on the anterior pituitary gland and triggers the synthesis and release of the gonadotropins luteinizing hormone (LH) and follicle-stimulating hormone (FSH). In females, these hormones regulate the activity of the ovary by stimulating oogenesis and secretion of sex steroids such as testosterone (T) and estradiol (E_2), which bind to the receptors of their target tissues in the HPGL axis creating a feedback system (Meethal et al. 2009). In addition to the hypothalamus, pituitary gland and ovary, the liver plays an important role in *female* birds. The liver initiates, dependent on E_2, vitellogenesis, i.e. the production of two yolk precursors; vitellogenin (VTG) and yolk targeted very-low density lipoprotein (VLDLy) (Williams 2012, Yoshimura 2006, 2010). These yolk precursors are the primary sources of yolk protein and lipid for the developing embryo (Walzem 1996).

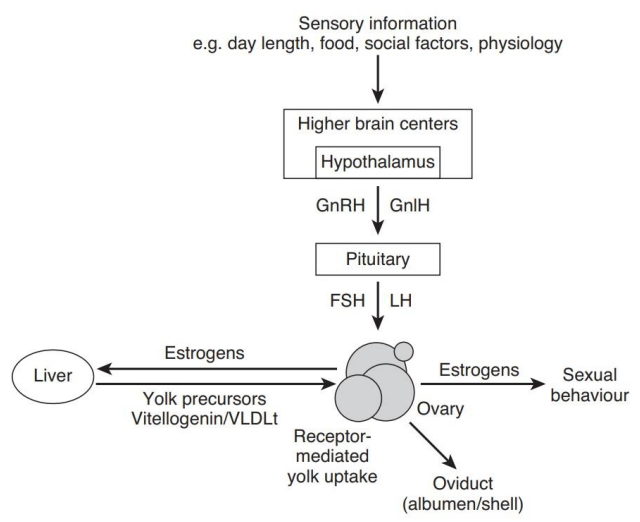


Figure 1. HPGL axis (from Williams 2012).

After transduction and integration of environmental cues, such as temperature, the neuro-endocrine cascade of physiological events (**Box 2**) triggered by the expression of genes, results in a wide range in breeding time among females. Therefore, the causation of variation in timing is rooted in the underlying physiology. It has been widely assumed that the hypothalamus, considered as the final integration point of environmental cues, the pituitary gland, and the neural centres 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). But is this top-down regulation entirely true? For example in great tits, individual variation in reproductive physiology (e.g. hormone

concentrations, gonadal development) did not predict variation in onset of egg-laying (Schaper et al. 2012). This emphasizes the ability to 'fine-tune' the timing of breeding to supplementary environmental information (see **Chapter 4** and **5**), irrespective of individual variation in developmental state. Understanding how different parts of the HPGL axis are regulated by environmental cues, what is the molecular basis of this and in which component, is hence important (and requires investigation of all components of the HPGL axis), especially in the context of adaptation to climate change, but this remains a major lacuna in our knowledge. One reason is that timing of breeding, as many other life-history traits, is a so-called 'complex trait' (Garland 2003). Complex traits are regulated by (the expression of) multiple, if not many, genes that cooperate in physiological processes best described as pathways or networks in interaction with the environment (e.g. temperature, see above).

From genotype to phenotype

In an organism, the genotypic level generates the heritable variation, whereas selection operates at the level of the phenotype changing trait distributions within a population. Identifying the genes underlying life-history traits, e.g. timing of breeding, is a major challenge, because as said above, the relationship between the genotype and the phenotype is not one-on-one, but much more complex. So, where the genotype sets the limits in variation in timing of breeding, environmental factors, when interacting with the epigenome, largely determine expression of the genotype through so-called 'endo-phenotypes' (te Pas et al. 2017): the transcriptome, proteome, metabolome and microbiome, which in their turn regulate the expression of the phenotype. The proteome, metabolome and microbiome will be discussed in **Chapter 7**, as these are beyond the scope of this thesis. Explanation of the 'black box' between the genetic variability and the observed phenotype (i.e. timing of breeding), warrants investigation of these 'endo-phenotypes' of the mechanisms underlying timing of breeding. Further, these endo-phenotypes do not function independently, but through feedback loops within and between the endo-phenotypes and as such, regulate the trait.

Genotype – nucleic acid sequence

Linking genes to traits ultimately requires genome sequence information, or an organisms' hereditary information. For great tits there is an excellent molecular toolbox available (Derks et al. 2016; Kim et al. 2018), among which a de novo high-quality genome assembly (Laine et al. 2016). Genetic determinants of phenotypic variation, i.e. mapping of quantitative trait loci (QTL), revolutionized the study of complex traits, because of the application of genome-wide association studies (GWAS). GWAS, aim to find markers, mostly single nucleotide polymorphisms (SNPs), associated with phenotypic variation by using SNP chips. Recently, aided by a previously developed 10k SNP chip (van Bers et al. 2010, 2012), a 650k SNP chip for great tits was developed to explore the genetic architecture

of exploration behaviour (Kim et al. 2018). In addition, using the 650 SNP chip, no significant association between SNPs and egg-laying date or annual spring temperature was found (Gienapp et al. 2017). However, a better fit with annual spring temperature suggests that the effects of genes shaping seasonal timing depended on annual spring temperature and taking these effects into account will lead to a better understanding of their evolutionary potential. In **Chapter 2** this SNP chip is further employed in relation to timing of breeding.

Epigenome – chemical modifications of DNA

Theodosius Dobzhansky (1973) wrote: "The entire evolutionary development of the living world has taken place not by invention of new 'letters' in the genetic 'alphabet' but by elaboration of ever-new combinations of these letters" (Dobzhansky 1973). True this may be, it is not the whole story, because genes need to be activated, or expressed, in order to regulate processes in life. For many complex traits, not only genomic variation in the gene coding sequences, but rather the differential expression of genes, is associated with their variability (Cookson et al. 2009; Frésard et al. 2013). Under the influence of the environment, the epigenome (and transcriptome and proteome, see below), i.e. chemical modifications of the underlying DNA sequence without modifying the sequence, regulates gene expression through several epigenetic mechanisms (Bender 2004; Bird 2007; Jaenisch & Bird 2003; Laird 2003; Richards 2006).

When genetic diversity (i.e. the number of 'combinations of letters in the genetic alphabet') is insufficient to adapt rapidly to environmental changes, epigenetic mechanisms might provide variation and may enable organisms to rapidly adjust their phenotypes in response to changing environments (Liebl et al. 2013; Tammen et al. 2013). Therefore, epigenetic mechanisms, like DNA methylation, may be involved in mediating the effects of climate change. Whether this is adaptive, however, remains to be determined. In birds, the underlying (epi)genetic basis for timing of breeding and plasticity in timing of breeding are currently unknown. However, several studies in plants (Law & Jacobsen 2010; Wilschut et al. 2016) and rodents (Stevenson & Prendergast 2013) suggest that epigenetic processes, such as DNA methylation, are involved in seasonal timing of life-history traits. Varying methylation patterns throughout the breeding season (Viitaniemi et al. 2019) and between reproductive stages (Lindner et al. *in prep.*) in female great tits, support this assumption. Further, variation in methylation in great tits distributed across Europe is correlated with rates of molecular evolution, indicating that methylation plays an important role in evolution (Laine et al. 2016).

In general, which changes in methylation result in variation in, for example, timing of breeding, is largely unknown. These changes could be detected by repeated measurements over time and relate the found changes to the (changes in) the phenotype studied. Blood, being a tissue that could repeatedly be collected from the same individual, provides such an opportunity, given that methylation patterns in blood reflect methylation patterns in other tissues. One study found a significant correlation between methylation in brain and blood in great tits (Derks et al 2016), which is promising for studies regarding the effect of climate

change in wild populations. In addition, whether methylation patterns reflect gene expression patterns within and between tissues, remains unknown too in species other than the human and rodents. To address this gap, we explore in **Chapter 6** correlations between changes in methylation levels and between changes in methylation and changes in expression levels within and between several tissues of the HPGL axis.

Transcriptome – RNA

The genotype and environmentally induced epigenetic mechanisms are regulate the gene activity in the genome, i.e the transcriptome or the sum of a genome's ribonucleic acid (RNA) transcripts. There are two main techniques, microarrays (quantification of a predefined set of sequences) and RNA sequencing (RNA-seq, quantifying all sequences) and may be validated using for example quantitative PCR (qPCR) (Lowe et al. 2017). However, the latter is also used as an independent transcriptomic technique. Studying the transcriptome enables linking genome-wide or candidate gene transcript or messenger RNA (mRNA) levels to variation in phenotypes, because mRNA serves as a kind of intermediary between genotype and phenotype.

For example, in captive Swainson thrush (*Catharus ustulatus*), a genome-wide study found differentially expressed genes in specific networks to be associated with migratory state, which had not been linked to avian migration previously (Johnston et al. 2016). Further, gene expression studies allow for identification of genes and pathways that respond to environmental factors. A single long day in two great tit populations from different latitudes, resulted in different expression of genes involved in photoperiodic time measurement, but only in the Northern population (Perfito et al. 2012). When housed with males, female European starlings (*Sturnus vulgaris*) showed elevated expression of specific genes involved in reproduction (Perfito et al. 2015).

Measuring the expression of genes in different tissues, conditions, or time points gives information on how genes are regulated. In general, studies investigating multiple or all components within the HPGL axis are scarce (Cánovas et al. 2014; MacManes et al. 2017; Maruska et al. 2011; Maruska & Fernald 2011; Perfito et al. 2015). Also, gene expression dynamics within the HPGL axis are not well-known in seasonally breeding birds. This, however, is essential in understanding where selection could act in order for organisms in general and birds in particular, to adapt to (rapidly) changing environments. **Chapter 4** explores the effect of date and temperature on genome-wide expression patterns within the HPLG axis. In **Chapter 5**, we set out to assess whether individual variation in egg-laying is reflected in individual variation in candidate gene expression levels and if so, where in the HPGL axis these differences occur.

Comparing physiological mechanisms of early and late breeding females

A promising way to study physiological mechanisms underlying complex traits, ideally in natural conditions, is by comparing individuals that differ (extremely) in their observed phenotype. Comparing individuals in wild conditions, though, makes pin-pointing which cues directly affect breeding, instead of through a third variable (e.g. food), impossible and manipulation of timing of breeding in wild birds has proven to be extremely difficult (Verhulst & Nilsson 2008). A powerful tool to create extreme phenotypes, alternatively in controlled conditions, are selection lines (Conner 2003). Selection on phenotypes, i.e. individuals selected based on their phenotypes, has proven to be successful in several studies in evolutionary and physiological ecology in several species (Kotrschal et al. 2013, 2016; Wang et al. 2018), including birds (Drent et al. 2003; van Oers et al. 2004, 2011). However, phenotypic selection is less accurate compared to genomic selection, i.e. individuals selected based on their genotypes, because the expected and observed response to selection show a higher correlation in the latter (Meuwissen et al. 2016; Wolc et al. 2015). Genomic selection is commonly implemented in domestic animal breeding and agriculture (Calus 2010; Jannink et al. 2010), to improve complex heritable traits of interest (e.g. milk yield in cows). It follows, that in theory, bi-directional genomic selection could be applied to avian timing of breeding, as it is a heritable complex trait (e.g. Gienapp et al. 2006; Mccleery et al. 2004). Using genomic selection in wild species is unprecedented, until recently in great tits (Gienapp et al. 2019), but will be invaluable in determining where the physiological mechanisms vary genetically. **Chapter 2** reports the genomic and phenotypic responses to this genomic selection.

This thesis

As Alexander von Humboldt said: "phenomena are only important in their relation to the whole" (von Humboldt & Bonpland 1814), which would also apply to for example the expression of a single gene in the grand scheme of mechanisms that underlie laying the first egg. It is, therefore, imperative to investigate how each component of the underlying mechanisms of seasonal timing respond to (changing) temperatures, together with how each 'endo-phenotype' contributes to this response, in order to gather knowledge and a deep understanding of the mechanisms underlying timing of breeding. Only then, is assessment of the adaptive potential to the impact of the intensified and novel selection pressures imposed by climate change in wild populations possible. As such, different scientific disciplines need to join forces to be able to ultimately determine whether species can and will adapt to climate change or not (but see Visser et al. 2010). Physiologists, evolutionary ecologists and molecular geneticists try to answer the same questions, but approach these from their own perspective and search for answers at different levels, as described above (and discussed in **Chapter 7**). In this thesis, we integrate the fields of evolutionary ecology and molecular genetics, but without fully ignoring the physiology, to address the overall aim of this thesis: **study the molecular basis of the physiological mechanisms underlying avian seasonal timing of breeding**.

In order to do so, we first created selection lines for timing of breeding in great tits, to be able to compare females that are (extremely) early breeders to females that are (extremely) late breeders. Here, great tits were bi-directionally selected for timing of breeding based on their genotypes, instead of their phenotypes, to avoid the risk of selecting on phenotypes that are heavily influenced by captive conditions. This is common practice in domestic animal breeding and agriculture, but not in wild populations. With the resulting *early* and *late* selection line, we were able to address the second aim of this thesis: evaluate both the phenotypic and genotypic response to artificial selection on timing of breeding. As such, I studied egg-laying dates under natural photoperiod and temperatures, the genomic response in terms of genomic estimated breeding values (GEBVs) and allele frequencies and possible correlated selection in two traits that could be associated with egg-laying date, which we present in **Chapter 2**.

In **Chapter 3**, we looked at whether temperature has a direct effect on timing of breeding, and subsequent plasticity in timing of breeding, *within* individual great tits (as opposed to Visser et al. (2009), where the effect was tested *between* individuals). This chapter described an experiment in both climate-controlled aviaries (i.e. artificial conditions) and outdoor aviaries (i.e. semi-natural conditions). In the climate-controlled aviaries, pairs of great tits were kept two breeding seasons under a natural photoperiod, but one season in a cold treatment and the other in a warm treatment, mimicking a cold and warm spring in the Netherlands, respectively. Great tit pairs housed in the outdoor aviaries were subjected to natural photoperiod and temperatures for two consecutive years. Because these pairs were from the F_3 generation of the breeding time selection lines (**Chapter 2**), we assessed whether selection on timing of breeding would result in correlated selection on phenotypic plasticity in timing of breeding. And, if so, whether selection on timing of breeding would result in a change in the average phenotype, the degree in plasticity or both.

The selection lines open up the possibility to work on the third aim: investigate whether individual differences in timing of breeding in females are reflected in differences in their molecular physiology and if so, where in the HPGL axis (Box 1) these differences occur. We housed individual F_1 generation females in the climate-controlled aviaries and subjected them to either the cold or the warm treatment as mentioned above. Within one year, these females went through two breeding seasons. In the first, we obtained egg-laying dates and in the second, groups of females were sacrificed at three different time points across the breeding season at which follicles were measured and brain, ovary and liver collected for subsequent candidate gene expression analysis (qPCR). We chose these time points based on reproductive behaviour in the first breeding season. We were able to link the two breeding seasons based on a correlation between egg-laying date and follicle size. The experiment was repeated with F_2 generation females. Within each tissue, for a tissue-specific set of candidate genes, expression levels were determined and analysed for differences between generations, time points and individual timing of breeding. The results are presented in **Chapter 5**. In **Chapter 4**, we measured genome-wide expression patters (RNA-seq) over time to investigate which genes and gene networks are involved in seasonal timing of breeding. We used the brain, ovary and liver samples from the same F_2 females

as in **Chapter 5**. However, in this study we did not look at individual females as in **Chapter 5**, but we analysed pools of three females, where each pool represents a unique combination of selection line, treatment and time point of sacrifice.

Gene expression, as investigated in **Chapter 4** and **Chapter 5**, is regulated by DNA methylation, a key epigenetic modification in genomes. It is currently largely unknown how DNA methylation influences or shapes complex traits such as timing of breeding, but slowly evidence starts to emerge. Further, the correlation between gene expression and DNA methylation is largely unknown. In **Chapter 6** we set out to assess whether (1) changes in DNA methylation in blood correlate to changes in DNA methylation in liver, (2) changes in DNA methylation correlates to changes in gene expression of candidate genes (qPCR data from **Chapter 5**) or genome-wide (RNA-seq data from **Chapter 4**) within liver and (3) changes in DNA methylation in blood correlate to changes in gene expression in the brain, ovary and liver.

Finally, in **Chapter 7**, I discuss how the different studies together give insight in, and improve our understanding of, the mechanisms underlying avian timing of breeding. Also, I discuss what is still unknown and thus, to where future studies should be directed.

CHAPTER 2

Genetic and phenotypic responses to genomic selection for timing of breeding in a wild songbird

Irene Verhagen, Phillip Gienapp, Veronika, N. Laine, Elizabeth M. van Grevenhof, A. Christa Mateman, Kees van Oers & Marcel E. Visser

ABSTRACT

The physiological mechanisms underlying avian seasonal timing of reproduction, a life history trait with major fitness consequences, are not well understood. Comparing individuals that have been selected to differ in their timing of breeding may prove to be a promising in studying these mechanisms, making selection lines a valuable tool. We created selection lines for early and late timing of breeding in great tits (*Parus major*) using genomic selection, i.e. selection based on multi-marker genotypes rather than on the phenotype. We took in nestlings (F_1 generation) from wild broods of which the mother was either an extremely early ('*early* line') or extremely late ('*late* line') breeder. These chicks were then genotyped and, based on their 'genomic breeding values' (GEBVs), we selected individuals for *early* and *late* line breeding pairs to produce the F_2 generation in captivity. The F_2 offspring was hand reared, genotyped and selected to produce an F_3 generation, which were then again genotyped and selected. This way we obtained egg-laying dates in aviaries for F_1, F_2 and F_3 birds. We studied the genetic response to the artificial selection and found increased genetic differentiation between the *early* and *late* reproducing selection lines over generations (F_1 to F_3), indicated by both diverging GEBVs and increased fixation indices (F_{ST}). We studied the phenotypic response to selection for birds breeding in outdoor breeding aviaries. We found that *early* line birds laid earlier than *late* line birds, and this difference increased over the generations (F_1 to F_3), with non-significant line effects for the F_1 and F_2, but highly significant line differences for the F_3. We also assessed whether there was correlated selection on two traits that are potentially part of the mechanisms underlying seasonal timing: the endogenous free running period of the day/night clock (tau) and basal metabolic rate (BMR), but found no correlated selection. We have successfully created selection lines on seasonal timing in a wild bird species and obtained an instrument for future studies to investigate the physiological mechanisms underlying timing of breeding, and the genetic variation in these mechanisms, an essential component for evolutionary change in timing of reproduction.

Functional Ecology (2019) ***33****(9): 1708-1721*

Introduction

The physiological mechanisms underlying avian seasonal timing of breeding, a life history trait with major fitness consequences, are not well understood. One reason for this is that timing of breeding, similar to many other life-history traits, is a 'complex trait' (Garland 2003), in the sense that it is the result of a cascade of physiological events (Visser et al. 2010; Williams 2012). While there is clear phenotypic variation in phenology among individuals within a population, it is unclear how these individuals differ in their underlying physiology. A better understanding of variation in timing of breeding and its physiology, and especially the variation which has a genetic basis, is important, because this determines the potential of timing of breeding to respond to selection, i.e. its 'evolutionary potential'.

Comparing individuals that differ in timing of breeding is a promising way to study these physiological mechanisms. Selection lines that create individuals that differ in their phenotype can be an especially potential powerful tool (Conner 2003). Selection lines have been used in evolutionary and physiological ecology to address a range of questions, such as the work on personalities in great tits (*Parus major*). Artificial phenotypic bi-directional selection on divergent levels of exploratory behaviour (Drent et al. 2003) and risk-taking behaviour (van Oers et al. 2004) showed correlated responses to selection, for example with systematic physiological changes on the neuro-endocrine stress axis. Faster exploring, more risk-averse individuals thereby had higher expression of mineralocorticoids, lower expression of glucocorticoid receptors in the brain and elevated plasma glucocorticoid levels (van Oers et al. 2011). Also, in the silver fox (*Vulpes vulpes*) long-term selective breeding in favour of, or against, aggressive behaviour resulted in selection lines with aggressive and tame responses to humans respectively (Wang et al. 2018). In guppies (*Poecilia reticulate*), divergent lines for large and small brain size showed that due to a negative genetic correlation with gut size, small brained fish need to trade-off relative brain size with feeding efficiency (Kotrschal et al. 2013) and immune function (Kotrschal et al. 2016). This is indicative of evolutionary trade-offs due to varying levels of predation (Reddon et al. 2018).

Here, we explore whether it is possible to create selection lines for early and late breeding in great tits from our long-term study population in the National Park de Hoge Veluwe (The Netherlands). Timing of egg-laying (i.e. the date the first egg is laid) is heritable (h^2 = 0.17) in this population (Gienapp et al. 2006) which means that we could expect a response to selection. We created an 'early' and a 'late' selection line for early and late timing of breeding using bi-directional genomic selection (Meuwissen et al. 2016) which is now commonly applied in domestic animal breeding and agriculture (Calus 2010; Jannink et al. 2010). In contrast to 'traditional' selection where individuals are selected based on their own phenotypes, 'genomic' selection selects individuals based on their 'genotypes'. In other words, selection is based on single nucleotide polymorphisms (SNPs), estimated as 'genomic breeding values' (GEBVs, see '*Materials and methods*' for details), rather than on their phenotypes. By selecting directly on GEBVs, we were able to select juvenile individuals who have not yet expressed the phenotype (egg-laying date), thereby speeding up the artificial selection. Additionally, we were able to select males who do not express the

phenotype at all. In general, genomic selection is more accurate, i.e. the expected and observed selection response show a higher correlation, compared to phenotypic selection (Meuwissen, Hayes & Goddard 2016; Wolc et al. 2015).

In this study we use the fixation index (F_{ST}, Holsinger & Weir 2009) to estimate the level of genetic differentiation and to detect the SNPs under selection between the early and late selection line. In other selection line studies it has been used successfully, for example in chicken, where the F_{ST} method detected regions with changes in allele frequencies (i.e. signatures of selection) between lines bred for either meat or eggs (Boschiero et al. 2018) and between three different lines of egg-layers (Heidaritabar et al. 2014). An additional 'sliding window analysis', where a window of a certain length slides along the genotypes, checks whether SNPs under selection cluster in certain genomic regions (Tajima 1991). After obtaining the regions under selection and the genes located there, we conduct a gene Ontology (GO) enrichment analysis to explore which functional groups (GO terms) are over-represented for a specific gene set (Gaudet & Dessimoz 2017; Primmer et al. 2013). In GO databases, the genes are assigned to predefined functional groups. In addition to the GO databases, the KEGG (Kyoto Encyclopedia of Genes and Genomes) is a database collection that links genomic information with higher order functional information, i.e. cellular processes and pathways (Kanehisa & Goto 2000).

Our selection line experiment also allowed us to estimate correlated selection responses in two physiological traits potentially related to egg-laying date: the endogenous free running period (tau), which is the period of time it takes for an organism's endogenous rhythm to repeat in artificial constant conditions, and basal metabolic rate (BMR). We could thereby test whether these traits were genetically correlated with egg-laying date and are hence a potentially heritable part of the underlying cascade of egg-laying date.

Tau underlies circadian rhythms, the way physiology and behaviour varies with daily changes in the environment, and is therefore an important prerequisite for successful breeding and survival (Dawson et al. 2001). The internal clock regulates the expression of a panel of 'clock' genes and tau lies close to, but still differs significantly from, 24h (Pittendrigh & Daan 1976; Reppert & Wever 2002). Tau has been shown to vary and to be highly heritable ($h^2 = 0.86 \pm 0.24$) in great tits (Helm & Visser 2010), suggesting that evolutionary changes in tau are, in theory, possible. The role of tau in circannual rhythms (such as seasonal timing) is, however, unclear. Although some studies did not find a link between circadian and circannual rhythms (Agarwal et al. 2017; Budki et al. 2014), others have found evidence that they are linked (Gwinner 1986; Myung et al. 2015). Therefore, we chose to investigate a possible response in tau in our breeding time selection lines.

Timing of breeding might be constrained by high energetic demands for egg-laying (te Marvelde et al. 2012; Monaghan & Nager 1997) as resources are scarce and temperatures low early in the season. Daily energy expenditure did not differ between wild early and late breeding females during egg-laying (te Marvelde et al. 2012) and females selected for early breeding may thus have a lower BMR, i.e. leaving more energy to produce eggs. As BMR

is heritable in birds (Nilsson et al. 2009; Tieleman et al. 2009) it could potentially respond to selection on timing of breeding if these traits are genetically correlated.

We evaluated the response to bi-directional artificial selection on egg-laying dates by (1) studying egg-laying dates (the phenotypic response) in aviaries under natural day length and temperatures, (2) studying the genomic response to selection by using the fixation index (F_{ST}) as a measure of genetic diversity between the early and late breeding birds, and (3) the phenotypic response to selection in traits potentially associated with egg-laying date: tau and BMR.

Materials and methods

SELECTION LINES

Obtaining the F1 generation

We created two selection lines: an early line that we selected for early egg-laying and a late line that we selected for late egg-laying. In the spring of 2014, 28 pairs from our long-term study population (Hoge Veluwe, the Netherlands) were selected as the 'parental' (P) generation based on their breeding values, estimated using the pedigree of the wild population (Figure S2.1). To calculate these pedigree-based breeding values we used the following animal model (Lynch & Walsh 1998) where $y_{i,j}$ is

$$y_{i,j} = \mu + age_{i,j} + year_j + ind_i + a_i + \varepsilon$$

the phenotype of individual *i* in year *j*, μ is the population mean, $age_{i,j}$ is the fixed effect of age ('first year breeder' vs. 'older') of individual *i* in year *j*, $year_j$ is the fixed effect of year *j* (to account for differences among year driven by phenotypic plasticity), ind_i is the random non-genetic effect (also called 'permanent environment effect') of individual *i*, a_i is the additive genetic effect of individual *i*, estimated from the pedigree, and ε the error term. We included all records of females breeding from 1973 to 2014 in the Hoge Veluwe study population into our analysis (Gienapp et al. 2006; Husby et al. 2010; Ramakers et al. 2019). Parents (except for two males) were identified and blood sampled for later DNA extraction (see Gienapp et al. 2017 for details) and genotyping (see '*Calculating the GEBV*' below).

From the parental generation, we brought all nestlings (F_1 generation) into the aviary-facilities at the NIOO-KNAW at 10 days post-hatching (Figure S2.2). Nestlings were ringed for identification, weighed and further hand-raised at the NIOO-KNAW (see Drent et al. 2003 for details). These chicks were then genotyped and based on their GEBVs we selected individuals for early and late line breeding pairs to produce the next generation in captivity.

Obtaining the F2 and F3 generation

All eggs laid by the F_1 generation (and the subsequent F_2 generation) were transferred to wild nests, were they were incubated by foster parents and hatched. These chicks were brought into the aviary-facilities at the NIOO-KNAW at 10 days post-hatching for further hand raising (208 chicks from 37 F_1 pairs and 300 chicks from 33 F_2 pairs). When birds reached the independent stage (approximately 30 days post-hatching), a blood sample was taken for DNA extraction and genotyping. The F_2 offspring were genotyped and selected to produce an F_3 generation, which was then genotyped and selected. After moult birds were temporarily housed in single-sex groups of 7 (male) or 8 (female) birds in outdoor aviaries (4.2 × 1.9 × 2.1m) under natural light conditions. There, birds were fed ad libitum and had water available for drinking and bathing.

Calculating the GEBV

To be able to select individuals without obtaining their phenotypes we predicted 'genomic breeding values' (GEBVs) for each individual in the selection lines with the 'genomic best linear unbiased prediction' (GBLUP) approach (Clark & van der Werf 2013). This approach uses genomic markers to calculate pairwise relatedness among all individuals, i.e. those in the training population and the selection candidates who may not have phenotypes. The genomic relatedness matrix (GRM) obtained in this way is then used to replace the pedigree-derived relatedness matrix in a standard animal model. The predicted breeding values, i.e. BLUPs for the additive genetic effect, from a model with a GRM are then the GEBVs. In short (but see Gienapp et al. 2019 for methodological details), we genotyped 2045 great tit females that bred (between 1995 and 2015) in our study populations that had recorded egg-laying dates. These were used as training population. These individuals, as well as all F_1 and F_2 individuals, were genotyped on a 650K SNP chip (Kim et al. 2018) to predict their GEBVs. We excluded SNPs from the Z chromosome and unassigned scaffolds (32 467 SNPs). Individuals with more than 5% of missing genotypes and SNPs with a call rate below 95% were discarded resulting in 665 individuals and 437 271 SNPs. The selected F_1 and F_2 individuals had egg-laying dates recorded in the aviaries but these were not used when predicting their GEBVS. To predict the GEBVs we first corrected the phenotypes of the individuals in the training population for year and area effects (based on all birds that bred in those years). Fitting these year and area effects directly in the GBLUP model would have led to biased estimates for these effects. As then, the year and area corrections would only been done on the (sometimes very limited number of) genotyped individuals in a year/area combination. We therefore used the complete data set to estimate area- and year-effects with the following model:

$$y_{i,j} = \mu + yr_j + ar_a + age_i + ind_i + \varepsilon$$

with $y_{i,j}$ being the phenotype of individual i in year j, μ the overall intercept, yr_j and ar_a the fixed effects for year (as factor) and area, respectively, age_i the age of individual i (as factor, 1st year breeder vs. older) and ind_i the random effect of individual i. Area refers to four different 'study populations' that are all part of the large and more or less continuous woodland area on the Veluwe near Arnhem and show small but consistent differences in timing of egg-laying. That these areas are referred to as different 'study populations' is more owing to 'historical' reasons as they all lie within five km of each other. We then fitted the following animal model, in which the pedigree-derived relatedness matrix was replaced by the GRM:

$$y'_{i,j} = \mu + age_i + pe_i + a_i + \varepsilon$$

with $y'_{i,j}$ being the pre-corrected phenotype of individual i in year j, i.e. . $y'_{i,j} = y_{i,j} - \widehat{yr}_j - \widehat{ar}_a$, and a_i the random additive genetic effect of individual i.

Selection procedure

GEBVs were calculated for all offspring alive in December of the year they were born. These selection candidates were ordered according to their GEBVs and suitable pairs (n = 20) for creating the next generation were made starting with the most extreme individuals (Figure S2.2). To maintain as much of the initial genetic variation as possible we tried to select within rather than among families by including offspring from each breeding pair (from the previous generation) but maximal two siblings (of each sex) in the selected individuals. The criterion on the maximum number of selected siblings was relaxed if necessary. For example, when there was an insufficient number of (fe)males from one family with extreme GEBVs, a (fe)male from another family with similar GEBVs already sufficiently (n = 2 individuals per sex) represented in the selected population would be supplemented in order to keep the GEBVs as extreme as possible. We also paired the selected individuals dis-assortatively to maintain genetic variation. To prevent inbreeding we never paired siblings.

The expected phenotypic response to genomic selection was calculated as the standardised selection differential on GEBVs multiplied with the accuracy (0.21) of the GEBVs (Gienapp et al. 2019). This gives the expected response in standard deviations of the trait. The standardised selection differential on GEBVs was calculated – analogous to the phenotypic case – as the difference between the means of the unselected population and the selected individuals divided by the standard deviation of the GEBVs in the population prior to selection (Lynch & Walsh 1998). Since not all males of breeding pairs from the P-generation were genotyped, we assumed random mating with regard to egg-laying date and therefore halved the calculated selection differential for this generation, as it was based on only females.

HOUSING CONDITIONS AND EGG-LAYING DATES

From January (2015-2017) onwards, the breeding pairs (n = 120; 40 aviaries times three years, for F_1, F_2 and F_3) were housed in 40 outdoor aviaries (4.2 × 1.9 × 2.1m) where the birds were subjected to natural photoperiod length and temperatures. From the 20th of February onwards, the birds received daily additional light from a single full spectrum daylight fluorescent lamp (58W, 5500K, True-light, The Netherlands) per aviary. Lights went on 2.5hrs before sunrise until 12:00 PM in order to synchronize their breeding with the wild population which fostered the eggs laid in captivity. The time of lights on changed daily with sunrise, but never earlier than 02:00 AM. We made a distinction between the North and South side of the aviary building as the latter experienced a different environment throughout the breeding season because of the daily rotation of the sun. Temperatures were recorded every 10-30 minutes using loggers (Thermochron iButton).

Nest boxes in the aviaries were checked daily for eggs. Eggs were collected and replaced by dummy eggs. When a female had incubated a complete artificial clutch for five days, it was removed and the female was allowed to relay. Egg-laying date is recorded as the day the first egg of the first clutch was laid.

IDENTIFYING LOCI DIFFERENTIATED BETWEEN SELECTION LINES

To quantify the level of genetic differentiation between the early and late line, we estimated F_{ST} (Holsinger & Weir 2009) for each SNP in every generation using a custom made Affymetrix great tit 650K SNP chip (Kim et al., 2018). We used the same SNPs as for the estimation of GEBVs (see '*Calculating the GEBV*' above). We calculated F_{ST} values for individual SNPs between early and late line using the program PLINK 1.9 (Purcell et al. 2007). In order to see if highly differentiated SNPs cluster to certain genomic regions we also used sliding window F_{ST} calculation. For this we used vcftools 0.1.14 (Danecek et al. 2011) with --fst-window-size 200000 and --fst-window-step 50000. To distinguish SNPs under selection from genetic differences between the lines due to drift, we used Arlequin version 3.5.2.2 (Excoffier & Lischer 2010), which uses coalescent simulations to get the p-values of locus-specific F-statistics conditioned on observed levels of heterozygosities (Excoffier et al. 2009). We used all the generations in the model (P-F_1-F_2-F_3) and two lines (early and late) and the run was conducted with default values, except that we increased the number of simulations and demes (100 000 simulations and 1000 demes). For the SNPs under selection we also calculated observed versus expected heterozygosities. A gene ontology (GO) enrichment analysis was done on genes linked to SNPs with F_{ST} p-values <0.01 in order to see the functional relatedness of GO terms and genes under selection. The significant SNPs could be assigned to 1743 great tit genes (NCBI Parus major genome version 1.1, GCA_001522545.2, annotation release ID: 101). Functional relatedness of GO terms was done using the Cytoscape plugin ClueGo 2.5.1 (Bindea et al. 2009). ClueGo constructs and compares networks of functionally related GO terms with kappa statistics. A two-sided hypergeometric test (enrichment/depletion) (Rivals et al. 2007) was applied with GO term fusion, network specificity and Kappa score were kept at default values and

false discovery correction was carried out using the Bonferroni step-down method. We used both human (30.08.2018) and chicken (21.09.2018) gene ontologies and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa & Goto 2000) pathway database for comparison due to differences in GO annotations.

PHENOTYPIC CHANGES ASSOCIATED WITH SELECTION - CORRELATED RESPONSE

Endogenous free running period length (tau)

Following the breeding season of 2015 and 2016 (Figure S2.3), 167 birds (F_1 = 66, F_2 = 101) were transferred in autumn to individual cages distributed over three rooms. The cages were equipped with two wooden perches, with one being connected to a computer to register perch-hopping activity (software developed by T&M Automation, Leidschendam, The Netherlands). All birds were entrained to a L:D schedule for six days with a 1000 lux light (18W Havells Sylvana Activa 172, East Sussex, UK) at perch level. Length of photoperiod was based on the amount of natural daylight on the day the experiment began. Subsequently, the light was turned off and birds received constant dim light with an intensity of 0.5 lux at perch level for 14 days, during which we measured the length of time it takes for an individual's endogenous rhythm to repeat in constant conditions, i.e. the endogenous free running period length (tau). White noise was played continuously to mask neighbouring vocalizations and activity to prevent birds influencing each other.

Basal metabolic rate (BMR)

After juveniles had completed moult, but before they were paired for the breeding season (Figure S2.2), BMR from 620 individuals (F_1 = 163, F_2 = 181, F_3 = 276) was measured in autumn and winter of 2014-2016. Birds were caught from their outdoor aviaries around 17:00 PM and transferred to a respiratory chamber within an open-circuit respirometer (see Amo et al. 2011 and Caro & Visser 2009 for details). In short, oxygen consumption (ml O2 min^{-1}) was calculated as the difference in oxygen concentration between air from the respirometer chambers and reference air from an empty chamber. Metabolic rate (kJ 24 h^{-1}) was calculated by converting oxygen consumption, assuming an energetic equivalence of 20 kJ per liter of O2 (Weir 1949). Birds were weighed after overnight measurement (i.e. morning mass), before being transferred back to their outdoor aviaries the next morning.

STATISTICAL ANALYSES

All analyses were performed in R (version 3.3.1) and animal models were run using ASReml-R (Butler et al. 2009). Effects were considered significant when $p < 0.05$. For the analysis of egg-laying dates, we fitted linear models. We then followed backward elimination of the model and used analysis of variance (ANOVA) to test for the effects of aviary side, selection line, generation and their interactions. We included aviary side because half of the

aviaries faced North and the other half South (later referred to as 'North side' and 'South side' respectively). The two sides experienced a different environment throughout the breeding season because of the daily rotation of the sun. Differences in mean daily temperatures were tested from 16 March – 20 April, following (Visser et al. 2006), by performing a t-test per year.

To calculate tau (in hours) from activity data we used Chronoshop 1.1 (Spoelstra et al. 2018). Actograms were plotted in Chronoshop based on the Lomb-Scargle algorithm (Ruf 1999). We excluded 72 individuals from analyses due to low individual activity, technical errors or because they were remaining birds from the F_3 generation, but their number was too low (n = 19) for proper statistical testing. This allowed data analysis of 115 birds for the F_1 and F_2 generation. Chronoshop distinguishes between qualitative (activity or no activity) and quantitative (different activity levels or no activity) data. We have chosen to analyse the quantitative data, as these contain more information. We performed Mantel tests (Mantel 1967) from the 'ade4' package to test for a possible influence of neighbouring birds. No neighbouring effects on the onset of activity were detected (p-values ≥ 0.06, Table S2.1). Fitting a linear model, we determined effects of generation, selection line, sex and their interactions by backward elimination and ANOVA for model selection. A Tukey test was performed for post hoc analysis.

Basal metabolic rates were analysed with linear mixed effect models using the 'lme4' package (Bates et al. 2015). Sex, selection line, generation, morning mass and temperature at 17:00 PM were fixed effects and respiratory chamber and date of BMR measurement as random effects. We then followed backward elimination of the model based on the F-test with Kenward-Roger approximation from the KRmodcomp function in the 'pbkrtest' package in R (Halekoh & Højsgaard 2014). The heritability of BMR was estimated as done previously with the heritability of tau in another study (Laine et al. 2019). In short, an initial mixed linear model was fitted with the 'ASReml-R' package to remove non-significant effects based on the Wald F test ($p<0.05$). Subsequently, an animal model (Henderson, 1986; Kruuk, 2004), using the GRM of the individuals (see '*Calculating the GEBV*' above), was fitted for BMR with fixed factors sex, selection line and their interaction, as well as morning mass and outside temperature at 17:00 PM.

Results

GENOMIC CHARACTERISTICS OF THE SELECTION LINES

GEBVs differed significantly between the two selection lines (Figure 2.1, $F_{1, 235} = 1428$, $p < 0.001$) and diverged significantly over the generations (interaction line × generation: $F_{1, 234} = 27.1$, $p < 0.001$, Figure 2.1). When analysing the selection lines separately, GEBVs decreased significantly over the generations in the early selection line ($b = -0.079 \pm 0.018$, $F_{1, 117} = 18.7$, $p < 0.001$) and the GEBVs increased significantly in the late selection line ($b = 0.112 \pm 0.032$, $F_{1, 116} = 12.3$, $p < 0.001$). For both the early and late selection line

standardised selection differentials were moderately strong (Kingsolver et al. 2001) and in opposite directions (Table 2.1).

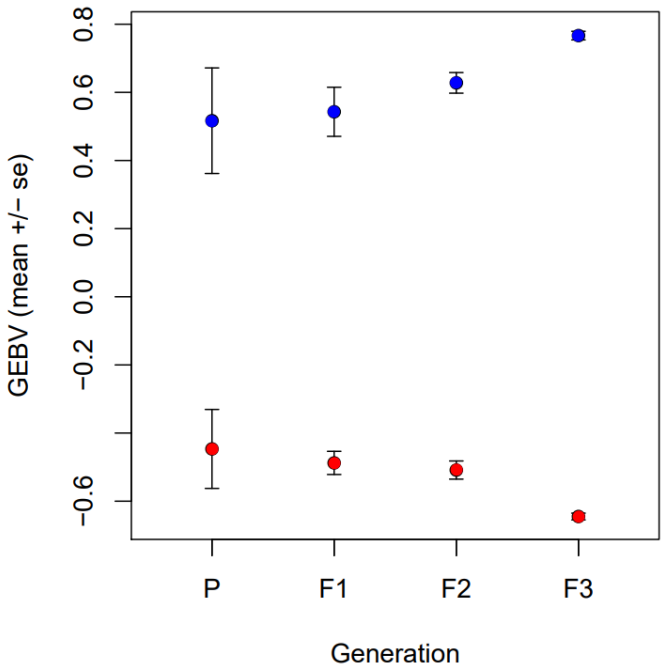


Figure 2.1. Change in GEBVs over generations in the selected individuals. GEBVs decreased significantly for the early selection line (red) and increased significantly for the late selection line (blue).

Table 2.1. Selection differentials on GEBVs within generations and the cumulative selection differentials over generations. Here, the selection differential is the difference in the trait between the unselected population and the selected individuals.

Selection line	*P*	*F_1*	*F_2*	*Cumulative*
Early	-0.445	-0.134	-0.14	-0.719
Late	0.49	0.227	0.126	0.844

IDENTIFYING LOCI DIFFERENTIATED BETWEEN SELECTION LINES

The F_{ST}-values between early and late lines ranged from -0.003 to 0.432 (Figure 2.2, Table S2.2). In the sliding window analysis clear peaks formed especially in chromosome 4A, but due to low linkage disequilibrium in the great tit genome (van Bers et al., 2012) the F_{ST}-values were lower in the sliding window setting (Figure S2.5). When distinguishing drift from selection altogether 4786 SNPs showed a significant signal of selection ($p < 0.01$), which showed increasing F_{ST}-values between lines over generations (Figure 2.3, Table S2.3). These SNPs covered 1753 (1743 unique) great tit genes (Table S2.4) of which 1525 and 1472 are also found in human and chicken GO-databases, respectively. When using the

human GO-database we found 204 significant GO terms associated with the genes under selection (Table S2.5). When using the chicken GO-database 126 significant GO-terms were found (Table S2.6). From the GO-terms, 95 were shared by both database results.

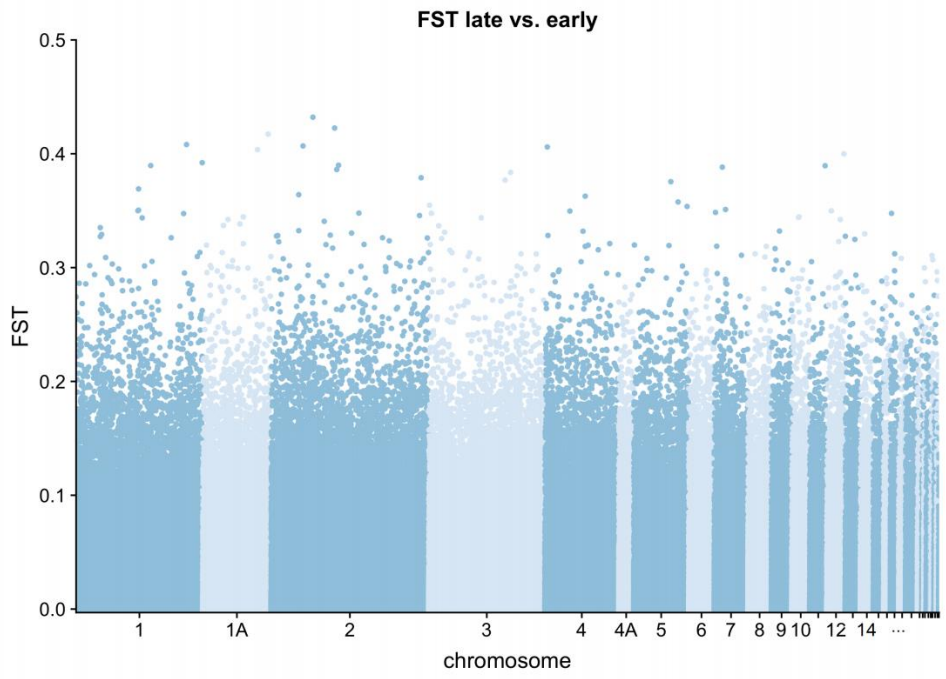


Figure 2.2. Manhattan plot showing on the y-axis the F_{ST}-values (ranging from -0.003 to 0.432) estimated for each of the SNPs between the early and late selection line in the genome (along the x-axis) including all generations. The colours are to distinguish chromosomes from each other.

RESPONSE TO SELECTION

The cumulative predicted response to genomic selection (i.e. the sum of the selection differentials) was -0.72 days for the early line and 0.84 days for the late line (Table 2.1). Assuming the unselected parental generation, i.e. the breeding population on the Hoge Veluwe in 2014, has an average GEBV of 0, we can compare the cumulative selection response to the mean GEBV in the F_3 generation. The average GEBV of the F_3 individuals of the early line was -0.50, while it was 0.61 for the late line. This corresponds reasonably well to the expected cumulative responses. Please note that the response to genomic selection cannot be directly compared with the phenotypic divergence in the egg-laying dates reported below because GEBVs are for egg-laying dates in the wild, while the phenotypes are for a different trait, namely egg-laying date in the aviary. For example, egg-laying dates in the aviaries have a considerably higher heritability (0.42 ± 0.22, LRT: χ^2 = 5.56, df = 1, p = 0.02) than 'wild' egg-laying dates (0.17, Gienapp et al. 2006).

Over the three years, 14 out of 120 females were excluded for the analysis on egg-laying dates as they did not initiate egg-laying. The females from the early line laid on average 6.2 ± 2 days (mean ± SE) earlier compared to late line (p = 0.003, Table 2.2, Figure 2.4). Although there is no difference in egg-laying dates between the selection lines at the F_1

generation (early = 13.9 ± 2.5, late = 15.9 ± 3.6, $t_{(27.8)}$ = -0.55, p = 0.585), nor at the F_2 generation (early = 14.0 ± 2.7, late = 20.0 ± 3.7, $t_{(28.4)}$ = -1.59, p = 0.123), egg-laying dates did differ significantly between the selection lines at the F_3 generation (early = 12.6 ± 2.4, late = 22.2 ± 3.4, $t_{(35)}$ = -2.82, p = 0.008). This response to selection in egg-laying dates was of a larger magnitude compared to the response in GEBVs (Figure 2.1; see also '*Discussion*').

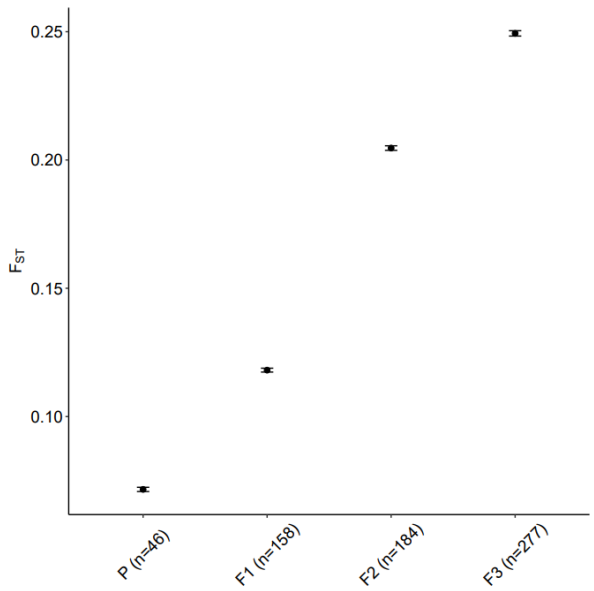


Figure 2.3. F_{ST} values between the selection lines over generations for the 4786 SNPs showing a significant signal of selection, i.e. SNPs that show an increase in F_{ST} over generations (generated from Table S2.3).

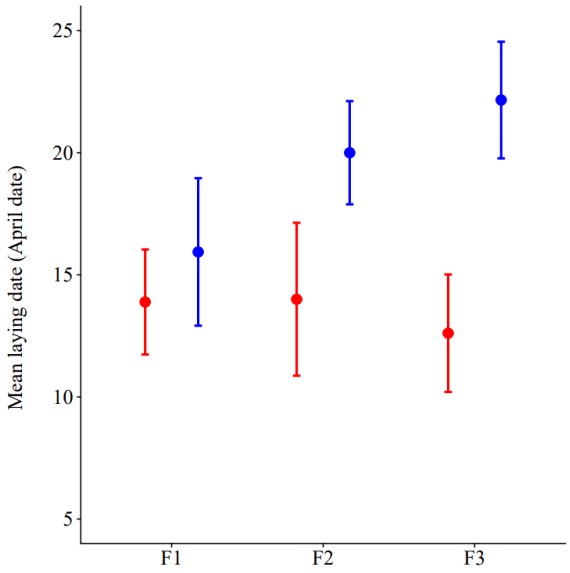


Figure 2.4. Mean laying dates (mean ± SEM) in April dates (01-04 is 1, 02-04 is 2 etc.), from selection line females in outdoor aviaries for the three generations. The mean laying dates per generation for the *early* selection line females (F_1 = 18, F_2 = 17, F_3= 18) are represented in red and the mean laying dates per generation for the *late* selection females (F_1 = 16, F_2 = 18, F_3= 19) are represented in blue.

CORRELATED RESPONSES IN TAU AND BMR

Endogenous free running period length (tau)

The heritability of tau was previously estimated to be h^2 = 0.48 ± 0.22 (Laine et al. 2019). We found no correlated response to selection on egg-laying date of tau (Table S2.7, Figure 2.5); selection lines did not differ (p = 0.211) in tau. In addition, no differences (p = 0.246) between generations were observed. There was a significant difference (p = 0.002) in tau between sexes, where males showed a slightly longer tau (5.82 ± 1.86 min) compared to females (Table S2.7, Figure 2.5).

Basal metabolic rate (BMR)

The heritability of BMR was estimated to be $h^2 = 0.08 \pm 0.08$. Sex, line, generation and line × generation, did not affect BMR (Table S2.8, Fig. 6), while correcting for morning mass ($p < 0.001$) and temperature outside at 17:00 PM ($p < 0.001$, Table S2.8, Figure 2.6). In addition, the random effects 'respirometer channel' ($\chi^2 = 3.4905$, df = 1, $p = 0.062$) and 'date of BMR measurement' ($\chi^2 = 2.9326$, df = 1, $p = 0.087$) did not explain the variation in BMR.

Table 2.2. Estimated parameters of the linear models investigating the degree of variation in egg-laying dates (n = 106) explained by selection line, generation, the side (North or South) of the aviary building and their interactions. All parameters were fixed effects in the model and statistics are given for the point of exclusion from the model. Bold p-values indicate significance.

Parameter	*Estimate*	*s.e.*	*F-ratio*	*Df, ndf*	*p-value*
Line × generation × side			0.84	2, 94	0.434
Late × F_2 × *South*	-3.048	10.112			
Late × F_3 × *South*	-12.325	9.976			
Generation × side			0.37	2, 96	0.689
F_2 × *South*	4.133	5.045			
F_3 × *South*	0.925	4.977			
Line (*late*) × side (*South*)	3.496	4.038	0.75	1, 98	0.389
Line × generation			1.10	2, 99	0.336
Late × F_2	3.469	4.999			
Late × F_3	7.308	4.930			
Generation			0.49	2, 101	0.614
F_2	1.887	2.500			
F_3	2.316	2.466			
Side (*South*)	-5.314	2.007	7.00	1, 103	**0.009**
Line (*late*)	6.157	2.006	9.42	1, 103	**0.003**
Intercept	16.198	1.748			**<0.001**

Discussion

We found genetic and phenotypic responses in timing of breeding to bi-directional artificial selection using genomic selection. Selection significantly decreased and increased GEBVs for the early and late selection lines respectively and we found increasing F_{ST}-values between selection lines over the course of three generations. In addition, we found a phenotypic response to genomic selection in egg-laying date where early line females in the outdoor aviaries laid about six days earlier compared to late line females. In the wild, six days is roughly one standard deviation of the within year variation in egg-laying date (Gienapp et al. 2006). We did not find a correlated response to selection for tau or BMR.

UNDERLYING GENOMIC CHANGES

Our genomic selection led to genomic differentiation between the selection lines, measured by F_{ST}. The genes in genomic areas where the F_{ST} is more strongly differentiated between the lines than expected by drift are linked to neuronal and developmental related GO groups (Tables S2.5 and S2.6). One of the most significant genes, the adipokine angiopoietin 1 (*ANGPT1*, Table S2.3), is an important gene in angiogenesis (Koh 2013) and has been shown to be part of follicular development in rats (Rudolph et al. 2016), pregnancy complications in humans (Andraweera et al. 2012) and may have a possible role in avian reproduction (Bornelöv et al. 2018). Another highly significant gene, glycerol-3-phosphate dehydrogenase 1 like (*GPD1L*, Table S2.3), has been shown to be linked to adaptive responses to temperature in chicken thyroids (S. Xie et al. 2018). Also zona pellucida glycoprotein 4 (*ZP4*) differentiated significantly between the selection lines. This gene codes for glycoproteins which constitute the avian perivitelline layer and plays various roles (e.g. oocyte protection) in reproductive functioning (Serizawa et al. 2011).

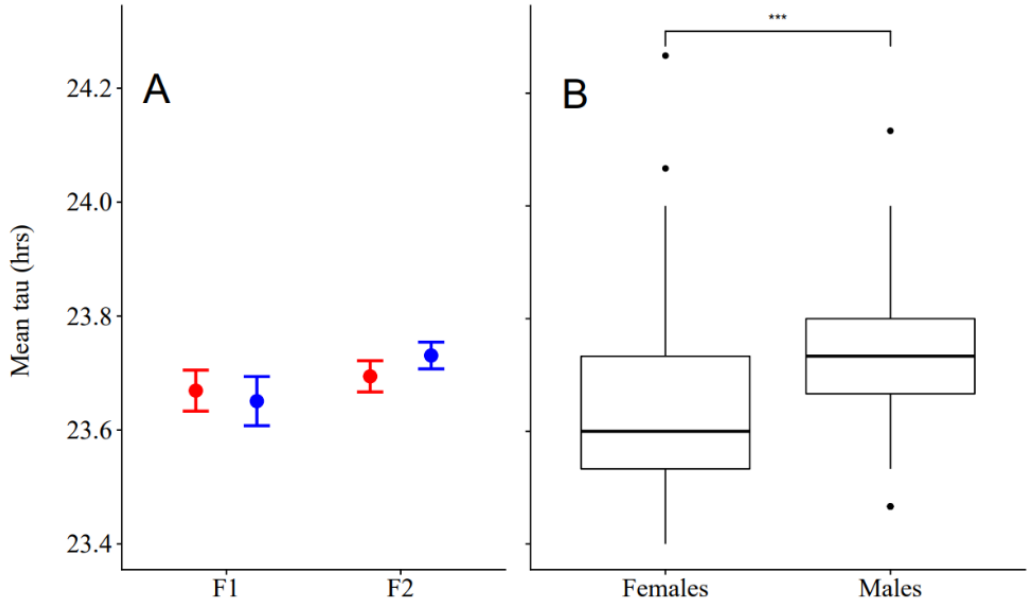


Figure 2.5. In panel A tau in hours (mean ± s.e.m.) is shown for all individuals (N = 115) over two generations. The mean individual tau per generation is represented in red and blue for the early and late selection line respectively. For panel A, we adjusted the horizontal position of the data shown in panel A to prevent overlap and so facilitate clarity of the graph. Panel B visualizes the significant effect (p = 0.002, Table S2.7) found between sexes ($N_{females}$ = 60, N_{males} = 55). For both panels the y-axis of panel A applies.

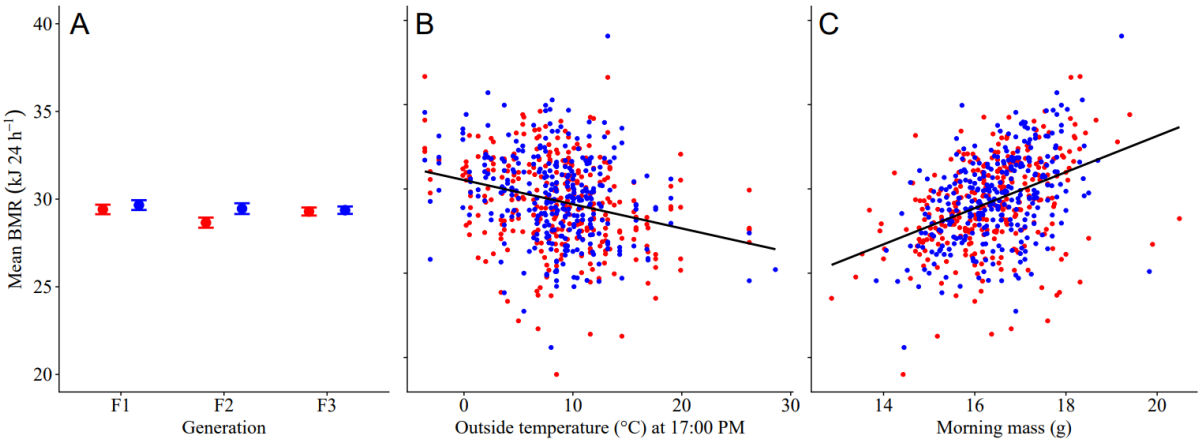


Figure 2.6. In panel A basal metabolic rates (mean ± s.e.m.) are shown for all individuals (N = 620) over three generations. For this panel, we adjusted the horizontal position of the data shown to prevent overlap and so facilitate clarity of the graph. The mean (panel A) and individual (panel B and C) BMRs are represented in red and blue for the *early* and *late* selection line, respectively. Panel B visualizes the significant effect found for outside temperature (in degrees Celsius) at 17:00 on BMR ($F_{1, 275.8} = 27.92$, $p < 0.001$, Table S2.8) and panel C the significant effect of morning mass (in grams) on BMR ($F_{1, 550.2} = 120.33$, $p < 0.001$, Table S2.8). For all three panels the y-axis of panel A applies.

In the regulation of the hypothalamic-pituitary-gonadal-liver axis (HPGL axis), i.e. the physiological mechanism underlying seasonal breeding in birds, both the circadian clock and thyroid play critical roles (Dawson et al. 2001; Nakane & Yoshimura 2014). Although no response in tau was observed, we found increased genetic differentiation between the selection lines in the allele frequencies of genes related to a circadian entrainment KEGG pathways (Table S2.5). Interestingly, in another study, we did not find a difference in onset of activity (as a measure of entrainment) in a L:D cycle between the selection lines nor a clear genomic signal underlying the variation in the circadian traits studied (Laine et al. 2019). One should note, however, that the genes investigated in Laine et al. (2019) differ from the genes in the circadian entrainment KEGG pathway, where the latter may thus give potential insights in genetic variation underlying phenotypic variation in circadian traits. Furthermore, neuronal and especially glutamate related GO groups and KEGG pathways are known to be important in learning (Audet et al. 2018), behaviour (X. Wang et al. 2018) and in the HPGL axis especially during reproduction (Maffucci & Gore 2009; Neal-Perry & Santoro 2006; Zhang et al. 2016).

RESPONSE TO SELECTION

The observed average difference in egg-laying dates between the selection lines was about six days, which is considerably larger than the difference of about one day in GEBVs. It is not completely clear why there is such a difference in the magnitude of the response but one potential explanation is that this is due to regression to the mean. A characteristic of GEBVs is that their mean is zero and their variance is set by their accuracy, which is ~0.2 in this study. The lower the accuracy, the smaller the variance around the mean and the

closer the GEBVs will be 'pulled' towards the mean (i.e. zero). Therefore, the selection lines are showing a smaller response in GEBVs (Figure 2.1) than in egg-laying dates (Figure 2.4), for which this problem does not occur and as such can be interpreted as the realized response to selection. An additional reason may be that the first selection step (from the parental to the F_1 generation) was based on pedigree-based breeding values and not on GEBVs, but when we calculated the cumulative response based on GEBVs we also included this selection step. Not all males of the pairs from the P-generation we selected were genotyped at the time of selection. Therefore, we had to assume that breeding pairs were mated randomly with regard to their GEBVs, which may not have been the case and could have influenced our expected genomic selection response. Lastly, the aviary egg-laying dates may not be exactly the same trait as wild egg-laying dates (see below). As a result of these three possibilities, a comparison between genomic selection response and phenotypic divergence between the selection lines in the aviaries is difficult to interpret.

Egg-laying dates in captivity differ from egg-laying dates in the wild as discussed in detail in another study (Visser et al. 2009), where captive females initiate egg-laying on average later compared to wild females. A probability for this delay in egg-laying, among others, could be that females lack specific cues (Lambrechts et al. 1999) or experience a disrupted correlation between predictive cues (Bentley et al. 1998) in (semi-)artificial conditions. We counteracted this successfully by increasing photoperiod with 2.5 hrs of extra light (see '*Housing conditions and egg-laying dates*' above), but we could question whether aviary egg-laying dates obtained in this study are comparable to those in the wild? Egg-laying date is the outcome of a neuro-endocrine cascade along the HPGL axis, which in birds starts months previously with the onset of gonadal growth triggered by increasing photoperiod (Ball & Balthazart 2002; Dawson et al. 2001). As the change in photoperiod is highly predictable every year, it cannot account for the variation in egg-laying dates across years or females. Females make use of supplementary cues (e.g. temperature, social cues) to fine-tune the onset of egg-laying (Dawson 2008) and thus, variation in egg-laying date can be caused by individual variation in underlying processes (Visser et al. 2010). Increasing photoperiod artificially, therefore, has a direct effect on the activation of the HPGL axis, but not on egg-laying itself. Unfortunately, the selection line female great tits have no wild egg-laying dates, and so we were unable to test for a within female correlation between wild and captive egg-laying dates. There are also too few aviary birds with relatives in the wild for which we have egg-laying dates and hence an analysis on genetic covariation of wild versus aviary egg-laying dates has too little power. There is, however, evidence that there is a strong correlation between egg-laying dates from wild great tits having initiated egg-laying both in the wild and in captivity (Visser et al. 2009).

The F_1-F_3 generation birds in this study were kept in outdoor aviaries but their siblings, with less extreme GEBVs, were kept in climate controlled aviaries, as done previously with great tits (Caro & Visser 2009; Schaper 2012; Visser et al. 2009). In these aviaries, birds were subjected to two contrasting temperature treatments mimicking an extremely cold and extremely warm spring in the Netherlands. Egg-laying dates of these females will be

published elsewhere (**Chapter 3 and 4**). It is however worth mentioning that the egg-laying dates in these breeding pairs did not differ between the early and late selection lines, nor that an interaction was found between selection line and temperature treatment. This lack of effect of selection line on egg-laying date is in contrast with the findings in this study where birds were kept in outdoor aviaries. Possible reasons are pointed out above, and the birds in climate controlled aviaries might experience an even higher reduction in environmental variability. The lack of an effect could potentially also be caused by the less extreme GEBVs these birds have or stress caused by the artificial environment (Caro et al. 2007). These contrasting results highlight how important and complex the influence of environmental cues on (complex) traits is.

CORRELATED RESPONSES IN TAU AND BMR

As avian timing of breeding is a complex trait, i.e. determined by many genes, it is likely that pleiotropy or linkage disequilibrium could generate genetic covariation among traits. As tau and BMR are possibly related to timing of breeding (Helm & Visser 2010; Nilsson & Nilsson 2016; Tieleman et al. 2009), selection on these traits could potentially affect timing of breeding, i.e. selection on correlated traits (Lande & Arnold 1983; Merilä, Sheldon & Kruuk 2001). Heritability of tau was high ($h^2 = 0.48 \pm 0.22$) (Laine et al. 2019), which complies with a previous study on captive great tits (Helm & Visser, 2010). However, we found no correlated directional response to genomic selection in BMR and tau (Table S2.7 and S2.8). In addition, the study by Laine et al. (2019), did not find a response to selection in other circadian activity rhythm parameters (i.e. phase onset and phase shift).

Though BMR has shown to be highly heritable in wild populations of blue tits (*Cyanistes caeruleus*) (Nilsson et al. 2009) and pied flycatchers (*Ficedula hypoleuca*) (Bushuev et al. 2012), no significant heritability was found in our selection lines ($h^2 = 0.08 \pm 0.08$). This minimizes, if not prevents, a response of BMR to genomic selection and not finding a correlated response in this study is therefore not surprising. In addition, in a study in great tits, winter to breeding season repeatability of BMR was shown to be close to zero, suggesting that winter and spring BMR are two unrelated traits (Bouwhuis et al. 2011). Therefore, autumn/winter BMR, the time that we measured BMR, might not be affected when selecting for timing of breeding and could make a possible correlation with egg-laying date impossible. This could also be the reason for tau not to correlate with egg-laying date, despite being highly heritable, as tau was measured in the autumn/winter as well. Unfortunately, studies investigating the influence of season on tau are scarce and their results inconclusive (Daan & Aschoff 1975; Gwinner 1975). Repeated measures of both tau and BMR throughout the year are necessary to get insight in possible seasonal fluctuations in these traits.

OUTLOOK

For micro-evolution in timing of breeding to occur, sufficient genetic variation in the physiological mechanisms underlying the date of egg-laying is a prerequisite, but it is currently unknown where in these physiological mechanisms genetic variation can be found. At present, a 'black box' exists between the genetic and phenotypic level of this key life-history trait, in which a cascade of (epi)genetic and physiological processes determines the phenotype expressed. In our selection lines, timing of breeding clearly responded to bi-directional genomic selection. We found genomic and phenotypic responses in egg-laying date and thus with these selection lines we now have a powerful tool to study the physiological mechanisms underlying timing of breeding. Future work using birds from these selection lines will therefore involve the physiology (incl. endocrinology), genetics (RNA-seq, qPCR) and epigenetics (DNA methylation) of timing of breeding. Exploring these should allow us to pinpoint in which part of the physiological cascade determining timing of breeding genetic variation exists and the amount of this genetic variation. Linking this information with (predicted) climate change should increase our understanding of how the evolutionary response to selection on seasonal timing, due to global climate change (Gienapp, Reed, & Visser, 2014), may be constrained by lacking or low genetic variation in crucial parts of the mechanism underlying timing of breeding.

Acknowledgements

We thank Marylou Aaldering, Coretta Jongeling, Franca Kropman, Anouk de Plaa and Ruben de Wit for taking care of the birds and many people for the hand rearing. We also thank Bart van Lith and Eva Maria Schöll for performing the BMR measurements, Tom Sarraude and Amrit Knoppers for fieldwork and Jeroen Laurens and Gilles Wijlhuizen for technical assistance prior to and during the experiments. We thank Jip Ramakers for kindly providing the heritability estimates of both tau and BMR, Kamiel Spoelstra for assisting with the chronobiology in this study, Mario Calus for sharing his insight in GEBVs and Arild Husby for overall comments on the manuscript. We are very grateful to Loeske Kruuk, and an anonymous associate editor and reviewer for their constructive comments that helped significantly improve the manuscript. This study was supported by an ERC Advanced Grant (339092 – E-Response to MEV).

Supplementary Information Chapter 2

Supplementary tables not presented here

These tables can be found at https://hdl.handle.net/10411/6Q1YDC

Table S2.2. Summary statistics of all the SNPs used in the study.

Table S2.3. Summary statistics table of SNPs with $p<0.01$.

Table S2.4. Genes linked to SNPs that had an F_{ST} $p<0.01$.

Table S2.5. Significant GO terms associated with highest F_{ST} when using the HUMAN GO-database.

Table S2.6. Significant GO terms associated with highest F_{ST} when using the CHICKEN GO-database.

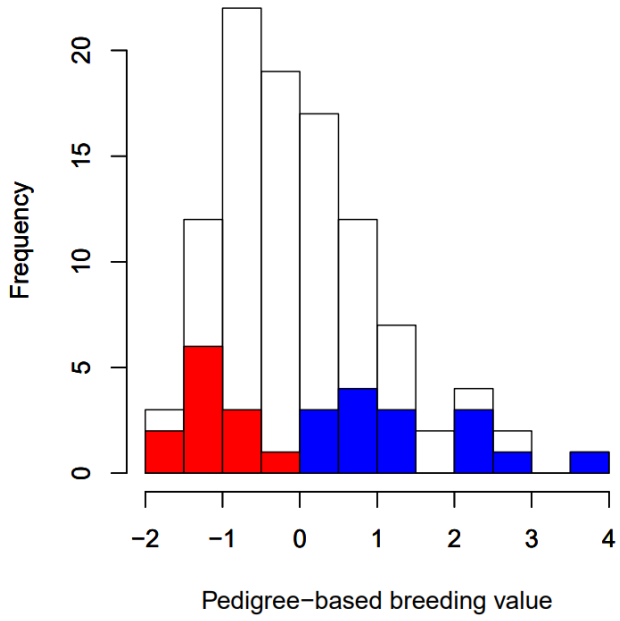


Figure S2.1. Histogram of the distribution of pedigree-based breeding values of females from the Hoge Veluwe population in 2014 (transparent bars). The females selected for the parental generation in red (*early* selection line) and blue (*late* selection line).

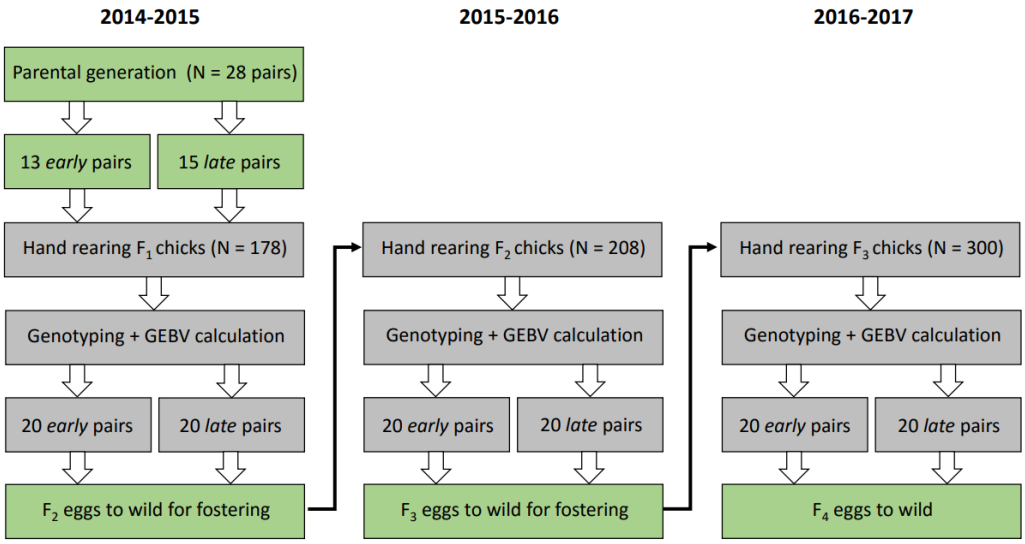


Figure S2.2. Simplified overview of the methodology from 2014-2017 behind the selection line experiment as described in the methods. Rectangles in green indicate when birds are in the wild, in grey when birds are at the NIOO-KNAW. In this figure, the 20 early and 20 late selection line pairs that are selected based on their most extreme GEBVs yearly to generate a new generation are shown. The remaining individuals from the yearly population of chicks have been selected for other purposes, which is beyond the scope of this paper and are described elsewhere (**Chapter 3-6**, Mäkinen et al. 2019).

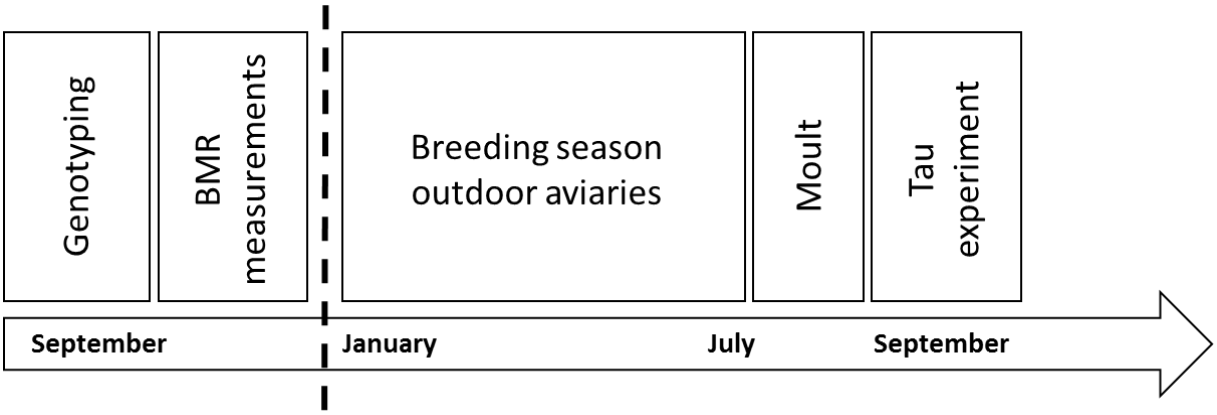


Figure S2.3. Representation of the chronological order of the several experiments conducted, which is identical for the years 2015 (F_1 generation) and 2016 (F_2 generation). Only some of the individuals from the F_3 generation went through the tau experiment, but were omitted from analyses (see '*Materials and methods*' for details). The vertical dashed line indicates the start of a new calendar year.

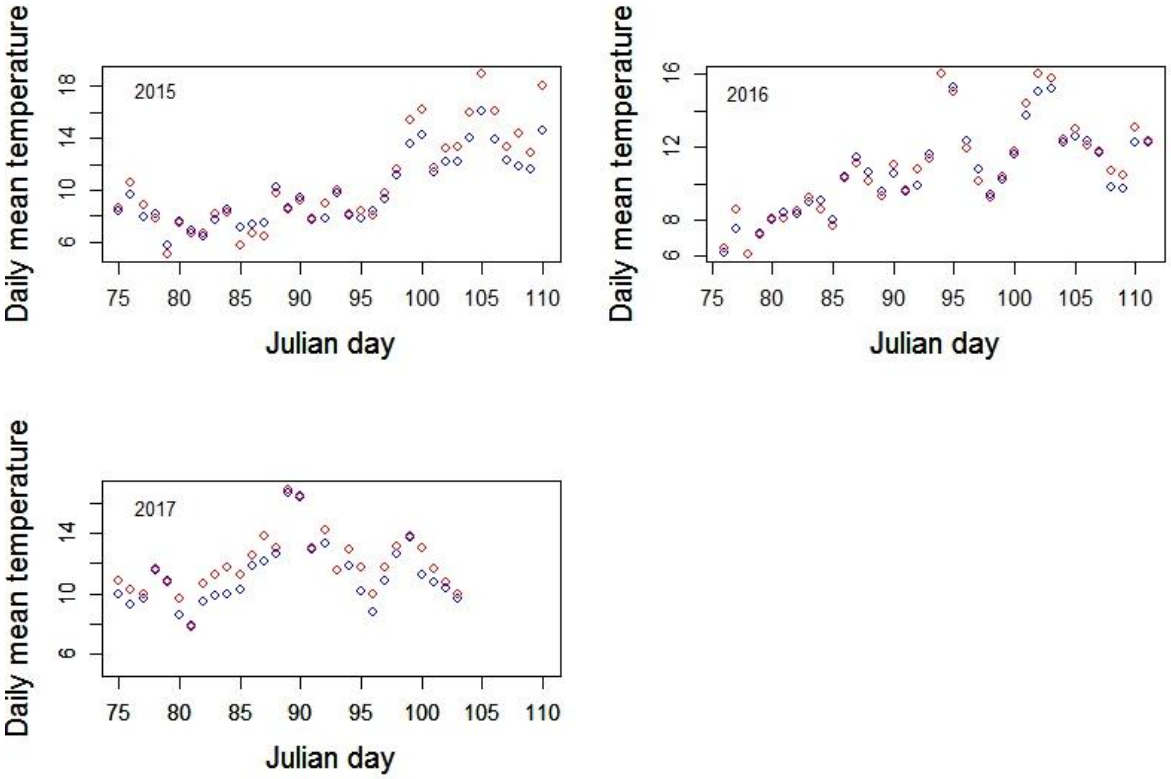


Figure S2.4. Mean daily temperature profiles throughout the breeding season in the outdoor aviaries for 2015, 2016 and 2017. Temperatures were collected using temperature loggers. The red and blue open circles represent the South and North side of the aviary building, respectively.

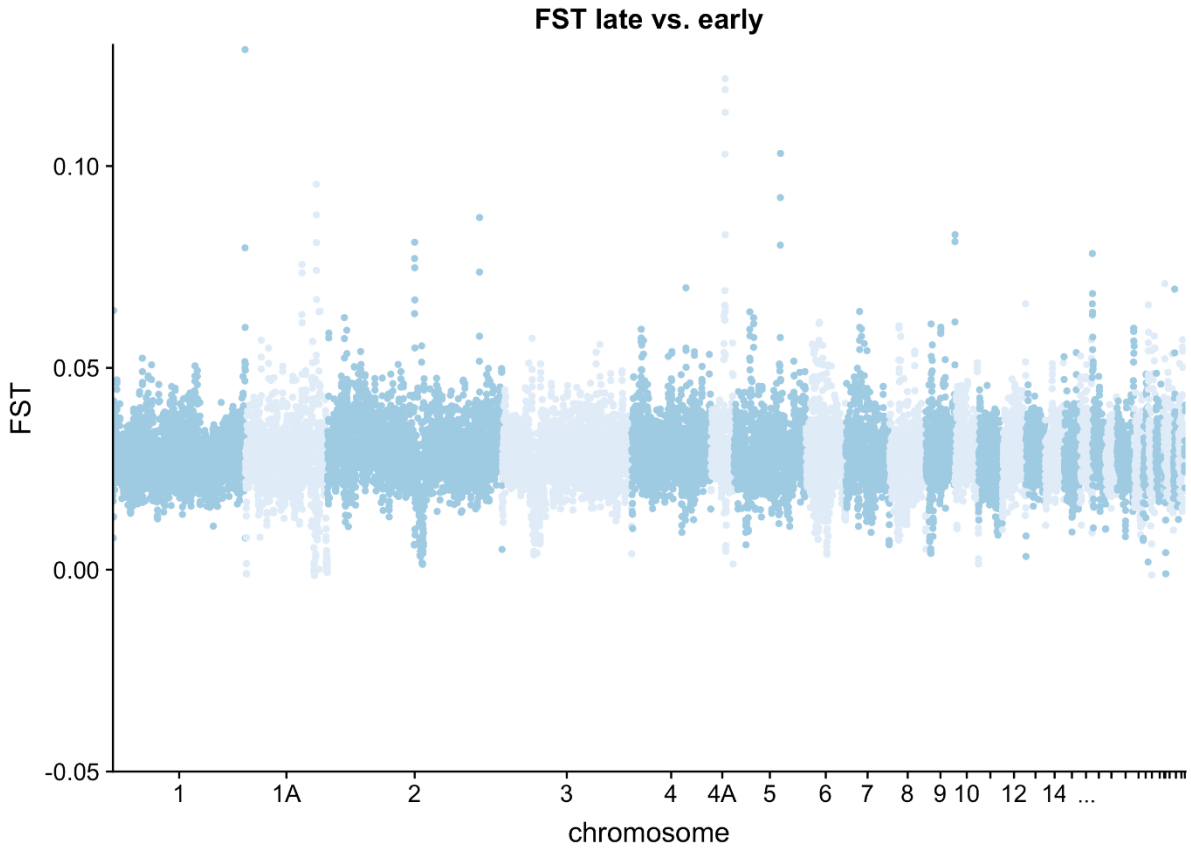


Figure S2.5. Manhattan plot showing the sliding window analysis on the F_{ST}-values estimated for each of the SNPs between the *early* and *late* selection line across the genome including all generations. The colours are to distinguish chromosomes from each other.

Table S2.1. Test results for influence of neighbouring birds on the individual onset of activity (Mantel test, 99 permutations). Tests are shown per room, where we used 12 cages per room in room 1 and 3 and 24 cages in room 2 divided over two walls. During the tau experiment, selection lines were mixed over the rooms, but always the same sex was tested per experiment.

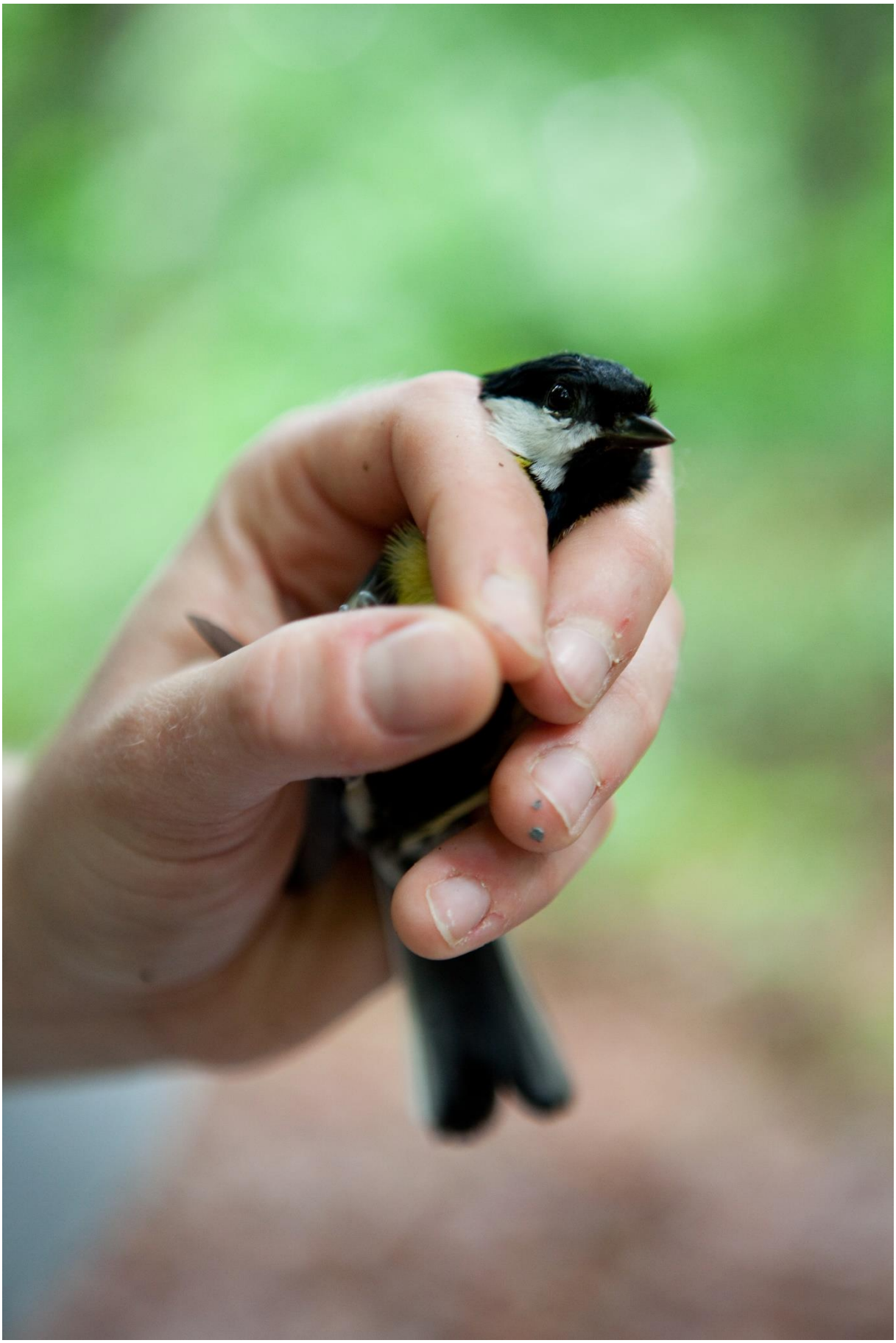
Subset of individuals	Observed	Expected	Variance	*p*-value
F_1 females				
Neighbouring birds room 1	-0.011	-0.004	0.049	0.55
Neighbouring birds room 2, wall 1	0.044	0.029	0.096	0.45
Neighbouring birds room 2, wall 2	-0.277	-0.022	0.051	0.80
Neighbouring birds room 3	-0.197	-0.001	0.065	0.77
F_2 females				
Neighbouring birds room 1	-0.281	0.051	0.051	0.98
Neighbouring birds room 2, wall 1	0.130	-0.023	0.021	0.17
Neighbouring birds room 2, wall 2	0.237	-0.013	0.035	0.12
Neighbouring birds room 3	0.632	0.121	0.128	0.09
F_1 males				
Neighbouring birds room 1	0.510	0.030	0.077	0.06
Neighbouring birds room 2, wall 1	0.804	0.024	0.289	0.21
Neighbouring birds room 2, wall 2	0.089	-0.029	0.047	0.28
F_2 males				
Neighbouring birds room 1	-0.404	0.030	0.079	0.97
Neighbouring birds room 2, wall 1	0.085	-0.040	0.023	0.25
Neighbouring birds room 2, wall 2	-0.329	-0.014	0.055	0.92
Neighbouring birds room 3	0.099	0.013	0.041	0.35

Table S2.7. Estimated parameters of the linear models investigating the degree of variation in tau (n = 115) explained by selection line, generation, sex and their interactions. All parameters were fixed effects in the model and statistics are given for the point of exclusion from the model. Bold p-values indicate significance.

Parameter	Estimate	s.e.	F-ratio	Df, ndf	*p*-value
Line × generation × sex	-0.020	0.133	0.02	1, 107	0.880
Line × sex	-0.006	0.065	0.00	1. 108	0.932
Generation × sex	-0.029	0.065	0.21	1, 109	0.650
Line × generation	0.048	0.063	0.58	1, 110	0.448
Line (*late*)	0.026	0.031	0.70	1, 111	0.405
Generation (F_2)	0.033	0.032	1.10	1, 112	0.296
Sex (*male*)	0.096	0.031	9.78	1, 113	**0.002**
Intercept	23.643	0.021			

Table S2.8. Estimated parameters of the linear mixed effect models investigating the degree of variation in basal metabolic rate (n = 620) explained by both fixed effects and random effects. Parameters are given for the point of exclusion from the model. For the fixed effects and random effects, the F-statistic (*F*) and the Chi-Square (χ^2) statistic are given, respectively. Bold p-values indicate significance.

Parameter	**Parameter**	**s.e.**	**Test-statistic**	***p*-value**
Fixed effects				
Line × generation			$F_{2, 596.4} = 2.31$	0.101
Late × F_2	0.551	0.501		
Late × F_3	-0.412	0.460		
Generation			$F_{2, 467.6} = 2.75$	0.065
F_2	0.118	0.267		
F_3	0.528	0.254		
Line (*Late*)	0.263	0.189	$F_{1, 613.1} = 1.16$	0.281
Seks (*male*)	0.110	0.246	$F_{1, 427.8} = 0.03$	0.853
Outside temperature at 17:00 PM	0.948	0.096	$F_{1, 275.8} = 27.92$	**<0.001**
Morning mass	-0.111	0.021	$F_{1, 550.2} = 120.33$	**<0.001**
Intercept	14.684	1.627		
	Variance	**s.d.**		
Random effects				
Channel of respirometer	0.088	0.297	$\chi^2 = 3.4905$	0.062
Date of BMR measurement	0.253	0.503	$\chi^2 = 2.9326$	0.087



CHAPTER 3

Temperature has a causal and plastic effect on timing of breeding in a small songbird

Irene Verhagen, Barbara M. Tomotani, Phillip Gienapp, & Marcel E. Visser

ABSTRACT

Phenotypic plasticity is an important mechanism by which an individual can adapt its seasonal timing to predictable, short-term environmental changes by using predictive cues. Identification of these cues is crucial to forecast species' response to long-term environmental change and to study the potential to adapt to such change. Individual great tits (*Parus major*) start reproduction early under warmer conditions in the wild, but whether this is effect is causal is not well known. To study this, we used 76 pairs of great tits from the F_3 generation of lines selected for early and late egg-laying. We housed 36 pairs in climate-controlled aviaries and 40 pairs in outdoor aviaries, where they bred under artificial contrasting temperature treatments or in semi-natural conditions, respectively, for two consecutive years. In the climate-controlled aviaries, temperature treatments were reversed in the second year. Pairs subjected to a warm treatment in the first year, were subjected to a cold treatment in the second and vice versa. The pairs in the outdoor aviaries bred in the same aviary for two years. We thus obtained egg-laying dates in two different thermal environments for each female. Our results show that females, independent of aviary type, bred earlier under warmer conditions. We then evaluated whether our selection for two lines of egg-laying dates also changed the birds' phenotypic plasticity. Early selection line females initiated egg-laying consistently ~9 days earlier compared to late selection line females in outdoor aviaries (i.e. late females had a higher reaction norm elevation), but we found no difference in the degree of plasticity in timing of breeding between selection lines (i.e. no difference in reaction norm slopes between lines). Given that temperature causally affects egg-laying date, climate change will lead to earlier egg-laying, but this advancement is unlikely to be sufficient to keep up with the advancement of resources, leading to selection on earlier egg-laying. Our results also suggest that while natural selection may lead to a change in phenotype in the average environment it is unlikely to result in a correlated response on the degree of plasticity in timing of breeding.

Manuscript

Introduction

Effects of global climate change are omnipresent and severely modify the environmental conditions for wild populations (Parmesan & Yohe 2003; Walther 2010). In particular, phenological traits are highly sensitive to these environmental modifications. This has led to phenological changes in numeral taxa, covering all trophic levels, and subsequent mismatches between them (Cohen et al. 2018; Parmesan 2006; Root et al. 2003; Visser & Gienapp 2019) because trophic levels do not necessarily respond at the same rate (Thackeray et al. 2010; Visser et al. 1998; Visser & Holleman 2001; Voigt et al. 2003). It remains largely unknown, however, to which processes to attribute these phenotypic changes, because it proves difficult to distinguish between genetic changes and phenotypic plasticity (Gienapp et al. 2008; Merilä & Hendry 2014), i.e. the environmentally induced production of different phenotypes by a single genotype (Pigliucci 2001), which hampers predictions of species adaptations to ongoing climate change.

In temperate zone birds, the breeding period is short and varies yearly due to yearly varying environmental conditions. Consequently, females need to track this inter-annual variation to optimally time their breeding to the time when food resources are most abundant in order to support successful rearing of offspring (Charmantier et al. 2008; van Noordwijk et al. 1995; Perrins 1965; Sheldon et al. 2003). Females thus need to be able to 'predict' when to initiate breeding, for which they use environmental cues. Photoperiod is an important cue, as it gives initial important information to track the time of the year (Dawson et al. 2001; Farner 1985) and starts up the reproductive system. However, its year-to-year invariability does not allow females to track year-to-year variation in optimal conditions (Bradshaw & Holzapfel 2007; Visser et al. 2004). Temperature is highly variable between years and does provide information about local conditions (Dawson 2008; Wingfield et al. 1992; Wingfield & Kenagy 1991), which allows female to fine-tune their timing of breeding. Therefore, temperature is the major driving force of biological seasonality in temperate zones (Parmesan 2007) and the most influential environmental cue in timing of avian seasonal breeding (Caro et al. 2013; Lambrechts & Visser 1999; Visser et al. 2009).

The temperature effect on timing of breeding seems to be direct, and not acting via e.g. food phenology (Schaper et al. 2011), as shown in a previous six-year study in great tits (Visser et al. 2009). In that study, breeding pairs were housed in climate-controlled aviaries and went through a breeding season in either a warm or a cold treatment, mimicking a warm and cold spring, respectively. Birds initiated egg-laying significantly earlier in the former. It is important to point out here, that in the study by Visser et al (2009) the breeding date comparison was performed *between* individuals, thus showing between-individual plasticity and that the effect of temperature on egg-laying varied strongly between years, despite controlled conditions. Here we used a more direct approach to determine whether plasticity in timing of breeding is due to the direct or indirect (e.g. via food phenology) effect of temperature on timing of breeding by keeping birds in contrasting temperatures for two years and gathering *within*-individual data.

To test for the within individual plasticity of egg-laying date to temperature, we housed pairs in climate-controlled aviaries for two consecutive years, in which they went through a warm spring followed by a cold spring or vice versa. For the cold and warm spring, we mimicked the temperatures from the year 2013 (extreme cold spring in The Netherlands) and 2014 (extreme warm spring in the Netherlands), respectively. In addition, we housed pairs in semi-natural conditions for two consecutive years. In our experiment, we used birds from selection lines for early and late timing of breeding in great tits (*Parus major*), which we created by using bi-directional genomic selection (Gienapp et al. 2019; **Chapter 2**). We have shown that early selection line females laid on average about six days earlier compared to late selection line females (**Chapter 2**). By using birds from these selection lines, we have the unique opportunity to test whether our selection for egg-laying date (the elevation of the reaction norm) had a correlated response to selection on the sensitivity to temperature (the slope of the reaction norm).

Currently, knowledge on how cues are perceived (Caro et al. 2013; Dawson 2008) and on how individuals vary in their perception (individual-by-environment interaction or I×E) is still scarce (Lyon, Chaine & Winkler 2008; Visser 2008; Visser et al. 2010). Identification of these cues, and understanding of the responses of breeding plasticity to selection, therefore, are crucial to forecast species' responses to long-term environmental change and to study the potential for adaptation to such change.

Materials and methods

Selection lines in timing of breeding

We created selection lines for early and late timing of breeding in great tits using genomic selection, which was moderately strong and in both directions (for details see Gienapp et al. 2019 and **Chapter 2**). Briefly, nestlings (F_1 generation) were taken in from wild broods of our long-term study population in the Hoge Veluwe of which the mother had initiated egg-laying either extremely early (*early* line) or extremely late (*late* line) in the wild. These chicks were genotyped using a 650 SNP chip (Kim et al. 2018) in order to predict their 'genomic breeding values' (GEBVs, i.e. the value estimating the relationship between genotype and phenotype based on genetic markers). Based on their GEBVs, F_1 generation individuals were selected for *early* and *late* line breeding pairs to produce the F_2 generation in captivity. Eggs were transferred to wild 'foster-nests', where they were incubated and hatched. Subsequently, 10 days post-hatching F_2 generation chicks were brought into the aviary-facilities at the NIOO-KNAW (Wageningen, the Netherlands) for further hand raising. In their turn, the F_2 offspring were genotyped and selected to produce the F_3 generation, which was then genotyped and selected.

The results of the selection line study are written elsewhere (**Chapter 2**), but in short, we found that *early* line birds laid earlier than *late* line birds, and this difference in egg-laying date increased (from about 2 to 10 days) over the generations (F_1 to F_3), with non-significant

line effects for the F_1 and F_2, but highly significant line differences for the F_3. On average *early* line birds laid about six days earlier compared to *late* line birds (**Chapter 2**). Note that these results were found in the birds housed in outdoor aviaries (see below). No differences were found between selection lines and treatments in birds that were housed in climate-controlled aviaries (see below, **Chapter 5**).

Outdoor aviaries

For a detailed description, see **Chapter 2**. In short, from January 2017 onwards, F_3 generation pairs (n = 40) were housed in 40 outdoor aviaries (4.2 × 1.9 × 2.1m) where the birds were subjected to natural photoperiod and temperatures. These pairs had the most extreme GEBVs (see above) within the F_3 generation and functioned as breeding pairs to produce eggs (F_4-generation) to be put in the wild as part of another study. Temperatures were recorded in 20 out of the 40 aviaries every 10-30 minutes using loggers (Thermochron iButton) throughout the breeding season.

Climate-controlled aviaries

The climate-controlled aviaries have been described in detail elsewhere (**Chapter 5**), but briefly, 36 pairs of the F_3 generation of the selection lines (see '*Selection lines in timing of breeding*' above) were housed in 36 climate-controlled aviaries in January 2017. These birds had less extreme GEBVs (see above) as compared to the birds housed in the outdoor aviaries. In the climate-controlled aviaries, birds received an artificial photoperiod following the change in natural photoperiod. In addition, two contrasting temperature treatments (Figure S3.1A) were provided mimicking an extreme cold (2013) or extreme warm (2014) spring in the Netherlands (for details see **Chapter 5**), which was reflected in egg-laying dates between these years (but see '*Results*' below). Every hour temperatures changed to follow as closely as possible the observed hourly temperatures in these years (note that the minimum temperature in the aviaries was 2°C so any temperature below 2°C in the temperature time series from outside was set to 2°C). Mimicking these natural temperature patterns is important to be able to infer realistic conclusions. Even though it is still a challenge to define what information in an experienced temperature profile is used to time breeding, previous research indicated that the seasonal increase in temperature, rather than the average temperature explains fine-tuning avian timing of breeding (Schaper et al. 2012). The combination of selection line and temperature treatment resulted in four groups (n = 9 pairs per group) in the climate-controlled aviaries: 'early-warm', 'early-cold', 'late-warm' and 'late-cold'.

Birds in both aviary types were fed *ad libitum* and had water available for drinking and bathing (for details see Visser et al. 2011).

Breeding seasons and egg-laying dates

*A*ll birds went through their breeding season in 2017 and were housed in single-sex groups afterwards. In January 2018, the same pairs were housed in the same outdoor or climate-controlled aviary they were in in the breeding season of 2017 to go through their second breeding season. In the climate-controlled aviaries, when pairs were subjected to the warm temperature treatment in 2017, they received the cold treatment in 2018 and vice versa. During these breeding seasons, nest boxes in all the aviaries were checked twice a week for nest building and daily, when a female had completed her nest, for eggs. The first day an egg was laid by a female is recorded as her egg-laying date. We were able to obtain two egg-laying dates recorded in two different environments for 34 out of 40 outdoor aviary females and 32 out of 36 climate-controlled aviary females.

Statistical analysis

While for the climate-controlled aviaries we know which temperature profile is associated with early egg-laying, as we mimic temperature from an early and a warm year, this is not the case for the outdoor aviary years. To determine a difference between temperatures in 2017 and 2018 in the outdoor aviaries, we tested with t-test the differences in mean daily temperatures between 2017 and 2018 from 16 March – 15 April, which is the period in which the temperatures correlate the best with mean annual egg-laying dates in our long-term wild population in the Hoge Veluwe (Visser et al. 2006).

To analyse the reaction norms, we used a mixed model analysis of variance (procedure 'lmer', package lme4, R 3.5.1, R Development Core Team 2018). Fixed effects were year and selection line for the outdoor aviary females and treatment, order of treatment and selection line for the climate controlled aviary females. The random effects were female identity and family with female identity nested in family. We could not test for female age, because for the outdoor aviaries, age is completely confounded with year. A significant selection line term indicates that selection lines differ in the elevation of the egg-laying date-environment relationship (i.e. lines differ in their average egg-laying date in the average environment). The interaction between year or treatment, depending on the aviary type, and selection line was tested. A significant interaction term indicates that the slopes of the egg-laying date-environment relationship are significantly different (i.e. selection lines differ in their environment-related reproductive flexibility). Non-significant effects were eliminated in a stepwise model reduction procedure (procedure 'KRmodcomp', package pbkrtest).

In addition, as a preliminary indication, we tested whether the average daily increase in temperatures differed between temperature environments in both outdoor and climate-controlled aviaries, as a previous study showed that great tits used the increase in temperature rather than the mean warm temperatures to time their breeding (Schaper et al. 2012). For this, we used the 11 days prior to egg-laying, starting within the period that correlates best with timing of breeding in the wild (see above, Visser et al. 2006) and the

earliest egg-laying date in the warm environment. We used year and treatment as a fixed effect for the outdoor and climate-controlled aviaries, respectively.

Results

Timing of breeding is directly affected by temperature

In the outdoor aviaries females laid earlier in 2018 compared to 2017 (year = -6.32 ± 1.57, $F_{1,33} = 16.24$, $p < 0.001$, Figure 3.1A). With average daily temperatures being significantly lower in 2018 compared to 2017 ($t = 2.27$, df = 38.78, $p = 0.029$, Figure S3.1B), the spring in 2018 is perceived as the warmer environment (but see '*Average daily increase in temperature prior to breeding'* below).

In the climate-controlled aviaries egg-laying dates were significantly affected by temperature, with birds laying earlier in the warm treatment, indicating a direct effect of temperature on timing of breeding (treatment = -7.19 ± 2.69, $F_{1,31} = 7.17$, $p = 0.012$, Fig. 1B).

Age and order of treatment

In the climate controlled aviaries, we found no effect of the order of the treatments to which the birds were subjected (treatment order = -0.34 ± 3.52, $F_{1,28.7} = 0.002$, p = 0.96), meaning that plasticity in timing of breeding is not influenced by first experiencing a cold spring, followed by a warm spring or vice versa. Further, we found no effect of age (age = 0.86 ± 2.64, $F_{1,30} = 0.09$, $p = 0.762$).

Selection on timing of breeding results in a change in reaction norm elevation between selection lines

In the outdoor aviaries there was no difference in breeding plasticity (i.e. reaction norm slope) between the *early* and *late* selection line (selection line × year = -1.28 ± 3.09, $F_{1,32} = 0.16$, $p = 0.69$). However, the *late* selection line females showed a significantly higher elevation in the reaction norm for timing egg-laying of about nine days (selection line = 9.31 ± 3.01, $F_{1,32} = 8.73$, $p = 0.004$).

Between the *early* and *late* selection line females in the climate-controlled aviaries, there was no difference in the reaction norm slope (selection line × treatment = -1.50 ± 5.29, $F_{1,30} = 0.08$, $p = 0.784$) or elevation (selection line = -1.35 ± 4.10, $F_{1,30} = 0.11$, $p = 0.747$).

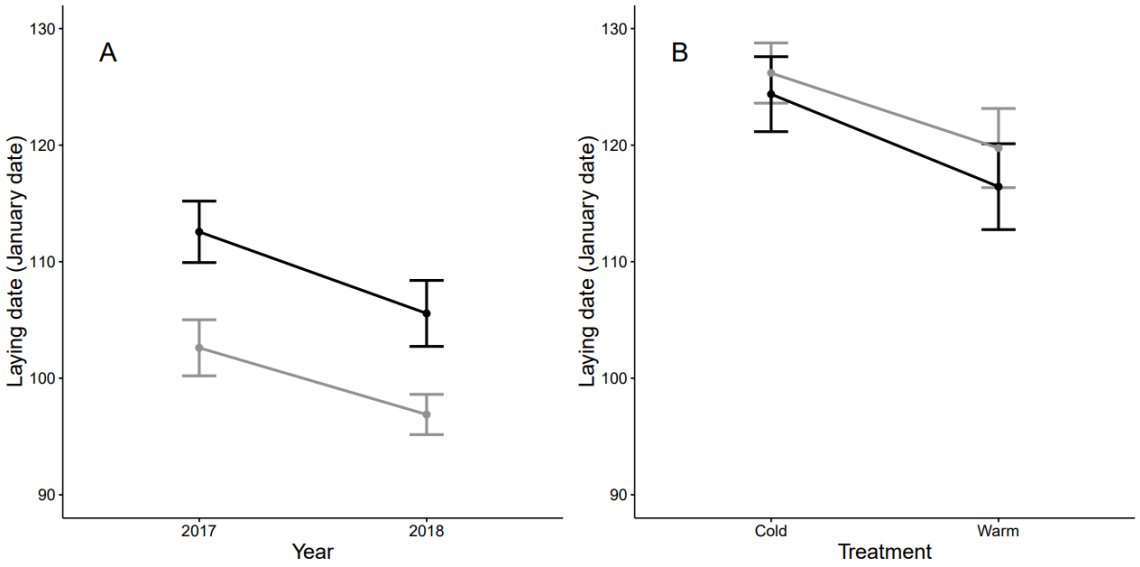


Figure 3.1. Reaction norms, presented as mean egg-laying dates (± s.e.m.) in two years for the outdoor aviary females (A) and in two temperature treatments in the climate-controlled aviary females (B) in. Egg-laying dates are presented in January dates (e.g. 90 January = 30 March). In the outdoor aviaries, late selection line females (black) show a higher elevation in egg-laying date compared to the early selection line females (grey). In the climate-controlled aviaries, there is no significant difference in elevation between selection lines.

Average daily increase in temperature prior to breeding

Outdoor aviaries: Both in 2017 and 2018, the earliest female started on 27 March, or 86 January (Figure S3.1B). The profiles of increasing temperature (Figure S3.2B) of the 11 days (75-86 January) until the start of breeding differed significantly between for 2017 and 2018 (year = -4.58 ± 0.87, $F_{1,22}$ = 27.5, p < 0.0001). Temperatures in this period increase on average with 0.09 ± 0.09 °C/day in 2017, whereas in 2018 this increase was 0.59 ± 0.16 °C/day.

Climate-controlled aviaries: The earliest females initiated breeding on 30 March (i.e. 90 January) and 16 April (i.e. 106 January) in the warm and cold treatment, respectively (Figure S1). We tested the mean increase in temperatures in the same period (79-90 January) for both treatments (Figure S3.2A). The main daily temperature over this 11-day period, increased 0.53 ± 0.29 °C/day in the warm treatment and 0.005 ± 0.02 °C/day in the cold treatment. The mean increase for the warm treatment was significantly different from the cold treatment (treatment = 6.05 ± 1.13, $F_{1,20}$ = 28.6, p < 0.0001).

Discussion

We studied whether temperature directly affects timing of breeding and whether selection on timing of breeding results in a correlated response to selection in plasticity in timing of breeding, using females from lines artificially selected for early and late egg laying. We found that females, independent of aviary type, initiated breeding earlier under warmer conditions showing that there is within-individual plasticity in the timing of reproduction in response

to temperature. We found no difference in the degree of plasticity between selection lines for both aviary types. However, in outdoor aviaries, early selection line females initiated egg-laying ~9 days earlier compared to late selection line females. Selection on timing of breeding, therefore, results in a change in phenotype in the average environment, but not in a correlated response to selection on the degree of plasticity in timing of breeding.

Currently, it is poorly understood what components of the temperature profiles, mean, minimum, maximum, change, etc., are used by birds to predict their breeding and how this information in perceived, transduced and ultimately translated into egg-laying. Interestingly, the temperature profiles provided in this study show periods of increasing mean daily temperature before females started initiating egg-laying (Figures S3.1 and S3.2). In the climate-controlled aviaries, the stable temperatures within the cold treatment (Figure S3.2A), when comparing the steep increase in temperatures in same 11-day period for the warm treatment, could explain why females in the cold treatment started breeding later Interestingly, despite that 2018 shows lower mean daily temperature compared to 2017 (Fig. S3.2B), females laid earlier in the outdoor aviaries in 2018. Also in these aviaries the increase in temperatures 11 days before the earliest female initiated breeding is steeper in 2018 (Fig. S3.2B). The results in both the climate-controlled and outdoor aviaries, are in concurrence with a previous study in which was shown that great tits used the increase in temperature rather than the mean warm temperatures to time their breeding (Schaper et al. 2012). Since we show that temperature directly affects egg-laying, future studies can try to pin-point which components of the temperature profiles birds use.

The breeding time reaction norm in the climate-controlled aviaries (controlled-aviaries: $\Delta_{\text{cold-warm}} = 7.2$ days), despite showing the same direction, differs greatly in slope compared to Hoge Veluwe females in 2013 and 2014 (Hoge Veluwe: $\Delta_{2013\text{-}2014} = 23.7$ days, n=48 females). This indicates that temperature is unlikely to be the only environmental driver affecting plasticity in timing of breeding and that other environmental factors are also involved, whether in interaction with temperature or not. A recent study in wild tree swallows (*Tachycineta bicolor*), for example, found that timing of breeding was mainly influenced by latitude and temperature, the latter in interaction with breeder density (Bourret et al. 2015).

Here, we found a difference in the timing of breeding in the average environment between the early and late selection line birds in the outdoor aviaries, but not in the climate-controlled aviaries. It is likely that the environments perceived in the outdoor aviaries, i.e. semi-natural conditions, give better or more complete information for timing of breeding. Further, genomic selection on timing of breeding resulted in a selection response in the outdoor aviaries not in the climate-controlled aviaries (**Chapters 2** and **5**). Possible reasons could be that females lack specific cues (Lambrechts et al. 1999) or experience a disrupted correlation between predictive cues (Bentley et al. 1998) in artificial conditions, and that this, in combination with a different genetic make-up (outdoor aviaries are more extreme), did not result in a difference in reaction norm elevation in climate-controlled aviaries.

While genomic selection on timing resulted in a change in phenotype in the average environment, at least in the outdoor aviaries, we found no correlated response to selection on plasticity in timing of breeding, independent of aviary type. This is in concurrence with a recently performed study in the long-term study population at the Hoge Veluwe from which these aviary birds descend (Ramakers et al. 2019). This study found directional selection on the elevation, but not the slope of the breeding time reaction norm. However, we must interpret the results from the aviaries with some reservation, because, as opposed to Ramakers et al (2019), we studied a limited number of females. With the strength of genomic selection on timing being moderate (**Chapter 2**), we may not have been able to detect changes in reaction norm slopes. In addition, due to this low sample size, we were unable to test the individual variation in plasticity (I×E) and whether it has a genetic basis (i.e. genotype-by-environment interaction or G×E). Further, both aviary types experienced two environments compared to other long-term studies performed in wild populations. Still, these results are promising for future studies (see below), especially due to the fact that they focus on patterns *within* individuals.

Global climate change will continue to disrupt the synchrony between interacting trophic levels, and therefore responding through phenotypic plasticity will likely not be sufficient in the long run (Thackeray et al. 2016; Visser 2008; Visser et al. 2004; Visser & Gienapp 2019). Genetic shifts in reaction norms are thus necessary for species to resolve the asynchrony in phenology between consumer and prey, but these shifts remain scarce (Merilä & Hendry 2014). In order for such a shift to occur, genetic variation in the mechanisms underlying phenological traits is necessary and we need to find where in these mechanisms this variation resides for selection to act upon. Experiments on temperature effects on timing of breeding contribute to our understanding of how birds respond to environmental cues. Here, by using a *within*-individual experimental approach, we show that temperature directly affects breeding time in a songbird. Thus, natural selection may lead to a change in phenotype in the average environment, but will likely not result in a correlated response to selection on the degree of plasticity in timing of breeding. Finding a direct effect of temperature on timing of breeding is exciting, as it advances our understanding of the mechanisms underlying breeding decisions under climate change. Data and results from this study will be important in future studies that, for example, investigate within-individual DNA methylation patterns in contrasting treatments analysing plasticity in timing of breeding (Lindner et al. *in prep.*).

Acknowledgements

We thank Marylou Aaldering, Coretta Jongeling, Franca Kropman and Ruben de Wit for taking care of the birds and Jeroen Laurens and Gilles Wijlhuizen for technical assistance prior to and during the experiments. We are grateful to Jip Ramakers for the attempt on the IxE and GxE analyses and Samuel Caro for constructive comments on the manuscript. This study was supported by an ERC Advanced Grant (339092 – E-response to MEV).

Supplementary Information Chapter 3

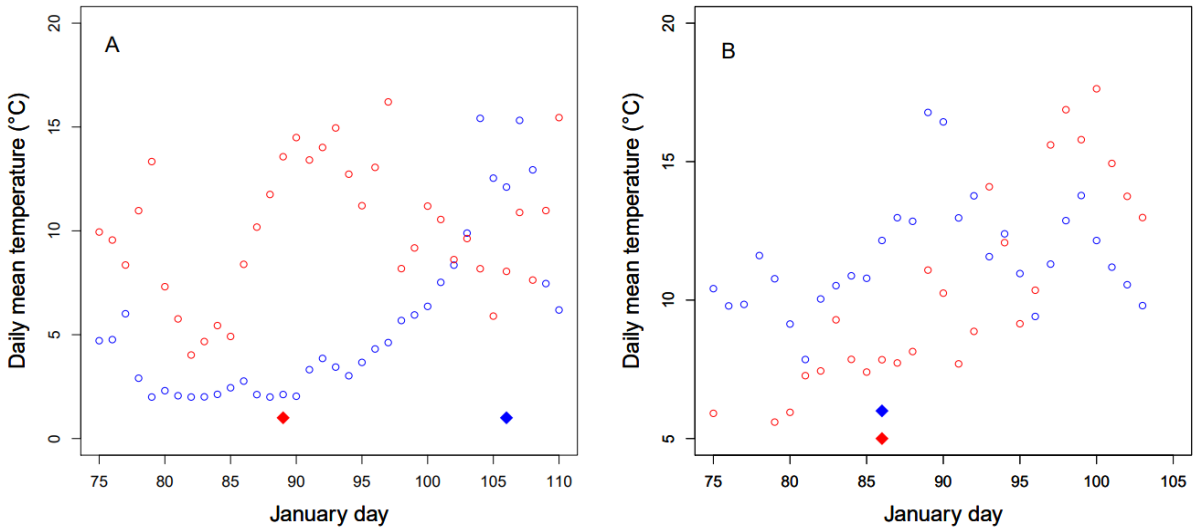


Figure S3.1. Mean daily temperature profiles during the breeding season (16 March – 20 April) in the climate-controlled aviaries (A) for the cold (blue) and warm (red) temperature treatment. For the outdoor aviaries (B) the temperature profiles of 2017 (blue) and 2018 (red) are shown. Note that for the outdoor aviaries temperatures are shown until 15 April. The filled diamonds indicate the first egg laid in the warm (red) and cold (blue) temperature treatment in the climate-controlled aviaries or the first egg laid in 2017 (blue) and 2018 (red) in the outdoor aviaries. Dates are in January days (January 1 = 1).

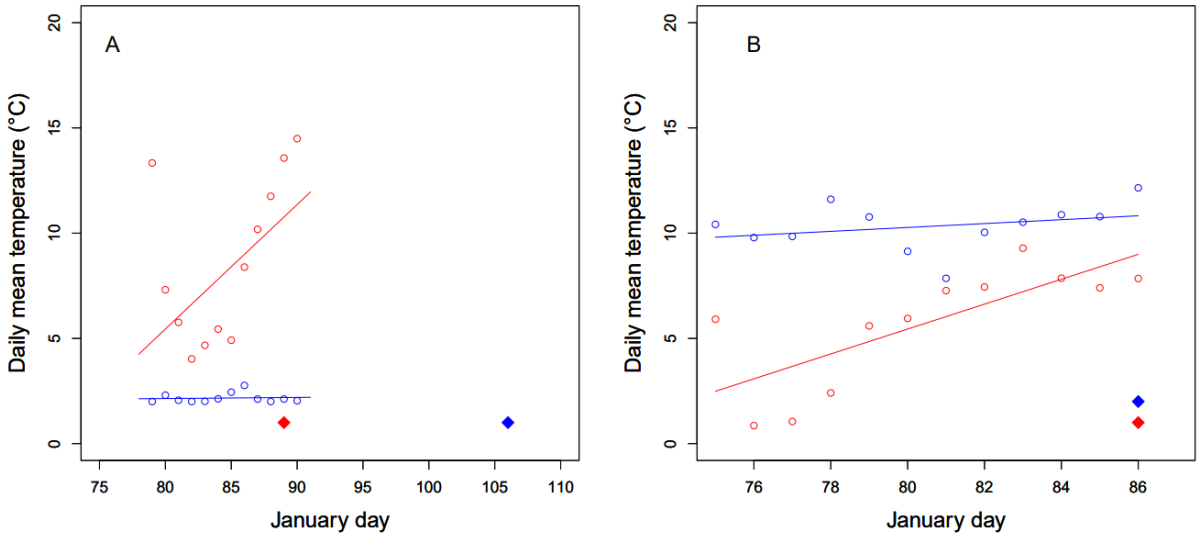


Figure S3.2. Mean daily temperature profiles of the 11-day period prior to breeding initiation in the climate-controlled aviaries (A) for the cold (blue) and warm (red) temperature treatment. For the outdoor aviaries (B) the 11-day temperature profiles of 2017 (blue) and 2018 (red) are shown. The filled diamonds indicate the first egg laid in the warm (red) and cold (blue) temperature treatment in the climate-controlled aviaries or the first egg laid in 2017 (blue) and 2018 (red) in the outdoor aviaries. Dates are in January days (January 1 = 1). Note that the January days shown differ between the climate-controlled and outdoor aviaries due to different first egg-laying dates between the aviary types.

CHAPTER 4

Exploration of tissue-specific gene expression patterns underlying timing of breeding in contrasting temperature environments in a songbird

Veronika N. Laine, Irene Verhagen, A. Christa Mateman, Agata Pijl, Tony D. Williams, Phillip Gienapp, Kees van Oers & Marcel E. Visser

ABSTRACT

Seasonal timing of breeding is a life history trait with major fitness consequences but the genetic basis of the physiological mechanism underlying it, and how gene expression is affected by date and temperature, is not well known. In order to study this, we measured patterns of gene expression over different time points in three different tissues of the hypothalamic-pituitary-gonadal-liver axis, and investigated specifically how temperature affects this axis during breeding. We studied female great tits (*Parus major*) from lines artificially selected for early and late timing of breeding that were housed in two contrasting temperature environments in climate-controlled aviaries. We collected hypothalamus, liver and ovary samples at three different time points (before and after onset of egg-laying). For each tissue, we sequenced whole transcriptomes of 12 pools (n=3 females) to analyse gene expression. Birds from the selection lines differed in expression especially for one gene with clear reproductive functions, zona pellucida glycoprotein 4 (*ZP4*), which has also been shown to be under selection in these lines. Genes were differentially expressed at different time points in all tissues and most of the differentially expressed genes between the two temperature treatments were found in the liver. We identified a set of hub genes from all the tissues which showed high association to hormonal functions, suggesting that they have a core function in timing of breeding. We also found ample differentially expressed genes with largely unknown functions in birds. We found differentially expressed genes associated with selection line and temperature treatment. Interestingly, the latter mainly in the liver suggesting that temperature effects on egg-laying date may happen down-stream in the physiological pathway. These findings, as well as our datasets, will further the knowledge of the mechanisms of tissue-specific avian seasonality in the future.

BMC Genomics (2019) ***20****(1): 693*

Introduction

Over recent decades, environmental change (e.g. climate change) has resulted in phenological shifts of spring events across trophic levels (Cohen et al. 2018; Parmesan & Yohe 2003; Root et al. 2003; Thackeray et al. 2010). In seasonally breeding birds, environmental change has the most profound effect on timing of breeding, i.e. timing of egg-laying, a life-history trait with major fitness consequences (Brinkhof et al. 1993; van Noordwijk et al. 1995; Visser et al. 2006). As such, seasonal timing of breeding has been under directional selection towards earlier egg-laying (Both & Visser 2001; Stevenson & Bryant 2000; Visser et al. 1998). In order to predict the responses to directional selection on timing of breeding via genetic changes, we need to understand both the novel and intensified selection pressures posed by environmental change on, as well as the genetic variation in, timing of breeding. Only those parts of the mechanisms underlying timing for which there is genetic variation can show a response to natural selection; these are the 'wheels' natural selection can turn (Visser 2008). Finding the genetic basis of timing of breeding is, however, complicated because there is a complex physiological mechanism underlying it, in which different organs and different environmental variables at different moments in time play a role (Visser et al. 2010).

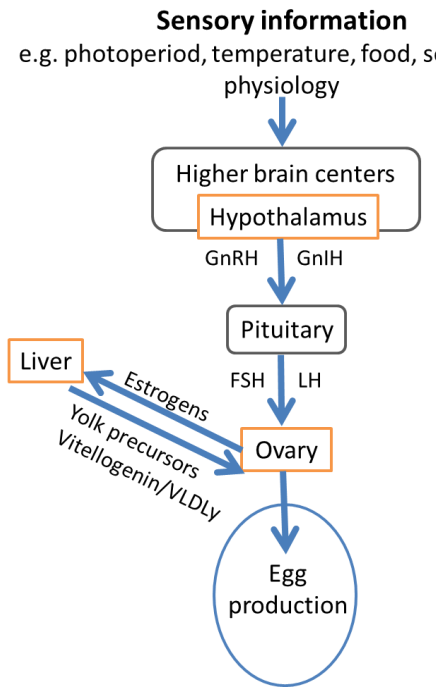


Figure 4.1. A schematic representation of the hypothalamic-pituitary-gonadal-liver axis in female birds (adapted from Williams 2012).

Photoperiod plays a main role in timing of breeding, as the yearly predictive increase in photoperiod in early spring provides precise information for birds to track the time of the year and stimulates the photoreceptors in the hypothalamus, which then send information along the photoperiodic signalling pathway (Nakao et al. 2008; Yoshimura 2010). This pathway, in turn, triggers the synthesis and secretion of gonadotropin-releasing hormone (GnRH) in the hypothalamus, which marks the activation of the hypothalamic-pituitary-

gonadal-liver (HPGL) axis (Figure 4.1, Nakao & Yoshimura 2008, Yoshimura 2010), a key pathway underlying gonadal growth and maturation in anticipation of the breeding season and ultimately timing of breeding (Dawson 2008, Williams 2012).

While the function of photoperiod is clear in timing of breeding (Dawson 2008) and its signalling pathway prior to the activation of the HPGL axis is well-known at the molecular level (Ubuka et al. 2013; Yoshimura 2010; Yoshimura et al. 2003), this remains largely elusive for temperature. We know, that it has a causal effect on (Visser et al 2009, **Chapter 3**) and 'fine-tunes' timing of breeding (Caro et al 2009, Williams 2012, **Chapter 5**), that breeding time varies greatly between and within females from one year to the next depending on spring temperatures (Charmantier et al. 2008; Gienapp et al. 2005; Gienapp et al. 2010), and that the effect of temperature varies throughout spring and across latitudes (Gienapp et al 2005, Gienapp et al 2010). Under global warming, seasonal breeding birds could use temperature information to adequately advance their egg-laying period. However, this advancement might at some point become constrained by the lack of responsiveness to the HPGL axis to an increasing temperature. This is implied by the weak relationship between the development of the HPGL axis and the onset of egg-laying (Schaper et al. 2012), suggesting that the way temperature acts on timing by-passes some major components of the reproductive system. However, it is unclear via which mechanism temperature is perceived and integrated (Caro et al. 2013). Thus, *how* temperature affects seasonal timing of breeding and if this is only in the brain, like photoperiod, or also elsewhere in the HPGL axis.

As pointed out above, changing environments pose selection pressures on phenological traits such as timing of breeding, and a better understanding of the regulation of different parts of the reproductive axis by environmental cues and its molecular basis is hence imperative, especially in the context of adaptation to climate change. For this study we use the great tit (*Parus major*), which is a model species in ecology and evolution, due to its willingness to breed in nest boxes, short generation time and large broods, and wide distribution (Perrins 1979). In addition, the study system of great tits, relying on caterpillars, which in turn rely on oak bud burst, is a well-known system (Perrins 1991) and showed different rates in shifts between trophic levels due to changing environments (van Asch et al. 2013; Both et al. 2009; Gienapp, Postma & Visser 2006; Visser et al. 1998). Recently, a comprehensive molecular toolbox became available, including a well annotated reference genome (Laine et al. 2016), whole transcriptomes and methylomes from several tissues (Derks et al. 2016; Laine et al. 2016; Santure et al. 2011) and two SNP chips, 10k and 650k (van Bers et al. 2012, Kim et al. 2018), making the exploration of the (epi)genetic architecture of life-history traits possible (Gienapp et al. 2017; Robinson et al. 2013; Santure et al. 2011, 2013; Viitaniemi et al. 2019). In addition to this toolbox, using the 650k SNP chip, selection lines for early and late egg-laying were created by genomic selection (Laine et al. 2016, **Chapter 2**) and, as such, can be used as an instrument to further investigate the mechanisms underlying timing of breeding. Nestlings (the F_1 generation) were taken from wild broods of which the mother was either an extremely early or extremely late breeder.

These chicks were genotyped and, based on their "genomic breeding values" (GEBVs), individuals were selected for early and late line breeding pairs to produce the F_2 generation in captivity (Gienapp et al. 2019; **Chapter 2**). The F_3 generation was then generated from the F_2 generation.

Here, making use of these tools, we measured overall gene expression levels by means of RNA-seq based expression profiling in three different tissues in great tit females housed in contrasting temperature treatments at three different time points related to egg-laying. As such, we explore time, temperature and tissue-specific gene expression patterns underlying timing of breeding. In order to identify molecular pathways likely to be involved in timing of breeding and the potential effect of temperature on these pathways, we performed functional gene enrichment analysis, network construction and hierarchal clustering of the RNA-seq datasets. In addition to exploring the molecular basis of seasonal breeding, our datasets and results will be an important starting point for future studies, especially on wild avian reproduction.

Methods

Experimental setup and samples

A detailed description of the experimental setup and sampling is described in (**Chapter 5**). In short, 36 great tit pairs (18 *early line* and 18 *late line* pairs) originating from the second generation (F_2) of lines artificially selected for early and a late timing of breeding (for details see (**Chapter 2**) were housed in 36 climate-controlled aviaries (2m × 2m × 2.25m) at the Netherlands Institute of Ecology. Birds were subjected to a photoperiod mimicking the natural photoperiod and to two contrasting environments mimicking a cold spring (2013) and a warm spring (2014) in the Netherlands (Figure S4.1.). Temperatures changed every hour to follow as closely as possible the observed hourly temperatures in these years. The combination of selection line and temperature environment resulted in four groups: 'early-warm', 'early-cold', 'late-warm' and 'late-cold'. Birds were fed *ad libitum* with a constant daily amount, had water available for drinking and bathing and their welfare were assessed twice a day by animal caretakers (Visser et al 2011). The pairs were used in two consecutive breeding seasons within one year (see **Chapter 5** for details); a first breeding season in spring and a second breeding season in autumn, after the birds went to a period of short-day length and low temperatures (see below).

First breeding season. In the first breeding season, initiated on 4 January 2016, the four groups were kept in pairs in the climate-controlled aviaries during spring. Nesting material (moss and hair) was provided from the second week of March onwards. Females could choose between three nest boxes of which two were accessible from the outside to minimalize disturbance. Females initiated nest building and subsequent egg-laying, which were recorded together with other reproductive traits (e.g. clutch size). In addition, both

sexes were blood sampled bi-weekly throughout the breeding season as part of another study (see Mäkinen et al. 2019).

Second breeding season. After the first breeding season, when birds were well on their way moulting (~mid-July), days were shortened to 9L:15D and temperatures decreased to 10°C for seven weeks to make the birds photoreceptive and temperature sensitive again. From September onwards, the pairs were subjected to the same photoperiod and temperature regimes again as in their first breeding season, to initiate their second breeding season. Four females were replaced with a sister, because they did not initiate egg-laying in the first breeding season. Females showed similar phenotypic responses in the first and the second breeding season (a significant correlation between lay date in the first breeding season and ovary size at time of sacrifice in the second breeding season; **Chapter 5**). Therefore, pairs were divided in three groups (n = 12 pairs per group) as such that the egg-laying date distribution (recorded in the first breeding season) were similar per group. Every group was sacrificed at a different time points (see '*Tissue collection and preparation*', Figure 4.2).

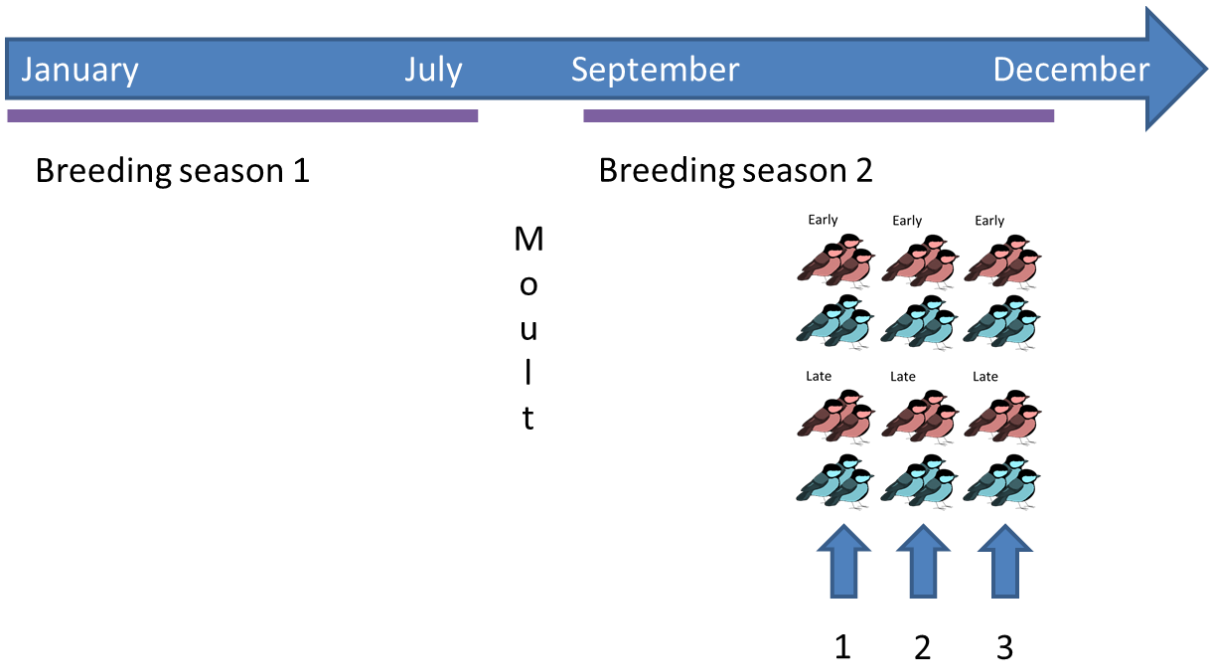


Figure 4.2. Visualization of the experiments through which the F_2 females (representing all four selection line × treatment combinations) in this study went. Females subjected to the warm and cold treatment are indicated in red and blue, respectively. The blue arrows indicate the three time point on which the tissues were collected: time point 1 when day length exceeds 11 hours, time point 2 when nest building occurs in the first breeding season and time point 3 when egg-laying was initiated in the first breeding season by 25% of the females.

Tissue collection and preparation. Three time points throughout the second breeding season were chosen, based on the reproductive behaviour from the first breeding season: (1) October 7 (resembling March 7) when photoperiod exceeded 11hrs (Silverin et al. 1993), (2) October 28 (resembling March 30) when nest building occurred in the first breeding season, but prior to egg-laying and (3) November 18 (or April 20) when about 25% of the

females in 2015 had initiated egg-laying in the first breeding season. We sacrificed one group (both males and females, but we focus on the females in this study) per time point (see **Chapter 5** for details, Figure 4.2). In short, birds were caught per pair between 9 and 12 AM from the aviaries, taken to the operation room and deeply anaesthetized with Isoflurane (IsoFlo, Zoetis, Kalamazoo, Michigan) using breathing mask during which a blood sample was also taken, followed by decapitation. Tissues, including brain, gonads and liver were dissected and stored in -80°C until further processing. At a later stage, the hypothalamus was isolated from the rest of the brain and, until further processing, stored in -80°C.

RNA extraction and sequencing

From hypothalamus, ovary and liver, RNA was isolated by Trizol extraction (see **Chapter 5** for details). We pooled RNA of three females per time/line/treatment group, resulting in a total of 12 pools (Figure 4.2). The library preparation and sequencing were performed at Baseclear, Leiden, The Netherlands. Libraries were made using the Illumina TruSeq strand-specific mRNA method ((Illumina, San Diego, CA, USA). We used one lane of Illumina HiSeq 2500 (single-end 50bp) for 12 pools. About 192 million single-end reads of 50 bp were generated for liver, 219 million reads for hypothalamus and 181 million reads for ovary.

RNA-seq analysis

Sequence data processing and differential gene expression analysis. Filtering of low quality reads was conducted at Baseclear by removing PhiX and adaptor sequences. The trimmed reads were mapped to the *Parus major* reference genome build 1.1.[1] using Hisat2 v2.1.0 (Kim et al. 2015) with default parameters. Transcript assembly was done using Cufflinks v2.2.1 (Trapnell et al. 2012), with default parameter settings and based on the *Parus major* annotation release 101[2] in NCBI. The obtained annotations were merged using cuffmerge. Unique reads that mapped to merged transcripts were counted using HTSeq v0.9.1 (Anders et al 2015).

All analyses were performed and figures made in R v.3.4.4. Clustering of the samples was done using Principal Component Analysis (PCA) using the 'regularized log transformation procedure' (rld) transformed expression values in order to diminish the number of variables and summarize the data. Differential expression of genes (DEG) between different time points, line and temperature were performed with DeSeq2 v3.6 (Love et al 2014) using the standard DeSeq2 protocol and Likelihood Ratio Test (LRT). LRT is useful for testing multiple terms at once compared to the default Wald test. The test examines two models, a full model with a certain number of terms and a reduced model, in which some of the terms of the full model are removed. The test determines if the increased likelihood of the data using the extra terms in the full model is more than expected if those extra terms are truly zero. Following the Deseq2 guidelines we created three main effect models: time point

[1] https://www.ncbi.nlm.nih.gov/assembly/GCF_001522545.2
[2] https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Parus_major/101/

(1 *vs.* 3), temperature and line, and included the two main variables that were not analysed in each of these three models as controlling variables. For the time point we also compared the expression patterns between one time point to the two others. We also had three interaction models using all two-way interactions of these three factors. Genes were considered differentially expressed if the adjusted p-value was below 0.05. Heatmaps were generated using the rld transformed expression values for DEGs using gplots and Pheatmap implemented in R.

Hierarchical clustering analysis and GO enrichment

Clustering of the DEGs was done separately for each tissue. A hierarchical dendrogram was generated using the 'hclust' function in R (R v.3.4.4), whereas the 'ward. D' objective criterion was used to merge a pair of clusters at each step. Trees were cut at k = 5, k = 3 and k = 3 in hypothalamus, liver and ovary time point models respectively and at k = 3 in hypothalamus for the time point-temperature interaction model to obtain clusters of genes that are expressed the similar way where k is the number of groups. Each cluster's fold change values at each time point were plotted as profile plots using ggplot2 in R.

For the significant DEGs a GO enrichment analysis was conducted per tissue using the Cytoscape plugin ClueGo 2.5.2 (Bindea et al 2009) with the human (30.9.2018) gene ontology and KEGG pathway databases (Kanehisa et al. 2017). Any transcripts that fell in multiple genes were removed from the analysis. Gene symbols starting with LOC ('LOC' + the GeneID is given when published symbol is not available, and orthologs have not yet been officially determined) were investigated by hand to determine if they had an ortholog in other species or if it was non-coding RNA (ncRNA). ClueGo constructs and compares networks of functionally related GO terms with kappa statistics. A two-sided hypergeometric test (enrichment/depletion) was applied with GO term fusion, network specificity and Kappa score were set as default and false discovery correction was carried out using the Bonferroni step-down method.

Weighed correlation network and hub gene analysis

The weighed correlation network analysis (WGCNA) (Langfelder & Horvath 2008) was used for getting the co-expression patterns among transcripts and we used the rld transformed data. WGCNA clusters genes with similar patterns of expression across samples to create modules of genes that are likely co-expressed. Because this method uses hierarchical clustering of expression values to group genes into modules, the connectivity of the genes in the modules could reflect the response to time, temperature, line or their interaction. After the modules were created, the correlation of the module eigengenes with time, temperature and line was calculated to examine the strength of the correlation of the module with a given trait. We first removed transcripts with low expression levels across time points (counts smaller than or equal to 4 at one time point) to only have high confidence transcripts and ran the function blockwiseModules to identify potentially co-regulated genes. We created a signed network using soft thresholding power based on the module results of the pickSoftThreshold function (hypothalamus = 6, liver = 20, ovary =

10), minimum module size of 30 transcripts, and a merge cut height of 0.25, 0.45, 0.35 in hypothalamus, liver and ovary, respectively in order to combine the similar modules from the same nodes to larger modules (Figure S4.2). After this we identified modules that were significantly associated with line, time point and temperature by correlating the module eigengenes with the treatments.

We further analysed the hub genes from the significant modules from each tissue and conducted a STRING pathway analyses (Szklarczyk et al. 2015) in order to see how co-expression translates to functional pathways. Hub genes were defined by module connectivity, measured by the absolute value of the Pearson's correlation (ModuleMembership > 0.8) and the significance of the relationship with treatments > 0.05. We analysed the hub genes in the STRING plugin (version 1.4.0) in Cytoscape, choosing confidence > 0.4 to construct a protein-protein interaction (PPI) network. In the PPI network, genes with a connectivity degree of ≥ 10 were also defined as hub genes. The common hub genes both in the co-expression network and the PPI network were regarded as "real" hub genes for subsequent GO enrichments analysis in STRING with default settings. In the PPI network we combined all the tissues together to see how the genes interact together between tissues.

Results

Phenotypic results

The phenotypic results are described in detail in **Chapter 2**. In short, we found no effect of either selection line, temperature treatment or their interaction (see '*Experimental setup*' and samples in '*Materials and methods*') on egg-laying dates (see '*First breeding season*' in '*Materials and methods*') and follicle size (see '*Second breeding season*' in '*Materials and methods*'). However, follicles were significantly larger at time point 3 compared to time points 1 and 2.

Sequencing and alignment

For the downstream analyses, we sequenced on average 18±3 million (mean ± s.d.) single end reads in hypothalamus, 16±2 million reads in liver and 15±2 million read in ovary and the overall alignment rate was on average 82.3% in hypothalamus, 79.8% in liver and 91.2% in ovary (Table S4.1).

Differentially expressed genes (DEGs)

When using the 'regularized log transformation procedure' (rld) transformed expression values from the DeSeq2 (Love et al 2014) package in the principal component anaylsis (PCA), we found that in the hypothalamus there was no clear clustering among time points,

line or treatment (Figure S4.3a). The PC1 explained 38% of the variance and that of PC2 is 22%. However, in liver and ovary the PC1 (with over 50% variation explained) clearly separated time point 3 samples from time points 1 and 2 (Figures S4.3b and S4.3c, respectively). Taken together, the PCA analysis provided the first evidence of a clear distinction of gene expression profiles between different time points especially within liver and ovary in our dataset.

In the differential gene expression analysis with DeSeq2, we found significant differences between time points in 491, 569 and 5175 transcripts in hypothalamus, liver and ovary, respectively (Table 4.1, Figure 4.3, Tables S4.2-S4.3, and Figures S4.4-S4.6). We also did pairwise comparison with one time point to the two other time points and most of the expression differences occurred in time point 1 and 3 (Table S4.5).

Table 4.1. The number of genes showing significant differential expression in the main effect models for the three tissues. Time point refers to contrasts between the three time points. Line refers to contrasts between the early and late selection line and temperature is the contrast between the warm and cold treatment.

Tissue	Time point	Line	Temperature
Hypothalamus	491	26	5
Ovary	5175	46	2
Liver	569	10	30

There was a line effect in hypothalamus and ovary (Table 4.1). In the line main effect model for ovary one gene, the zona pellucida glycoprotein 4 (*ZP4*), clearly stood out having a strong differentiation between lines (Figures S4.6 and S4.14, Table S4.4).

Most of the DEGs between warm and cold treatments were found in liver while, interestingly, the hypothalamus showed a significant interaction effect between time point and temperature forming two clear clusters of upregulated genes (Figure 4.3D, Table S4.5). The pools from the warm condition were shifted between time points compared to the cold treatment (Figure 4.3D). For the other interaction models there were 0 to 14 differentially expressed genes in all of the tissues (Table S4.5).

Hierarchical clustering of DEGs and GO enrichment analysis

We used hierarchical clustering of the DEGs to determine clusters of genes that changed through time in a similar way. We identified four, three and two clusters in hypothalamus, liver and ovary time point models respectively, and two groups in the hypothalamus time point-temperature interaction model (Figure 4.3, Figures S4.7-S4.10 and Tables S4.2-S4.4). Each cluster had a particular expression profile over time (and temperature in the

hypothalamus interaction model). In most clusters there was a linear increase or decrease of expression towards time point 3, but there were one cluster in hypothalamus (clusters 2) and one cluster (cluster 3) in liver which showed relatively higher/lower expression at time point 2 compared to time points 1 and 3.

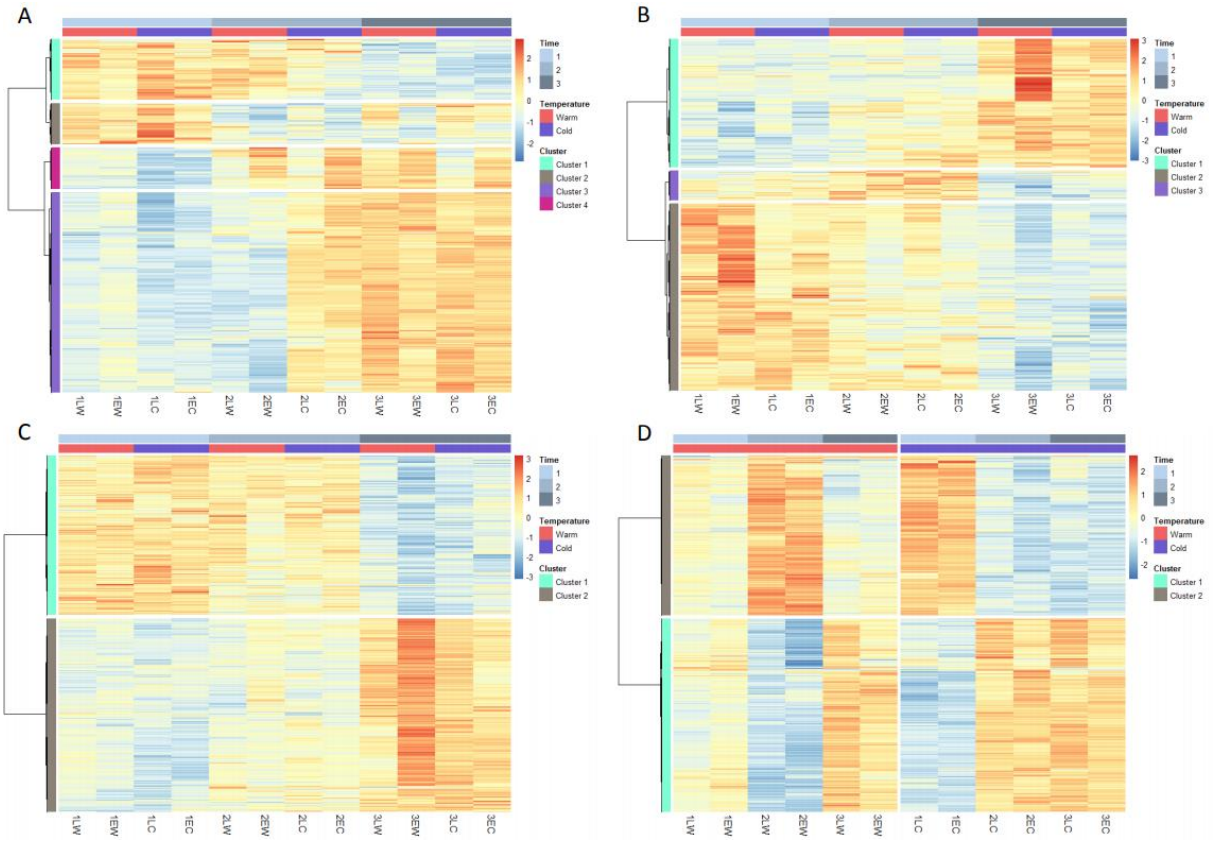


Figure 4.3. Heat map of genes that were significantly differentially expressed in the time point main models in hypothalamus (A), liver (B) and ovary (C) and in the time point x temperature interaction model in hypothalamus (D) (Note that in (D) samples were grouped based on temperature treatment, which differs from A-C). Genes were clustered by distances based on Pearson correlation coefficients in both figures. Lighter colours indicate lower differential expression; row Z- score scales from –3 (dark blue) to 3 (dark orange).

A functional enrichment analysis was possible for the time point main-effect models for all of the tissues and also for the time point-temperature interaction model for hypothalamus when the significance level was set to $p<0.05$ for DEGs. Enrichment analysis for the main effect model in hypothalamus showed that in genes that were upregulated in time point 1 (cluster 1) 11 different GO categories and KEGG pathways were overrepresented. These were related especially to circadian rhythm related GO terms and pathways (Table S4.6). Genes that had increased expression towards time point 3 (clusters 3 and 4) had 45 different GO terms and KEGG pathways overrepresented and from these especially GABA activity and other neuronal function related GO groups were significantly enriched. Cluster 2 (low expression at time point 2) had one GO category overrepresented.

In the interaction model between time point and temperature for hypothalamus, there were two gene clusters. Cluster 1 genes contained upregulated genes at time point 3 in both temperatures but the expression pattern in time points 1 and 2 differed between warm and cold treatments. In Cluster 2 the expression pattern was the opposite; upregulated genes at time points 1 and 2 compared to time point 3 and differing patterns between temperature treatments in time point 3. The genes and GO groups in cluster 1 (242 functional terms) were similar to the main model results with functions related to neuronal activity and the GABA pathway. However, the upregulated genes in cluster 2 (323 functional terms) were related to ribosomal, mitochondrial and ATP related metabolic functions (Table S4.7).

In liver, there were 130 GO terms and KEGG pathways enriched in genes that were upregulated in time point 1 (cluster 2). These terms and pathways were related to immunological functions, hormone responses and insulin response. Genes, upregulated at time point 2 (cluster 3), were linked to two GO terms: carbon-nitrogen lyase activity and oxidoreductase activity. In time point 3 (cluster 1) 32 GO groups and KEGG pathways were enriched which were especially related to protein processing and amino acid response (Table S4.8). Furthermore, egg-laying related genes, cathepsin E-A-like gene (LOC107205210, *CTSEAL*), vitellogenin 2 (*VTG2*; LOC107208431 and LOC107208432) and apovitellenin 1 (*APOV1,* LOC107200088) were expressed at this time point (Figure S4.11 and Table S4.3). The expression level increase of *VTG2* and *APOV1* had fold change of nine and *CTSEA*L fold change of 7 from time point 1 to time point 3 where early line showed larger increase.

In ovary, the genes that were upregulated at time point 1 (cluster 1 with 130 functional groups) were related to cell cycle, chromosome functions and spindle formation (Table S4.9). Five bird-specific egg related genes; *VTG2*, ovalbumin (*OVAL*; LOC107215075), ovalbumin-related protein Y (*OVALY*, LOC107214443), lamin-L(III)-like (*LMINA*; LOC107209405) and avidin (*AVD*; LOC107198337), were expressed at time point 1. In time point 3 (cluster 2 with 803 functional groups) genes were related to morphogenesis and development. The "egg-laying gene" *APOV1* was expressed at time point 3 and also bird specific major histocompatibility complex class II beta chain (*BLB2*; LOC107199337) gene (Table S4.4).

To explore the tissue specific circadian gene activity, we compared our DEGs to the genes from the super pathway 'BMAL1-CLOCK, NPAS2 activates circadian gene expression' from Path Cards, a pathway unification database (http://pathcards.genecards.org; Belinky et al. 2015), which lists 86 genes that activate the circadian gene expression pathway. In total, we found 41 genes of this pathway that were significantly differentially expressed between time points or in interaction with temperature in hypothalamus. Most of these circadian genes were found in ovary (28 genes) (Table S4.10).

Table 4.2. Summary of gene modules identified with weighted correlation network analysis (WGCNA). Only showing modules with significant correlation with the treatments. Network/module is number of genes found by STRING analysis out of the whole set of module genes that passed the selection threshold (ModuleMembership > 0.8 and treatment p-value >0.05). Top hub genes were chosen based on high modular membership (kME) value and highest degree in PPI network.

Tissue	Module color	Number of genes	Most significant correlation	Network/module	Hub gene symbol	Gene name
Hypothalamus						
	Brown	1668	Time point	117/130	*ADCY2*	adenylate cyclase 2
	Turquoise	5631	Time point	395/493	*HSPA8*	heat shock protein family A (Hsp70) member 8
	Blue	5068	Time point	784/911	*MAPK1*	mitogen-activated protein kinase 1
	Green	1022	Time point	65/72	*EPRS*	glutamyl-prolyl-tRNA synthetase
	Yellow	1031	Line	55/58		
Liver						
	Pink	714	Temperature	41/49		
	Turquoise	3149	Time point	202/264	*GART*	phosphoribosylglycinamide formyltransferase…
	Magenta	537	Time point	30/31	*SNAP25*	synaptosomal-associated protein 25
	Red	901	Time point	32/39		
	Green	951	Time point	39/46	*HSPA4*	heat shock protein family A (Hsp70) member 4
	Blue	3115	Time point	263/289	*PTPRC*	protein tyrosine phosphatase, receptor type C
	Purple	434	Time point	21/25		
	Salmon	363	Line	25/34	*POLR3B*	RNA polymerase III subunit B
	Midnightblue	160	Line	6/13		
Ovary						
	Yellow	2229	Time point	147/174	*HSP90AA1*	heat shock protein 90 alpha family class A member 1
	Turquoise	6579	Time point	913/1372	*ACLY*	ATP citrate lyase
	Blue	5573	Time point	1722/2252	*SRC*	SRC proto-oncogene, non-receptor tyrosine kinase
	Brown	3093	Time point	557/658	*AKT1*	AKT serine/threonine kinase 1
	Cyan	123	Line	5/15		
	Midnightblue	102	Line	3/3		

Weighed correlation network and hub gene analysis

To investigate the patterns of co-expression among transcripts, we analysed the rld transformed data using weighed correlation network analysis (WGCNA) (Dufour et al. 2005). We constructed five, nine and six co-expression modules for hypothalamus, liver and ovary, respectively that were significantly associated with the treatments (Table 4.2). There were modules that were significantly correlated with every treatment in liver but in hypothalamus and ovary none of the modules correlated with the temperature (Table 4.2, Figures S4.11-S4.13). There was an overlap with the genes between modules and DEGs where most of the overlap was with the time point model in every tissue type and also the interaction of time and temperature in hypothalamus suggesting time point being the most

driving effect of co-expression in our samples (Table 4.3). Details of the transcripts belonging to the modules are provided in Tables S4.11-4.13.

Table 4.3. The number of significantly differentially expressed genes from time point, temperature and line models overlapping with the members of gene modules from WGCNA (see Table 4.2). Only showing modules with significant correlation with the treatments.

Tissue	Module color	Total	Time	Temperature	Line
Hypothalamus					
	Brown	96	57	0	6
	Turquoise	770	78	1	1
	Blue	1070	280	1	0
	Green	41	32	0	6
	Yellow	44	21	0	4
Liver					
	Pink	10	5	5	0
	Turquoise	185	176	4	0
	Magenta	8	7	0	0
	Red	83	71	10	0
	Green	53	48	3	0
	Blue	214	205	4	1
	Purple	16	16	0	0
	Salmon	7	2	0	5
	Midnightblue	1	1	0	0
Ovary					
	Yellow	184	172	0	7
	Turquoise	1996	1967	0	17
	Blue	2242	2233	1	2
	Brown	579	576	0	2
	Cyan	7	6	0	1
	Midnightblue	3	3	0	0

We could determine 13 'real' hub genes out of 21 modules based on combination of co-expression and PPI network connections (Table 4.2). The network analysis of "real" hub genes from each module significantly associated with the treatments showed that all the genes (Tables S4.14-4.16), were in the same PPI network with all of them belonging to molecule binding GO term and most significant pathways were oestrogen signalling and progesterone-mediated oocyte maturation pathways (Figure 4.4 and Table S4.17).

Discussion

We measured gene expression levels in 12 pools of three female great tits each, from two lines selected for early and late egg-laying, which were kept at two contrasting temperature treatments, and were sampled at three time points across the breeding season. Most of the DEGs varied between time points 1 (well before egg-laying) and 3 (at the time of egg-laying). Gene expression levels of females from both lines and temperature treatments were following similar patterns in ovary and liver. In hypothalamus, however, we found a significant interaction between time point and temperature, which could indicate that temperature affects the timing of certain gene expression levels mainly in the brain. We found no effect of temperature on either egg-laying date in the first breeding season or follicle size in the second breeding season (see '*Downstream regulation of timing of breeding*' below). Many of the highly DEGs had an unknown function; either being non-coding RNA or the gene has an unknown function especially in birds. Furthermore, in every tissue we identified hub genes that may play a central role in timing of reproduction in great tits.

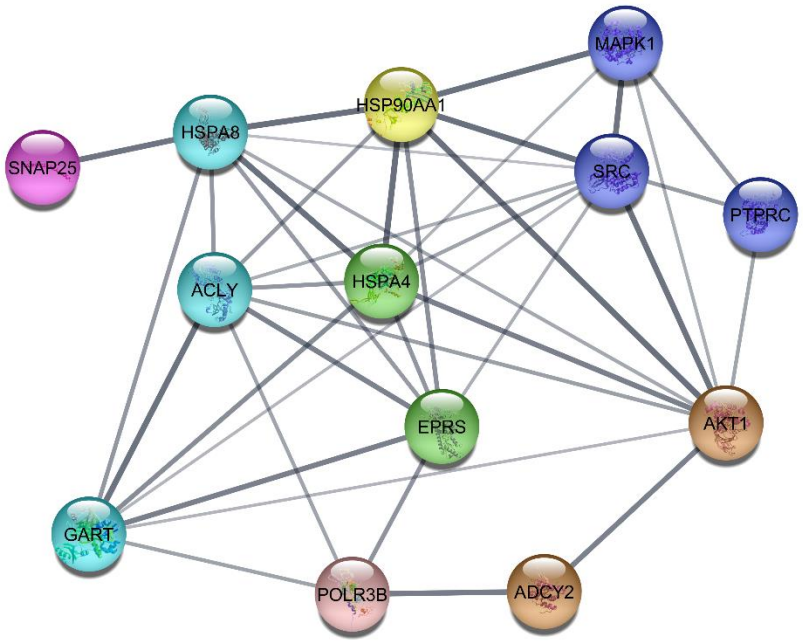


Figure 4.4. The PPI network of the 'real' hub genes from all tissues combined. The line thickness indicates the node connection score; thicker line means more evidence for the connection from existing research (experimental, co-expression, database). Colours correspond with modules from Tables 4.2 and 4.3.

Limitations of the data

Due to the small sample size for each time point in this study, the statistical analysis likely suffered from low power to detect differences between time points, temperature and line. We used pooled data without any replication and especially the interaction in hypothalamus would have benefitted of having individual level expression data with replication. Unfortunately, it is not possible to obtain tissue samples from the same individual in every time point to see how the expression patterns change in one individual. In great tits it has been shown that there is some genetic variation both in the onset and termination of egg-laying, and in the underlying mechanisms, and sometimes there is also an interaction with temperature (Schaper et al. 2012; Visser et al. 2011). Some families of birds are able to respond more quickly than others to the increasing temperature, which leads to differences in timing of breeding between families (Gienapp et al. 2006; Mccleery et al. 2004; Nussey et al. 2005; Sheldon et al. 2003; Visser et al. 2011). However, in our case we tried to minimize the relatedness within the pools and both lines were grouped in a similar way with regards to the temperature so relatedness might not play an important role for these results. We are thus positive that our results give a comprehensive overview of the different genes being expressed during seasonal timing as the PCA, differential expression and WGCNA give similar results. In addition, the results found also match with those described in the literature. In future studies, however, it is essential to confirm these results at the individual level and have additional time points before time point 1, as used here, in order to pin point the exact moment when preparing for breeding starts (i.e. activation of the HPGL axis). Furthermore, having gene expression levels in both the ovary tissue and follicular tissue could help us to recognise specific ovarian and/or follicular functions. Also, the addition of other tissues would help building the whole network of interacting genes (Bornelöv et al. 2018).

Downstream regulation of timing of breeding?

Although gene expression levels in the hypothalamus seem to be affected by temperature, this does not *directly* lead to earlier egg-laying, because we found no effect of temperature on either egg-laying date (first breeding season) or follicle size (second breeding season). Our data are in line with the hypothesis that downstream processes in the liver and ovary play a more important role in the fine-tuning of egg-laying date than hypothalamic processes (**Chapter 5**; Bergeon Burns et al. 2014; Caro et al. 2009; Needham et al. 2019). In this sense, the absence of an effect on egg-laying date of temperature is informative on where in the neuro-endocrine cascade fine-tuning occurs, rather than that it hampers new insights.

Time point effects

Time point 1. At time point 1, the genes expressed in the hypothalamus were related to circadian rhythm and photoperiodism. In fact, in every time point and every tissue, also in the interaction model in hypothalamus, several genes involved in circadian rhythms were

differentially expressed. In addition to the HPGL axis, the role of the circadian clock in annual cycles has been suggested for some time (Bünning 1969; Schultz & Kay 2003). The circadian phase at which light affects the photoreceptive elements causes reproductive changes (Cassone & Yoshimura 2014). The core of the avian circadian system is located in the pineal gland (Cassone 2015). This core clock acts as a master regulator of the rhythms of peripheral tissues. In birds a rhythmic expression of the clock genes has been identified in the mediobasal hypothalamus, suggesting that this structure contains the circadian pacemaker associated with photoperiodic time measurement (Yoshimura 2010, 2013).

Interestingly, there was not much overlap in circadian genes between tissues and also between the two models (main effect and interaction) in hypothalamus. The more downstream tissues (i.e. ovary and liver) also possess their own circadian clockworks and entrain their tissue-specific rhythms through their own, the core or both outputs of the circadian system (Asher & Schibler 2011; Nakao et al. 2007; Sellix & Menaker 2010). Especially the circadian clocks in the ovary may play a role in the timing of ovulation (Ball 2007; Nakao et al. 2007; Sellix & Menaker 2010). The circadian genes from the hypothalamus main model were mostly related to regulation and entrainment of the circadian rhythm. In the hypothalamus genes related to activin receptor signalling pathway were also upregulated at time point 1. Activin which is produced by gonads but also in extragonadal tissues, can enhance FSH biosynthesis in the pituitary gland and in hypothalamus activin stimulates GnRH release and thereby affects the levels of FSH and LH (Bilezikjian, Vaughan & Vale 1993; Dalkin et al. 1999; DePaolo 1997; Gregory & Kaiser 2004).

In liver, eight differentially expressed molecular clock related genes were mainly expressed at time point 1 with their majority being circadian regulators of gene expression (nuclear receptor subfamily 1 group D member 1, *NR1D1;* neuronal PAS domain protein 2, *NPAS2;* period circadian clock 2*, PER2;* period circadian clock 3*, PER3* and basic helix-loop-helix family, member e41, *BHLHE41*). In birds, changes in circadian gene expression in liver has been linked to alteration in the seasonal state (Trivedi et al. 2014). However, timing of the circadian clock in liver is often controlled by feeding rather than by the core clock system the brain (Peek et al. 2012; Stokkan et al. 2001). At time point 1 also the estrogen signalling pathway and hormone stimulus related GO terms were enriched suggesting that the liver could be processing hormonal signals from the ovary in order to start vitellogenesis later in the season. In addition, at time point 1 there were also genes upregulated that were belonging to immunological and insulin related functions. Both adaptive and innate immune responses produced by liver have been found in chicken ovaries, and these systems function to protect against colonization and infection by microbial pathogens, as well as to maintain normal functions of the ovary (Johnson 2015)Insulin is suggested to be one of the key regulators of reproductive function by having an effect on GnRH/LH secretion (Sliwowska et al. 2014).

We found that the ovary exhibited the most DEGs and co-expressed genes. In the pools at time point 1 ovary was expressing genes that were related to cell cycle, mitosis and meiosis suggesting that it already started with the ovarian maturation, along with follicle development (Johnson 2015b; Sánchez & Smitz 2012). For example, the expression of the genes important for follicular development such as the transforming growth factor beta (TGF-β) superfamily (such as growth differentiation factor-9, *GDF9* and bone morphogenetic protein 15, *BMP15*) (Elis et al. 2007; Juengel et al. 2004) and other genes such as forkhead box L2 (*FOXL2*) and NOBOX oogenesis homeobox (*NOBOX*) (Sánchez & Smitz 2012) were already high so it seems that great tits start folliculogenesis six weeks before laying the first eggs. There is an intermediate pre-vitellogenic follicle development phase in chicken which resides between the slow stage which is the development of primordial follicles and can last several months and the rapid follicle/rapid yolk development growth stage which can happen just few days before laying the first eggs (Johnson 2015b). During this intermediate phase small amount of lipoprotein rich white yolk are incorporated to the follicles increasing slightly their size and some of them are selected to final maturation stage (Johnson 2015b). Because vitellogenesis in liver and the increased expression of LH receptors in ovary happens at time point 3, our birds might indeed be in the pre-vitellogenic phase at time point 1 (and also 2), as shown by follicular measurements in the same females (**Chapter 5**).

Time point 2. Many of the DEG clusters from the time point 1 were also upregulated at time point 2 such as the circadian and activin related genes in hypothalamus and ovary's cell cycle related genes. In hypothalamus there was also a cluster of genes that were starting to be expressed at time point 2 and continued to be highly expressed at time point 3 as well. These genes were related to female reproduction such as the genes progesterone receptor (*PGR*) and prolactin receptor (*PRLR*, see below). There were also genes part of angiogenesis and one of them being fibroblast growth factor (*FGF1*) which has also been shown to be linked to egg fecundity in chicken albeit from the bone RNA samples (Johnsson et al. 2016).

In liver there was a specific upregulated gene cluster on time point 2. These genes were related to oxidoreductase and carbon-nitrogen lyase activity which do not have known function in reproduction. Both GO groups shared one gene, the aldo-keto reductase family 1 member B10 (*AKR1B10*) which is known to be part in detoxifying compounds under oxidative stress conditions and it has also been shown in humans that aldo–keto reductases are part of steroid hormone action and nuclear receptor signalling (Penning & Drury 2007). Oxidoreductase related functions continued being important as well at time point 3 where also amino acid metabolism and protein processing related GO groups were associated in which both oxidoreductase enzymes are important factors.

Time point 3. In time point 3 in hypothalamus the upregulated genes were related to many neuronal function groups but also to GABA receptor functions. GABA, the main inhibitory neurotransmitter, and glutamate, the main stimulatory neurotransmitter, set a level of

sensitivity in the hypothalamus that decreases or increases the likelihood that GnRH will be synthesized or released based on the reproduction status of the females (Maffucci & Gore 2009). Other HPGL axis genes that are known to be expressed in hypothalamus such as gonadotropin-releasing hormone 1 (*GnRH1*) was not expressed in our hypothalamus samples. However, *GnIH* (but annotated as neuropeptide VF precursor, *NPVF* in great tit), iodothyronine deiodinase 2 (*DIO2*) and thyroid stimulating hormone receptor (*TSHR*) were active in hypothalamus and from these *TSHR* was especially expressed on time point 3 indicating HPG cascade going towards egg production (Yoshimura 2013)(Yoshimura 2013).

At time point 3 in liver in addition to above mentioned metabolism and protein processing functional groups, vitellogenesis related genes were upregulated such as *VTG2* and *APOV1* which also showed line differences in expression levels where early lines had higher expression especially in the early-warm condition at this time point. Furthermore, cathepsin E-A-like gene (LOC107205210, *CTSEAL*) was upregulated at time point 3, which has been shown to be over-expressed during vitellogenesis in chicken liver and is regulated by estrogen (Bourin et al. 2012; Zheng et al. 2018). Next to *CTSEAL* in the great tit genome is bestrophin 3 (*BEST3*) which was also upregulated at time point 3. *BEST3* is an important gene in chloride channel activity but there is no known function in regards to reproduction. The similar expression pattern between *BEST3* and *CTSEAL* and their closeness in the genome suggests that they might be co-regulated but it is unclear in the current study if mRNA from *BEST3* used in the liver in the end. It is known that mRNA goes through several regulatory processes after it is made and this is often seen when comparing the expression levels from transcriptomes and proteomes (Payne 2015; Vogel & Marcotte 2012). In addition to *BEST3,* we found additional genes from every tissue that have unknown function in bird reproduction. There were also transcripts that are annotated as ncRNA by the NCBI. This type of RNA has been shown to be important in eukaryotic gene regulation and also in hormonal pathways and meiosis during reproduction (Taylor et al. 2015). Furthermore, there is evidence that miRNAs are differentially expressed in the ovary from sexually immature versus mature chickens, and in developing ovarian follicles relative to the stage of maturation (Kang et al. 2013).

Most of the circadian genes were expressed in the ovary and especially at time point 3 supporting the idea that these genes are important in starting the ovulation in birds (Ball 2007, Sellix & Menaker 2010). At time point 1 the two period genes, *PER2* and *PER3*, were upregulated. In poultry these two have been linked to preovulatory follicle expression (Nakao et al. 2007). In general, it is suggested that expression of ovarian circadian clock genes may be influenced by the increase of LH which may be a mechanistic link for communicating circadian timing information from the core clock in the brain to the ovary (Nakao et al. 2007; Tischkau et al. 2011). The receptors for FSH and LH (follicle-stimulating hormone receptor *FSHR*/LOC107202460 and lutropin-choriogonadotropic hormone receptor, *LHCGR*/LOC107201154) were expressed in ovary and especially the expression of *LHCGR* increased towards time point 3 suggesting increased LH activity in our ovary

samples. In birds the increased expression of LH receptors in ovary starts the final follicle maturation (Perfito et al. 2015). In addition to of the circadian genes, many of the upregulated genes were also related to developmental and morphogenesis GO groups and pathways. Interestingly, the mitogen-activated protein kinase (MAPK) signalling pathway was active at this time point as well. MAPK is proposed to inhibit *FSHR* transcription and is part of the cascade where pre-hierarchal follicles are selected into the preovulatory hierarchy (Johnson & Woods 2009) which is important at the rapid follicle development stage.

Temperature and line effects

In the hypothalamus gene expression was affected by the interaction between time point and temperature. However, due to limitations of the dataset the results should be treated as suggestive. Circadian genes were mostly expressed in time point 3 but there was a set of five circadian genes that were expressed at time point 1 which were mostly related to ubiquitination. Mutation in ubiquitin related genes can cause either elongation or shortening of the endogenous circadian period (tau) (Stojkovic et al. 2014). Interestingly, while photoperiod, nutrient and redox status can entrain the clock (Asher & Schibler 2011), temperature can affect the endogenous circadian period in great tits (Lehmann et al. 2012). Furthermore, in the interaction model in hypothalamus there were circadian genes that are regarded as the core genes in the circadian rhythm pathway such as clock circadian regulator (*CLOCK*), *PER2* and RAR related orphan receptor A (*RORA*) which also have pleiotropic effects to many metabolic processes (Asher & Schibler 2011; Buhr & Takahashi 2013).

In general, the interaction model had two gene clusters that showed distinctive patterns. The genes that were upregulated more during time point 1 and 2 in cold and warm environments, respectively were associated with metabolic-related terms and pathways such as ATP, NADH and ribosomal metabolic processes. The molecular clock constantly receives feedback from the metabolic signals in the cells (Asher & Schibler 2011; Mauvoisin et al. 2015) and can affect metabolism of the organism and is also controlled by metabolic pathways. The terms related to the second cluster which had genes upregulated more at time point 3 were similar to main effect time point model in hypothalamus by having the GABA pathways but also the circadian related terms. However, in this cluster there were also dopaminergic synapse pathway related genes upregulated. Dopamine together with prolactin influences the HPG axis primarily at the level of the hypothalamus and pituitary, by regulating the release of the gonadotropic hormones (Ben-Jonathan & Hnasko 2001; Dufour et al. 2005). *PRLR* was indeed also upregulated at time point 3 in our samples suggesting that both dopamine and prolactin were active in hypothalamus.

In contrast to the hypothalamus, no convincing effect of an interaction between time point and temperature (or just temperature alone) was found in liver and ovary, which was not surprising as no difference in egg-laying was observed between the temperature treatments.

Liver had 30 differentially expressed genes between the temperatures and it was the only tissue with a co-expression module associated with temperature. However, no GO enrichment analysis could be conducted with the genes and hub gene was not found in the module.

All the tissues showed some line differences in gene expression but in ovary one gene was highly differentially expressed. This was zona pellucida sperm-binding protein 4 (*ZP4*) which had a two to three times higher expression in early line compared to late line. It also appears in the co-expression results and is also under selection in these selection line birds (**Chapter 2**). *ZP4* is one of the genes responsible to making the zona pellucida (in mammals) or vitelline envelope (in fish, amphibians and birds), a glycoprotein layer surrounding oocytes (Litscher & Wassarman 2015). The zona pellucida mediates sperm–egg interaction, provides a post-fertilization block to polyspermy, and protects the embryo prior to implantation (Okabe 2013). In our selection line birds it is not known what role this gene plays between the lines.

Real hub genes for every tissue

All the "real" hub genes that shared high interaction both in the co-expression and the PPI networks were all transcribing binding molecules and they were all in the same final PPI network. Six genes were found in the estrogen signalling pathway (three from hypothalamus; *MAPK1, HSPA8, ADCY2* and three from ovary; *AKT1, HSP90AA1, SRC*). In addition, MAPK pathway being important in the ovary, *MAPK1* is estrogen activated in the brain and is important in female sexual behaviour (Kelly & Qiu 2010). Both *MAPK1* and *HSPA8* have been found to be differentially expressed in hypothalamus during spring migration in black-headed buntings (*Emberiza melanocephala*) (Sharma et al. 2018). *ADCYC2* in hypothalamus and *SNAP25* in liver are important genes in insulin secretion and four genes are important in temperature detection (two in hypothalamus: *MAPK1, HSPA8* and two in ovary: *AKT1, HSP90AA1*). In addition to estrogen signalling pathway, other hormonal pathways related to reproduction were associated with these hub genes such as progesterone, thyroid, prolactin and oxytocin binding/signalling pathways suggesting that our hub genes are important in female reproduction.

Conclusions

We generated comprehensive RNA expression data from a set of three tissues important in the neuro-endocrine cascade underlying avian seasonal timing of breeding, from three different time points and from two temperature treatments and two selection lines for breeding time. Time was the strongest driving variable in our dataset, as we would expect, but there was an interesting interaction between time and temperature in hypothalamus which should be studied more intensively in the future studies. It could be possible that gene expression in the brain is affected by temperature, perhaps through changes in

expression of genes involved in the circadian clock which affect the sensitivity to photoperiod. However, because egg-laying dates were not directly affected by temperature, the effect of temperature on timing of breeding is likely fine-tuned downstream in the reproductive axis, i.e. the liver and/or the ovary, rather than upstream, in the hypothalamus. These findings, as well as our datasets, will further the knowledge of the mechanisms of tissue-specific avian seasonality in the future.

Data and material accessibility

Sequence data: All quality-trimmed reads used in this study are available for download at the Short Read Archive (accession numbers SRR9644032-SRR9644067, bioproject PRJNA208335).

Acknowledgements

We thank Marylou Aaldering, Coretta Jongeling, Franca Kropman, Anouk de Plaa and Ruben de Wit for taking care of the birds. We also thank Renske Jongen, Bart van Lith, Heidi Viitaniemi,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. This study was supported by an ERC Advanced Grant (339092 – E-response to MEV).

Supplementary Information Chapter 4

Supplementary tables not presented here

These tables can be found at

https://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-019-6043-0

Table S4.2. The Likelihood Ratio Test results of the nine models and annotations for the transcripts in hypothalamus.

Table S4.3. The Likelihood Ratio Test results of the nine models and annotations for the transcripts in liver

Table S4.4. The Likelihood Ratio Test results of the nine models and annotations for the transcripts in ovary.

Table S4.6. Significant GO terms associated with the time point main effect model gene clusters in hypothalamus.

Table S4.7. Significant GO terms associated with the time point - temperature interaction model gene clusters in hypothalamus.

Table S4.8. Significant GO terms associated with the time point main effect model gene clusters in liver.

Table S4.9. Significant GO terms associated with the time point main effect model gene clusters in ovary.

Table S4.10. The genes from the super pathway 'BMAL1-CLOCK, NPAS2 activates circadian gene expression' found in our time point main effect models and from the time point - temperature interaction model.

Table S4.11. Modules of genes significantly correlated with time, temperature or line in hypothalamus.

Table S4.12. Modules of genes significantly correlated with time, temperature or line in liver.

Table S4.13. Modules of genes significantly correlated with time, temperature or line in ovary.

Table S4.14. List of highly connected module genes in hypothalamus that have at least one connection degree in the PPI network.

Table S4.15. List of highly connected module genes in liver that have at least one connection degree in the PPI network.

Table S4.16. List of highly connected module genes in ovary that have at least one connection degree in the PPI network.

Table S4.17. Significant GO terms associated with the real hub genes.

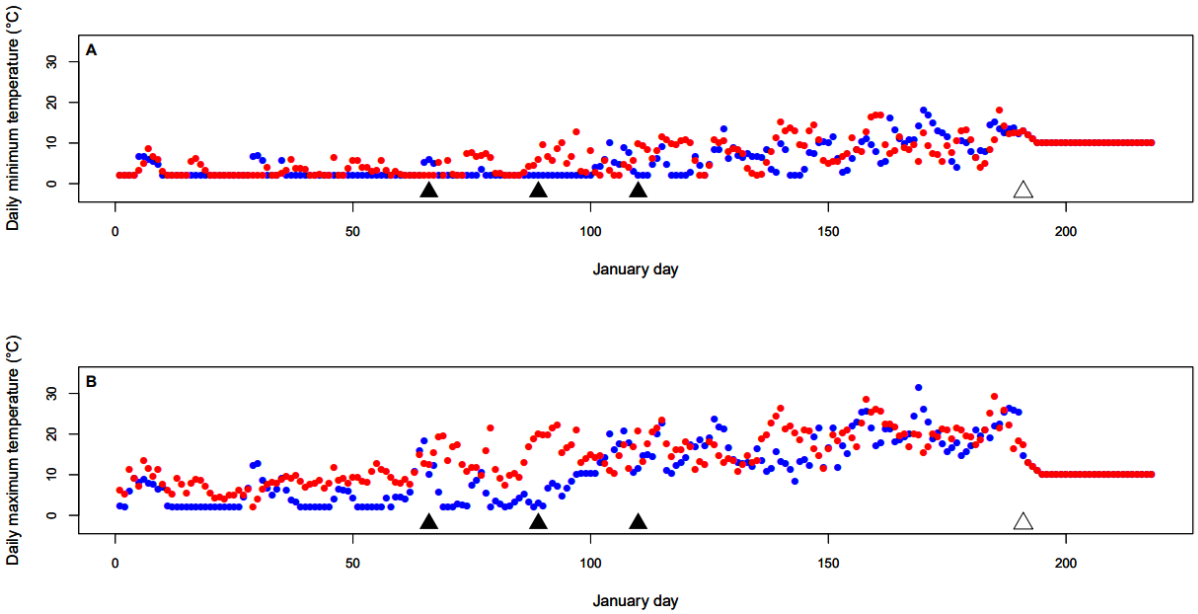


Figure S4.1. Daily minimum (A) and daily maximum (B) temperatures for the cold (blue) and warm (red) spring provided in the first and second breeding season. The open triangle indicates the day on which the first breeding season stopped and birds went into the phase of the experiment where days were shortened and the temperature set at 10 °C (see '*Second breeding season*') in to prepare them for the second breeding season. The black triangles indicate the three time points (66 January = 7 March, 89 January = March 30, 110 January = April 20) on which the birds were sacrificed in the second breeding season.

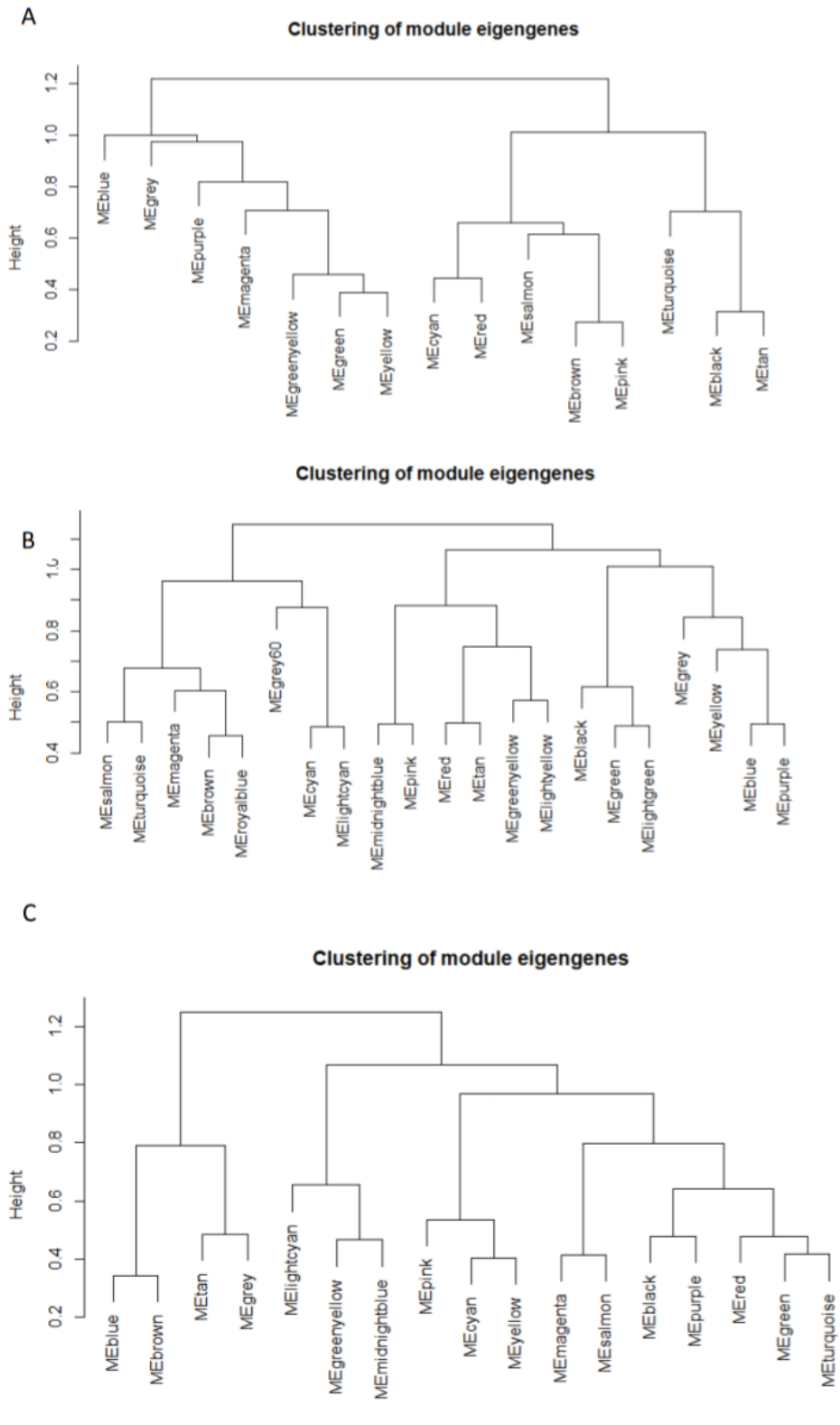


Figure S4.2. Hierarchical clustering tree based on WGCNA module eigengenes in A. hypothalamus, B. liver and C. ovary.

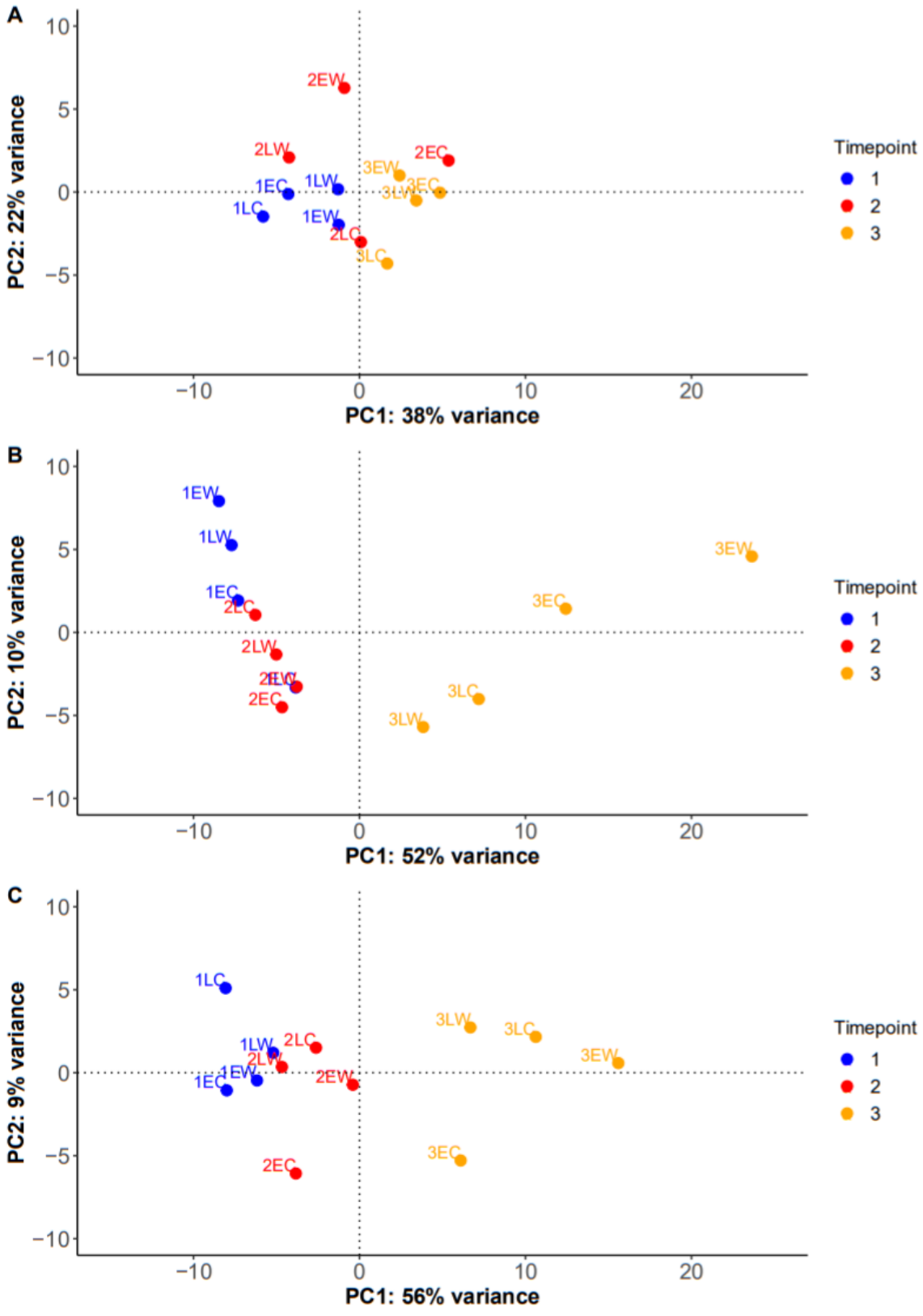


Figure S4.3. Clustering of samples based on principal component analysis (PCA). Samples collected from warm (W) and cold (C) temperature treatments from two different lines, early (E) and late (L), from three different time points and from three different tissues: A. hypothalamus, B. liver and C. ovary.

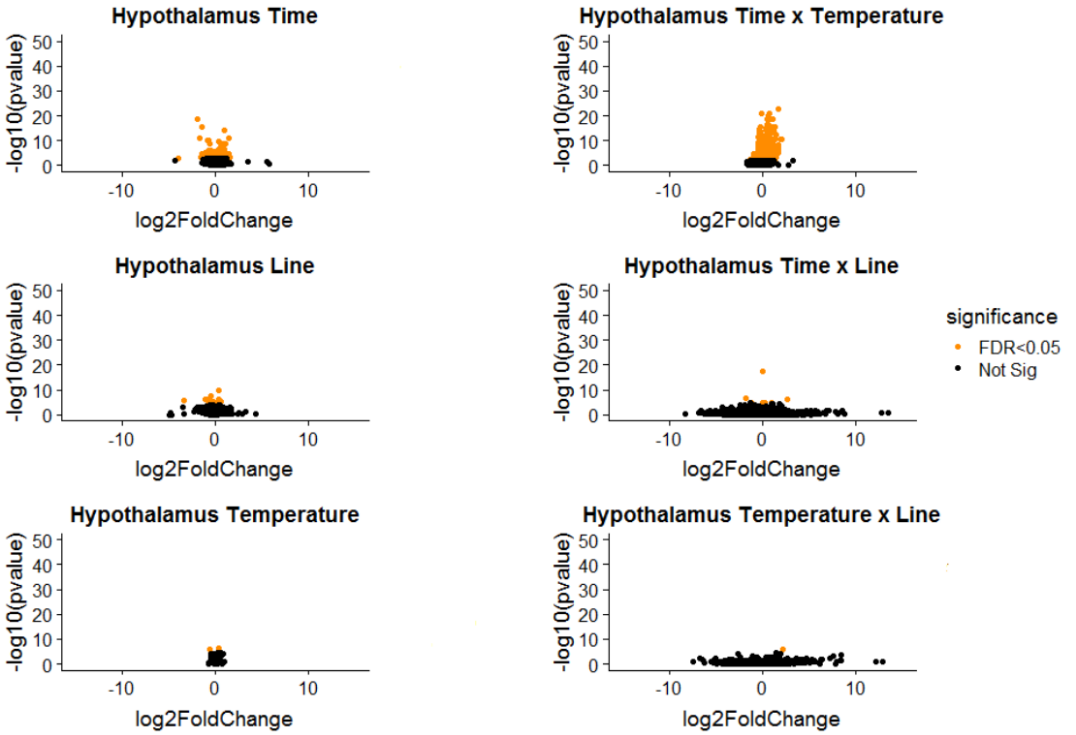


Figure S4.4. Volcano plots of all the transcripts analysed in hypothalamus RNA-seq in six different models. Genes differentially expressed with p<0.05 after correcting for false discovery rate are in orange. Genes with a p>0.05 after correcting for false discovery rate are in black.

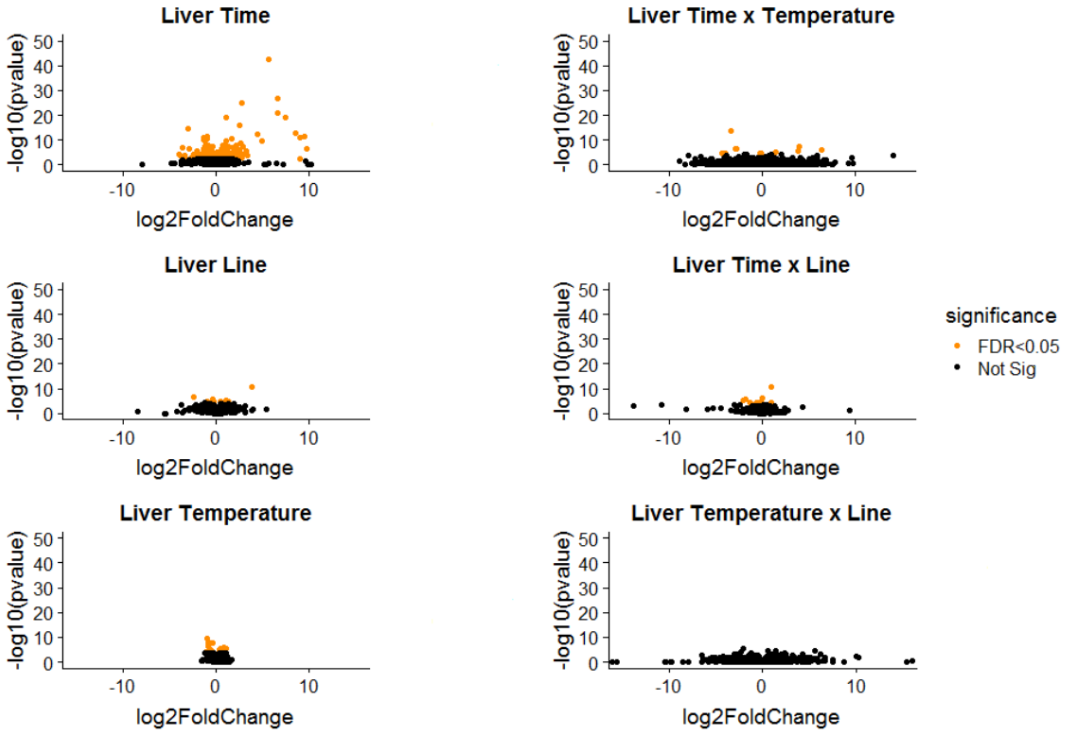


Figure S4.5. Volcano plots of all the transcripts analysed in liver RNA-seq in six different models. Genes differentially expressed with p<0.05 after correcting for false discovery rate are in orange. Genes with a p>0.05 after correcting for false discovery rate are in black.

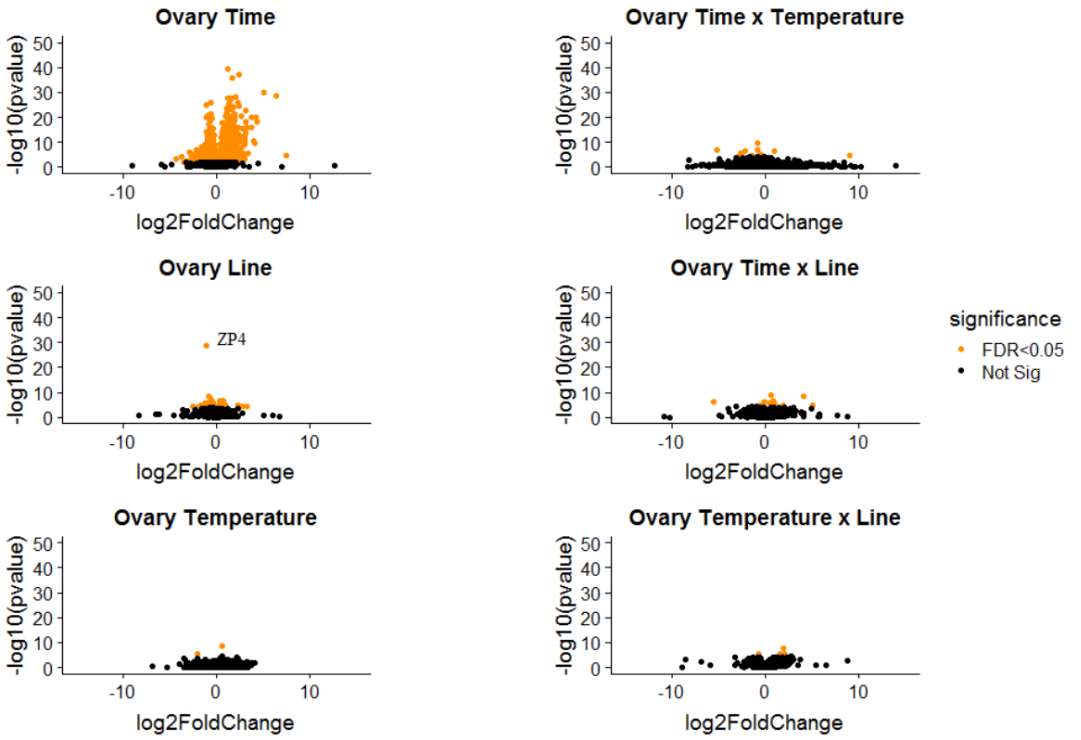


Figure S4.6. Volcano plots of all the transcripts analysed in ovary RNA-seq in six different models. Genes differentially expressed with $p<0.05$ after correcting for false discovery rate are in orange. Genes with a $p>0.05$ after correcting for false discovery rate are in black.

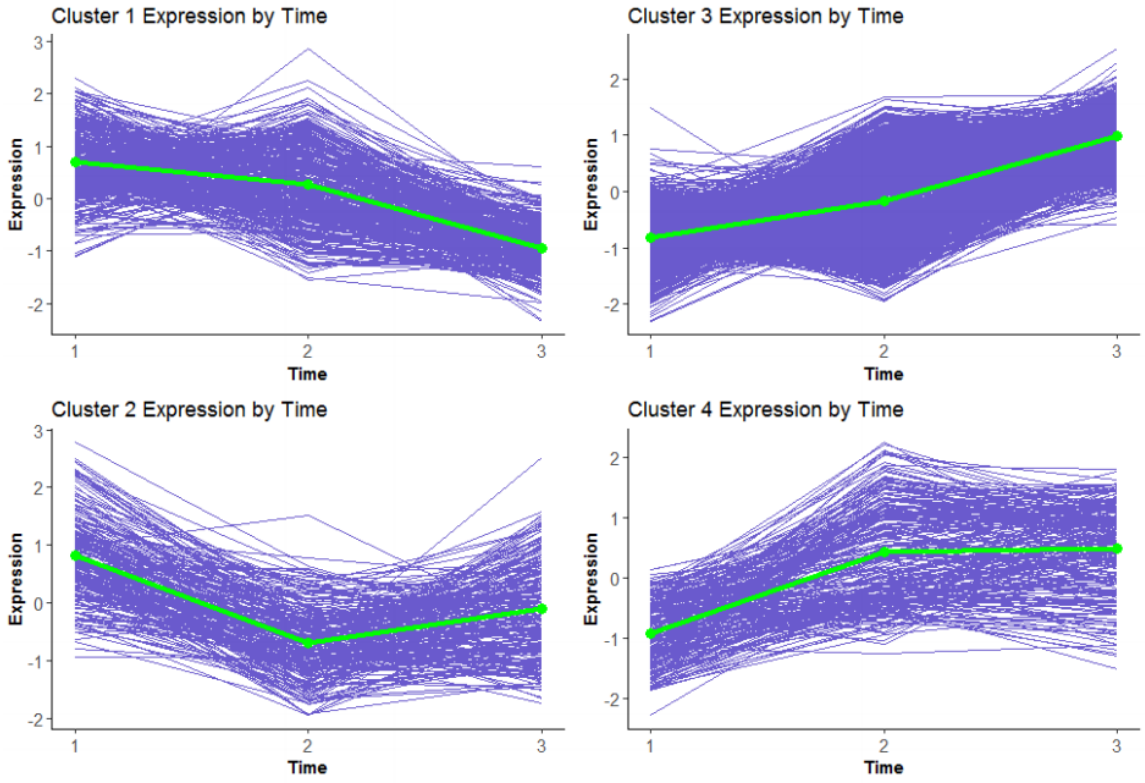


Figure S4.7. Expression patterns of DEG clusters in hypothalamus for the time point main effect model.

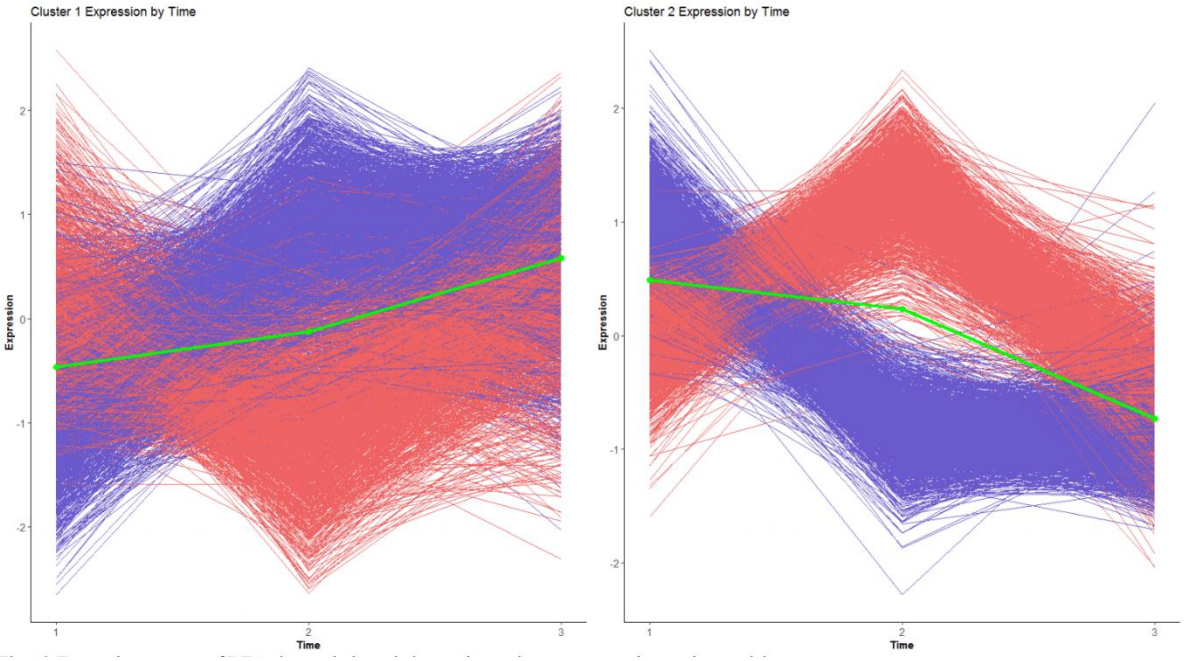


Figure S4.8. Expression patterns of DEG clusters in hypothalamus for the time point-temperature interaction model.

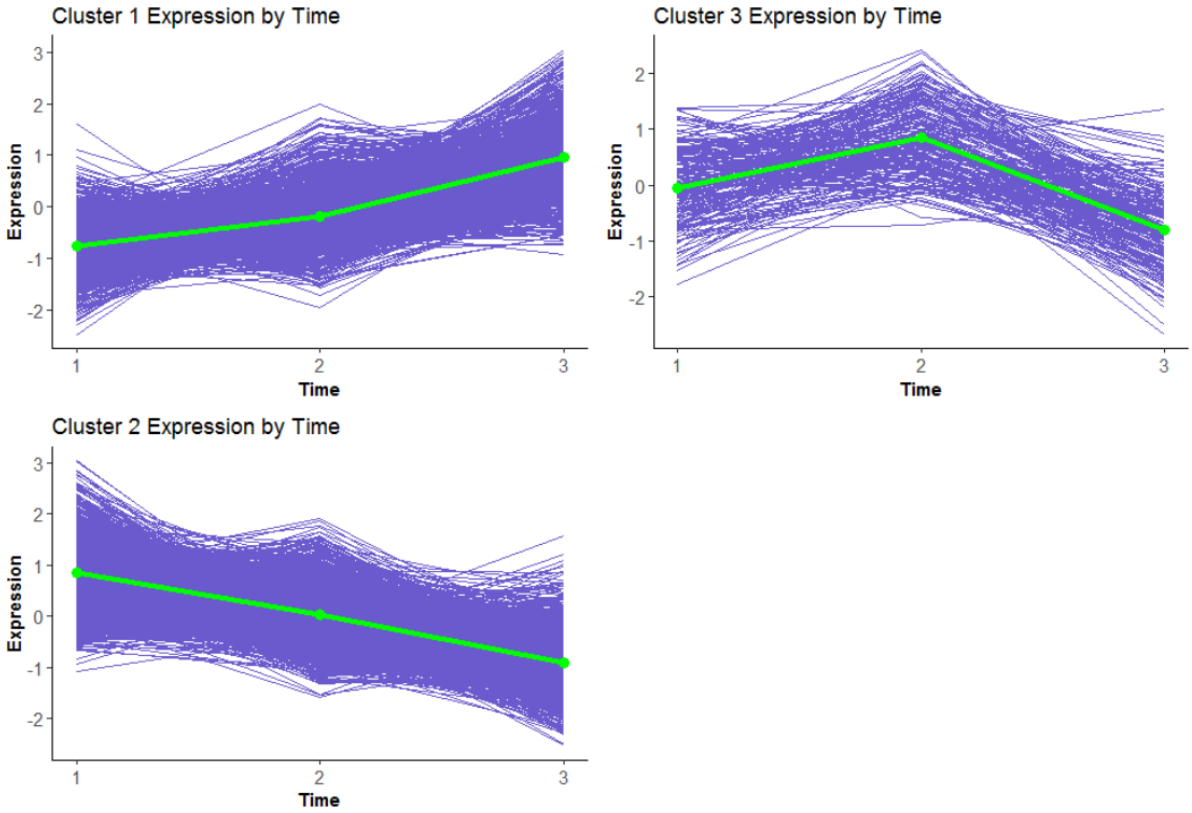


Figure S4.9. Expression patterns of DEG clusters in liver for the time point main effect model.

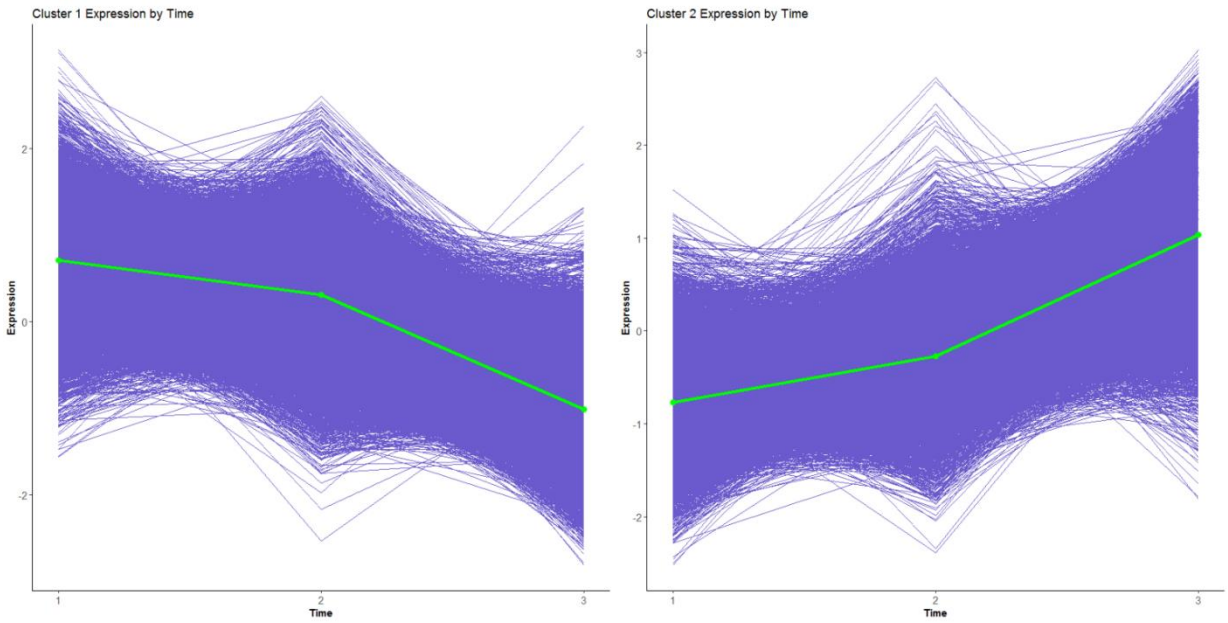


Figure S4.10. Expression patterns of DEG clusters in ovary for the time point main effect model.

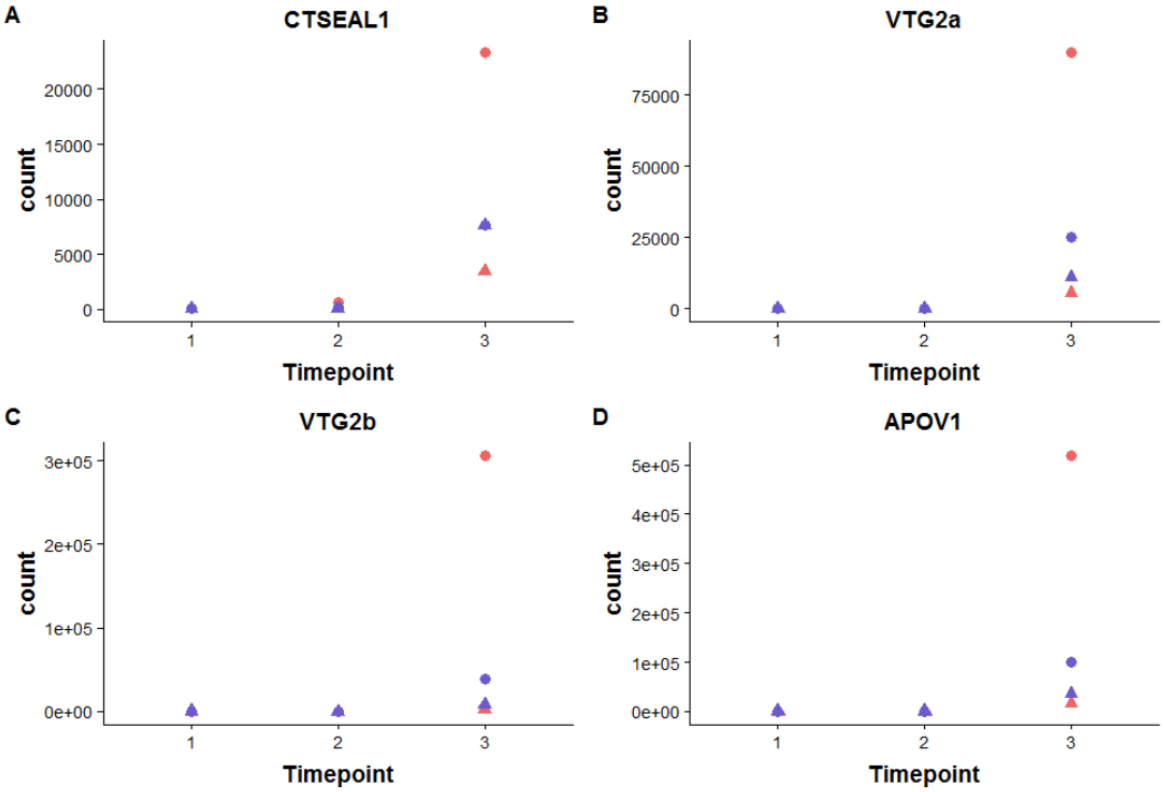


Figure S4.11. The raw expression levels of *CTSEAL*, *VTG2* (VTG2a - LOC107208431and VTG2b - LOC107208432) and *APOV1* in liver for the early (circles) and late (triangles) selection line pools in the warm (red) and cold (blue) treatment.

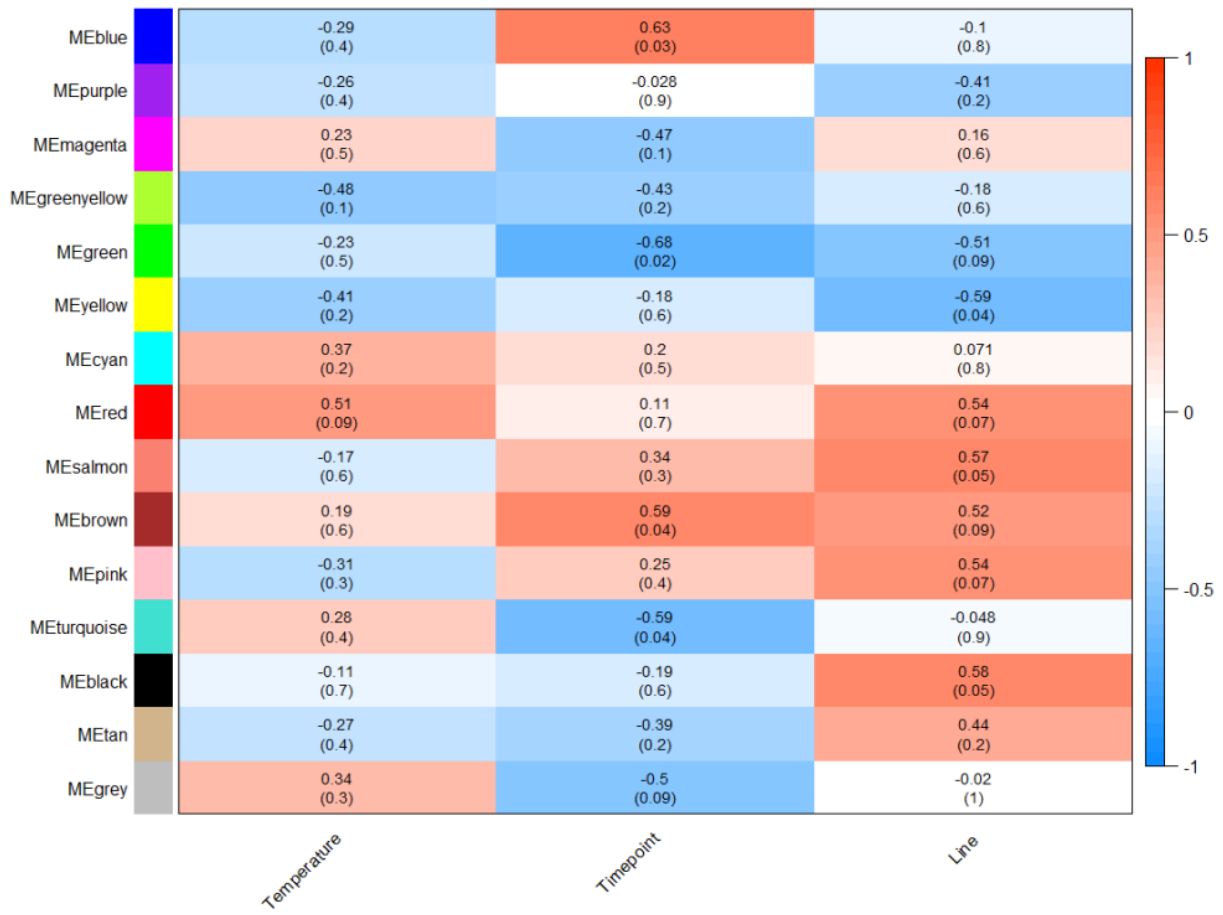


Figure S4.12. Matrix with the module-treatment relationships and corresponding p-values between the detected modules on the y-axis and treatments on the x-axis based on hypothalamus RNA-seq. The relationships are coloured based on their correlation: red indicates a strong positive correlation, while blue is a strong negative correlation. The value at the top of each square represents the correlation coefficient between the module eigengene and the treatment with the correlation p-value in parentheses.

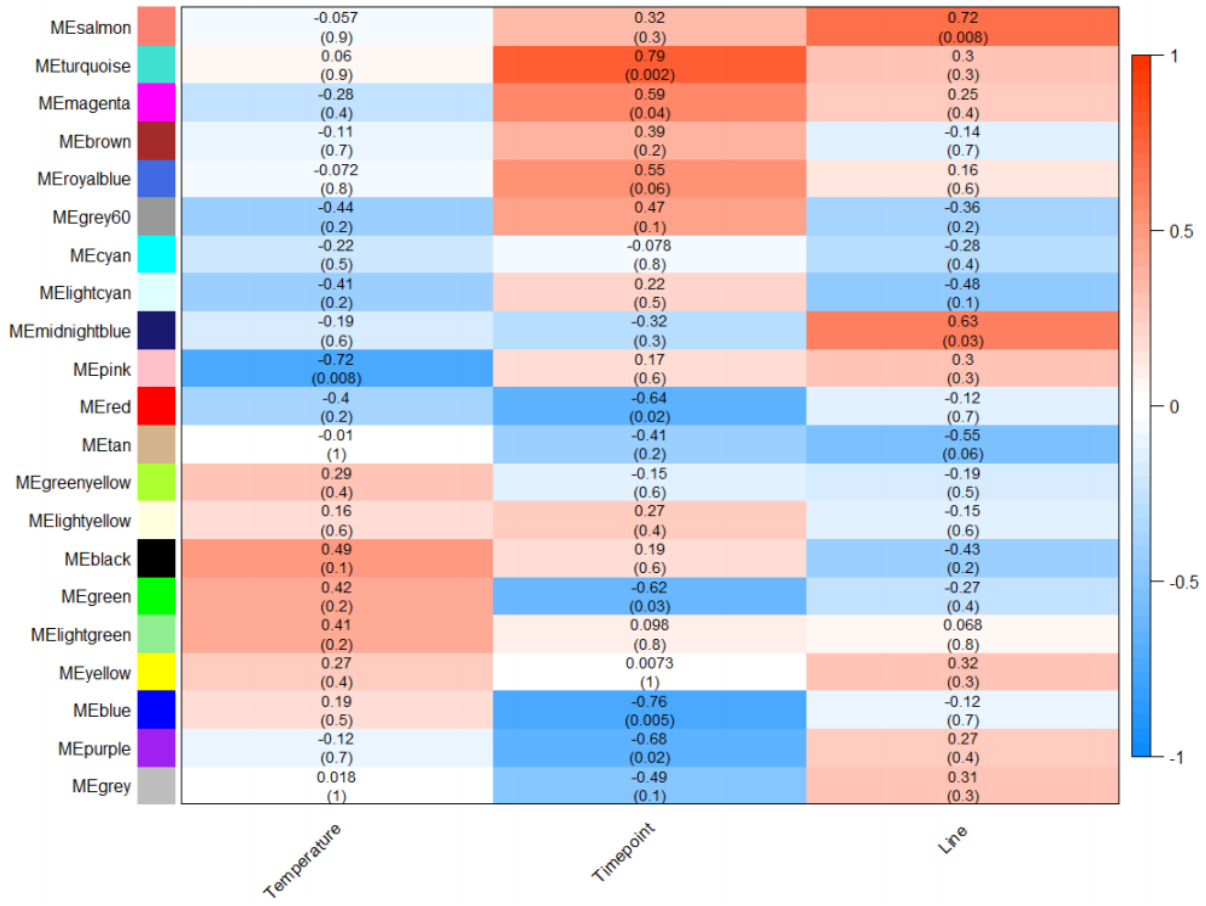


Figure S4.13. Matrix with the module-treatment relationships and corresponding p-values between the detected modules on the y-axis and treatments on the x-axis based on liver RNA-seq. The relationships are coloured based on their correlation: red indicates a strong positive correlation, while blue is a strong negative correlation. The value at the top of each square represents the correlation coefficient between the module eigengene and the treatment with the correlation p-value in parentheses.

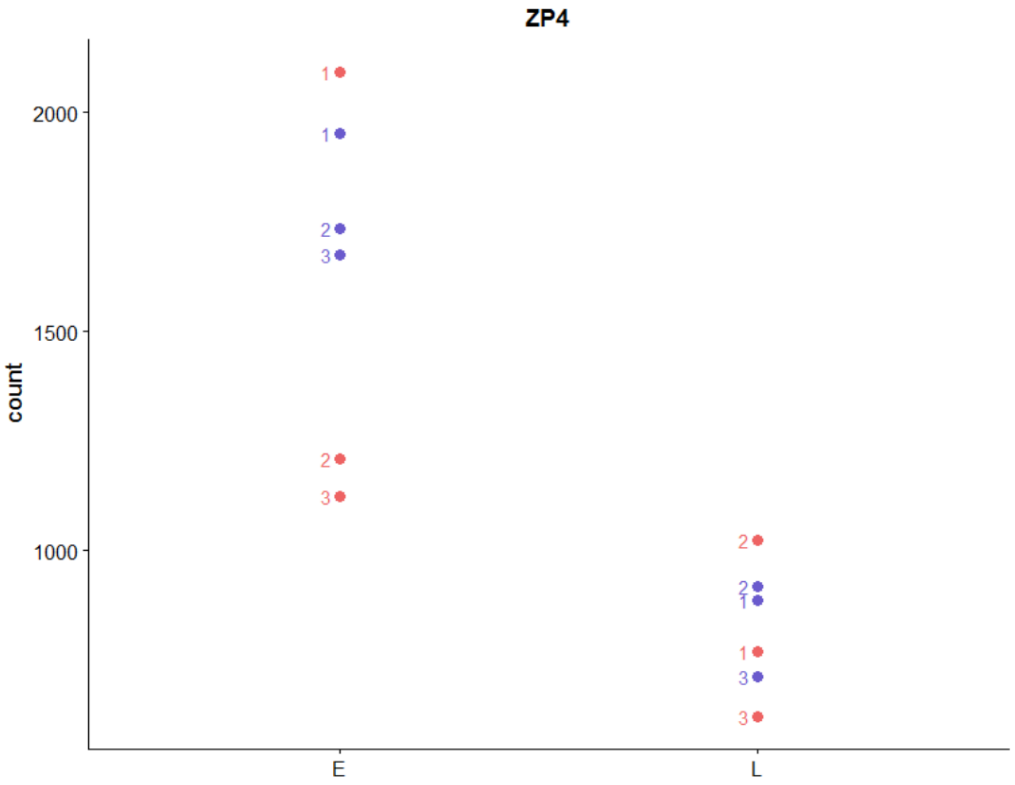


Figure S4.14. The raw expression levels of *ZP4* in ovary for the early (E) and late (L) selection line pools in the warm (red) and cold (blue) treatment. Numbers indicate different time points from the study.

Table S4.1. Summary of the sequencing and alignment of the three tissue types and 12 pools.

Pool	Tissue	Total sequences	Overall alignment %
1EC	Hypothalamus	19,060,146	82.12
1EW	Hypothalamus	19,135,588	82.23
1LC	Hypothalamus	20,002,181	81.39
1LW	Hypothalamus	21,049,529	81.25
2EC	Hypothalamus	13,892,273	82.94
2EW	Hypothalamus	20,671,433	82.42
2LC	Hypothalamus	14,297,648	82.82
2LW	Hypothalamus	17,606,352	81.24
3EC	Hypothalamus	18,954,331	82.80
3EW	Hypothalamus	13,023,076	82.66
3LC	Hypothalamus	19,657,504	82.84
3LW	Hypothalamus	21,905,692	82.52
Average		**18,271,313**	**82.27**
Standard deviation		**2,834,212**	
SUM		**219,255,753**	
1EC	Liver	15,027,929	80.37
1EW	Liver	14,102,973	79.89
1LC	Liver	14,841,357	79.43
1LW	Liver	12,459,889	78.81
2EC	Liver	17,427,338	81.17
2EW	Liver	17,024,244	78.51
2LC	Liver	17,441,069	80.93
2LW	Liver	13,116,994	78.35
3EC	Liver	18,113,293	79.16
3EW	Liver	17,330,580	82.09
3LC	Liver	18,760,366	79.37
3LW	Liver	16,636,408	79.23
Average		**16,023,537**	**79.78**
Standard deviation		**1,961,030**	
SUM		**192,282,440**	
1EC	Ovary	15,627,998	92.09
1EW	Ovary	16,766,194	91.68
1LC	Ovary	18,785,971	92.12
1LW	Ovary	14,169,965	91.51
2EC	Ovary	14,534,915	91.63
2EW	Ovary	10,800,894	91.5
2LC	Ovary	14,083,032	91.68
2LW	Ovary	15,979,825	91.56
3EC	Ovary	15,941,565	90.61
3EW	Ovary	14,631,394	89.27
3LC	Ovary	16,812,073	90.22
3LW	Ovary	13,781,979	91.02
Average		**15,159,650**	**91.24**
Standard deviation		**1,907,655**	
SUM		**181,915,805**	

Table S4.5. The number of genes showing significant differential expression in the time point comparisons and interaction models for the three tissues. Time 1 is comparison of time point 1 to time point 2 and 3, Time 2 is comparison of time point 2 to time point 1 and 3, Time 3 is comparison of time point 3 to time point 1 and 2.

Tissue	Time 1	Time 2	Time 3	Time × Temp	Time × Line	Temp × Line
Hypothalamus	165	2	711	1539	7	1
Liver	152	5	498	12	13	0
Ovary	1090	3	6266	10	14	4

CHAPTER 5

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

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

ABSTRACT

Timing of breeding is currently under selection in wild populations due to climate change, and improved knowledge of underlying physiological processes mediating timing will help understand the potential rate of adaptation. Our current knowledge on this variation in physiology is, however, mostly limited to males. Here, we assess 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 (i.e. hypothalamus), or downstream (ovary and liver) parts of the neuroendocrine system. We used 72 female great tits from the first 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 we sampled several tissues, in the same individuals, 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 three tissues with most pronounced differences in the hypothalamus. Differences between time points, and early and late egg-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.

*Journal of Experimental Biology (2019) **222**: jeb202481*

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). It is widely assumed that the hypothalamus, considered as the final integration point of environmental cues, the pituitary gland, and the neural centres 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 (**Chapter 4;** Cánovas et al., 2014; MacManes et al., 2017; Maruska & Fernald, 2011; Maruska et al. 2011; Perfito et al. 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 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 shortly 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 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 females' "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; **Chapter 2**). 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 (**Chapter 4**). 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 **Chapter 4**, 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 (**Chapter 4**).

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; **Chapter 2** 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 (**Chapter 2**). Briefly, we found that on average early line birds laid about six days earlier than late line birds. Further, the difference in average egg-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 generation (**Chapter 2**).

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'*). As opposed to outdoor aviaries (**Chapter 2**), neither selection line, temperature environment, nor their interaction, explained females' reproductive phenotypes (i.e. egg-laying dates and follicle widths) in these climate-controlled aviaries (Appendix 5.1). 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.

Experimental setup

F_1 generation (n = 36) and F_2 generation (n = 36) selection line of great tits(Gienapp et al. 2019; **Chapter 2**) were housed in 36 climate controlled aviaries (2m × 2m × 2.25m) at the Netherlands Institute of Ecology (NIOO-KNAW) in 2015 and 2016, respectively. Birds were subjected to an artificial photoperiod mimicking the change in natural photoperiod. Per aviary light was provided by one full spectrum daylight fluorescent lamp (58W, 5500K,

True-light, The Netherlands) and two fluorescent lamps (58W, Philips, The Netherlands). A roof shaft (SolaTube) provided additional natural light (total average daily light intensity ~500 lux per aviary, Table S5.1). A light bulb (7W, Philips, The Netherlands) mimicked dawn and dusk, which turned on half an hour before lights went on and stayed on half an hour after lights went off respectively (Caro & Visser 2009). In addition, breeding pairs were subjected to two contrasting environments mimicking an extreme cold spring (2013) and an extreme warm spring (2014) in the Netherlands (Figure S5.1): average egg-laying dates were May 5 ± 5.18 days (n = 112) and April 11.8 ± 5.46 days (n = 124) for 2013 and 2014, respectively in the wild long-term study population at the Hoge Veluwe. Temperatures changed every hour to follow as closely as possible the observed hourly temperatures in these years (note that the minimum temperature in the aviaries was 2°C so any temperature below 2°C in the temperature time series from outside were set to 2°C). The combination of selection line and temperature environment resulted in four groups of nine pairs: 'early-warm', 'early-cold', 'late-warm' and 'late-cold'. We like to state here again, that the variables selection line and temperature environments are left out of further analyses (see above), but are mentioned here to explain the experiment. Birds were fed *ad libitum* with food sources reported elsewhere (Visser et al. 2011) and had water available for drinking and bathing. All pairs went through two experimental breeding cycles; a 'first breeding season' and a 'second breeding season' (see below and Figure S5.2). This study was performed under the approval by the Animal Experimentation Committee (DEC), Amsterdam, The Netherlands, protocol NIOO 14.10 addendum 1.

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. egg-laying date) were recorded. Egg-laying dates were recorded as April dates (i.e. 31 March = 0, 1 April = 1, 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 moulting (~mid-July), days were shortened to 9L:15D and temperatures decreased to 10°C for seven weeks to make the birds photoreceptive 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 (in 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 et al. 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 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 for tissue collection (see '*Tissue collection and preparation*').

Tissue collection and preparation

For both generations, pairs were categorized in three groups (n = 12 pairs per group) based on their egg-laying dates from the first breeding season, resulting in groups with a roughly similar average egg-laying date and distribution (Figure S5.3). Three time points were chosen based on the egg-laying dates in 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 et al. 1993), (2) October 28 (i.e. March 30) when nest building occurred in the first breeding season, but prior to egg-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 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://www.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.

Reverse transcription 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, for hypothalamus and liver and ovary respectively. 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 (**Chapter 4**, Table S5.2). 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 S5.2). Primers were then built based on the great tit reference genome build 1.1‡ ((Laine et al. 2016) and annotation release 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.

qPCR 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 S5.3. 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).

‡ https://www.ncbi.nlm.nih.gov/assembly/GCF_001522545.2
§ https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Parus_major/101/

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 5.2). 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 et al. 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 egg-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 egg-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. 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). Significant differences were determined by performing ANOVA, with the following model PC_x ~ time point × egg-laying date + generation. P-values were adjusted for multiple comparisons using Benjamini and Hochberg's False Discovery Rate (FDR) (Benjamini & Hochberg 1995), accepting an FDR of 0.05. Females did not differ in either egg-laying dates or largest follicle widths between selection lines, temperature treatments or their interaction within a generation (see above, Appendix 5.1). 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 **Chapter 4**, 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 S5.4) 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 $expression_{gene}$ ~ time point × egg-laying date (from the first breeding season) + generation. All statistics were performed in R (version 3.3.1).

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 egg-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 egg-laying date and largest follicle (Figure S5.4A). When analyzing per time point (Figure S5.4B), the relationship between egg-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 relationship at time point 3 ($r = -0.66$, $F_{1,20} = 15.18$, $p < 0.001$). In addition to the significant difference in follicle widths, we are confident that the mRNA expression levels from the second breeding season are representative of the egg-laying dates recorded (Appendix 5.3).

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 S5.5-S5.7).

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 S5.5). In addition, the similar loadings (Table S5.5) and the small angle between the vectors of *OPN5* and *DIO2* (Figure 5.1A) 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 S5.8), as shown by two distinct, but overlapping clusters along PC1 in hypothalamus (Figure 5.1A). No distinction in expression profile was found when clustering females per 'time point' for PC1 or PC2 (Figure 5.1B), nor did time point explain variance in any of the PCs (Table S5.8) and also no association between expression of these genes and the interaction between egg-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 5.1C, Table S5.6). 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 1 panel C, Table S5.6). Although not shown in Figure 1, but based on similar loadings in PC1 and 2 (Table S5.6), 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 1 panel C, Table S5.9) 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 5.1D, Table S5.9). PC1 and PC2, together accounting for ~48% of the total variance, are highly significantly associated with both generation and time point (Table S5.9).

Liver. PC1, accounting for ~46% of the variance among females, is associated with an egg-laying date × time point interaction ($F_{2, 53} = 11.019$, FDR corrected $p < 0.001$, Table S5.10) 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 S5.7). Generation explains the variation in gene expression (~28%) along PC2 ($F_{2, 53} = 38.09$, FDR corrected $p < 0.0001$, Table S5.10). Although not shown in Figure 5.1, but based on similar loadings in PC1 and 2 (Table S5.7), *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 5.1E). Similar to ovary, but again along opposite PCs, females show a gradual change in expression profile over time points (Figure 5.1F).

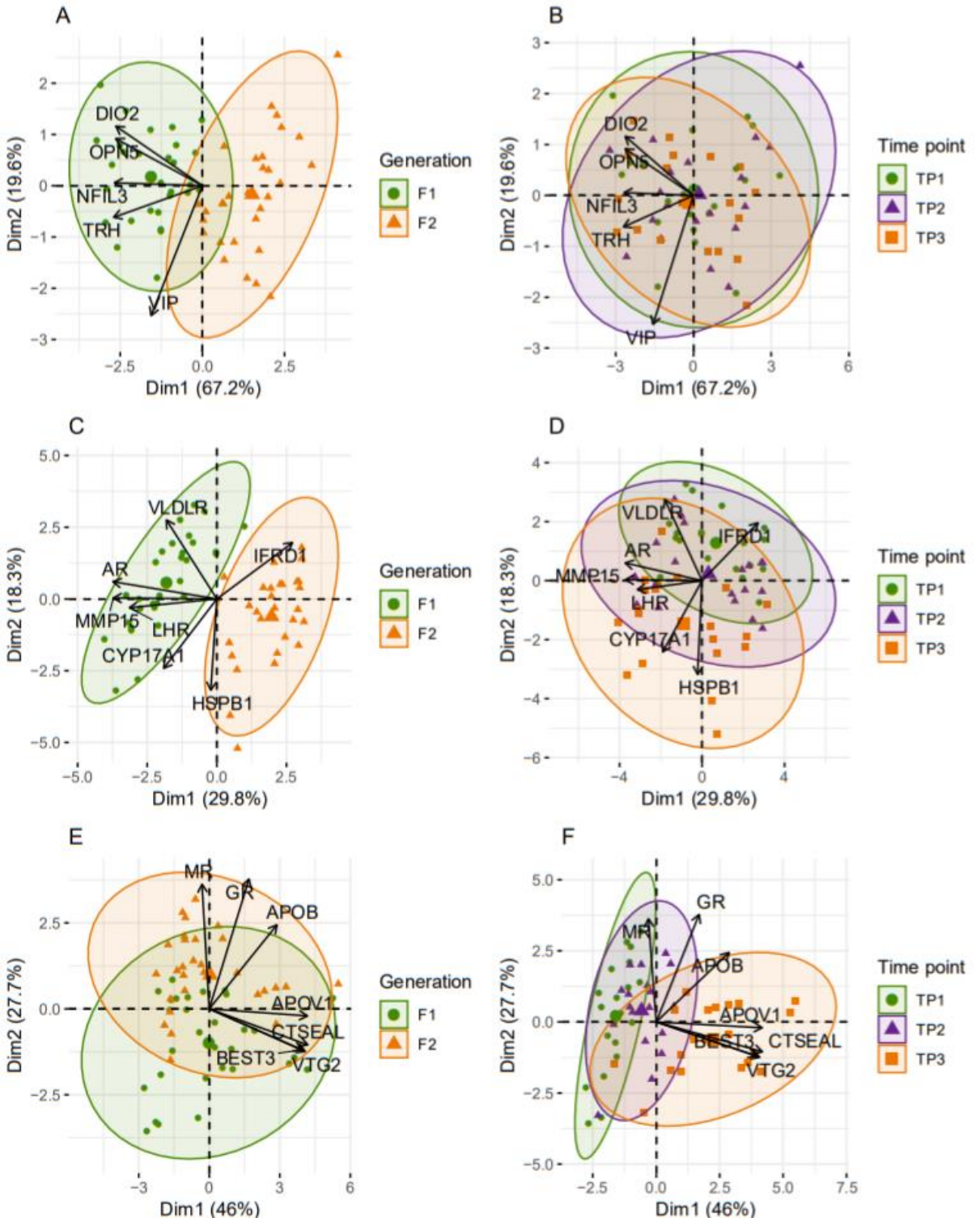


Figure 5.1. PCA was performed on normalized and subsequently transformed gene expression data from individual females for hypothalamus (A, B), ovary (C, D) and liver (E, F). Every data point represents an individual female. Potential clusters associated with generation are shown (panels A, C, E); green = 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).

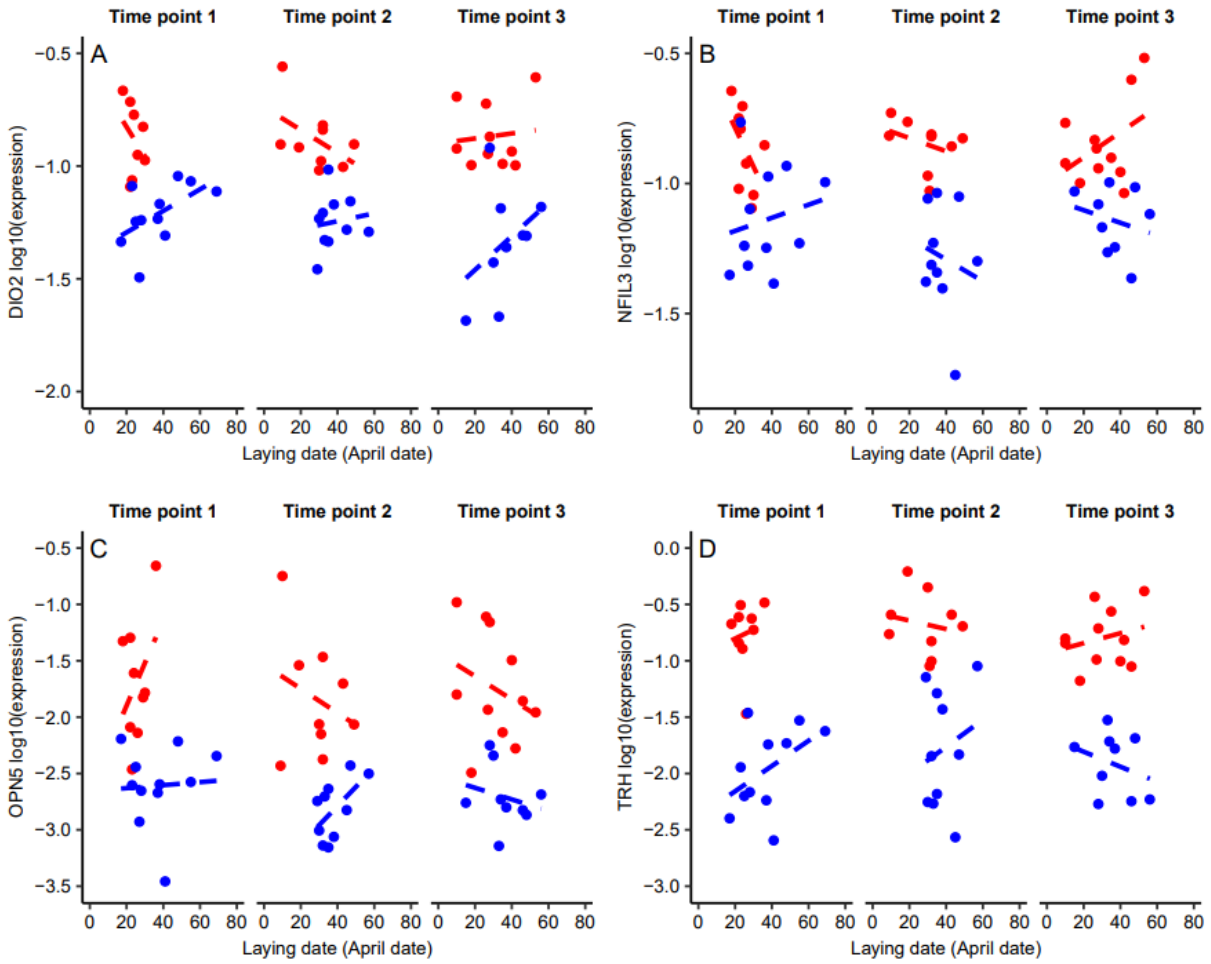


Figure 5.2. Normalized and subsequently log10 transformed mRNA levels of candidate genes in hypothalamus in three time points. Females from the F_1 generation (red) have significantly higher expression levels in hypothalamus for *DIO2* (A), *NFIL3* (B), *OPN5* (C) and *TRH* (D) compared to the F_2 generation (blue) in the three time points, but independent of time point (Table 5.1). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

Variation in hypothalamic, ovarian and liver candidate gene expression

Hypothalamus. We found no differences in candidate gene expression in hypothalamus between time points, egg-laying dates or their interaction (Table 5.1). 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 5.2, Table 5.1). We found no difference in expression levels of *VIP* between generations (data not shown).

Table 5.1. The degree of variation per gene explained by egg-laying date, time point, their interaction and generation in hypothalamus. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

	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

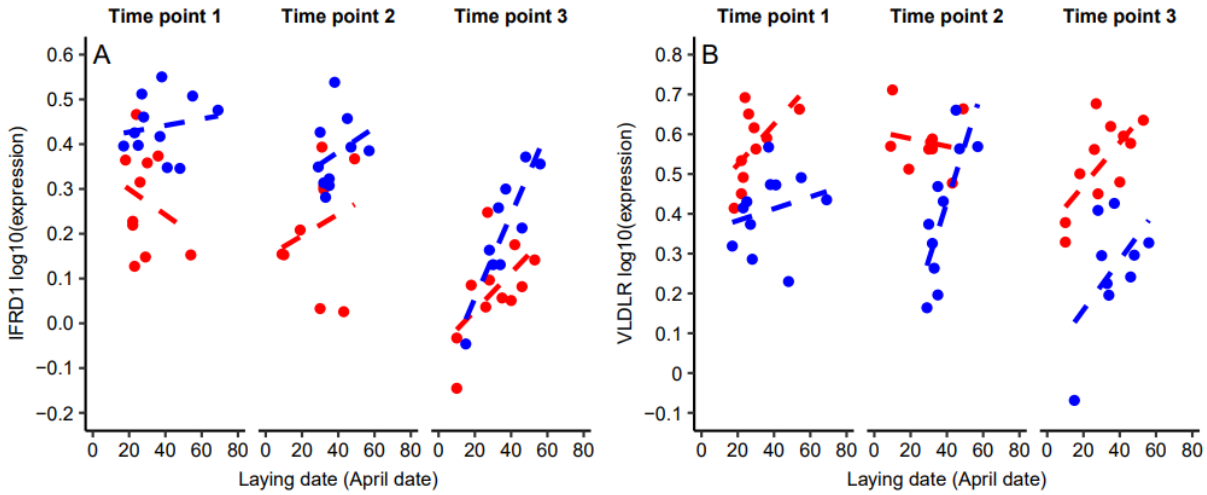


Figure 5.3. Normalized and subsequently log10 transformed mRNA levels of candidate genes in ovary in three time points. F1 generation (red) females have lower expression levels for *IFRD1* (A), but higher expression levels for *VLDLR* (B). Early breeding females have lower expression of *IFRD1* in time points 2 and 3 and late egg-laying females show and increased *VLDLR* expression in time point 2 (Table 5.2). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

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 5.2), but only variation in mRNA expression of *IFRD1* and *VLDLR* 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 $p < 0.001$, Figure 5.3, Table 5.2).

Table 5.2. The degree of variation per gene explained by egg-laying date, time point, their interaction and generation in ovary. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

	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**

Liver. Early breeding females show increased mRNA expression for both insulin-like growth factor 1 (*IGF1*) and *VTG2* in liver (*IGF1*: $F_{1,55}$ = 6.53, FDR corrected p = 0.03, *VTG2*: $F_{1,58}$ = 6.62, FDR corrected p = 0.03, Figure 5.4, Table 5.3) compared to late breeding females. Only in liver, we found differences in mRNA expression levels explained by egg-laying date × time point interactions (Figure 5.5, Table 5.3). 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 early egg-laying females compared to late egg-laying in time points 2 and 3.

Pairwise correlations between gene pairs

Within and among the tissues examined, candidate genes, whether they reflect differences in timing or not, tend to change in a strong and/or significantly similar way (Figure 5.6). For example, CYP17A1 expression in the ovary tends to change in a strong and similar way as APOV1, CTSEAL and VTG2 in liver. In addition, expression of HSPB1 in ovary resembles that of APOB and APOV1 in liver. The mRNA expression of GNIHR in ovary shows a weak positive, however significant, correlation with VTG2 in liver. Interestingly, the genes examined in the hypothalamus show a high and significant correlation among each other, but less so when correlated to genes in the ovary and liver. Between the ovary and liver, more genes tend to change in a similar way, both positively and negatively.

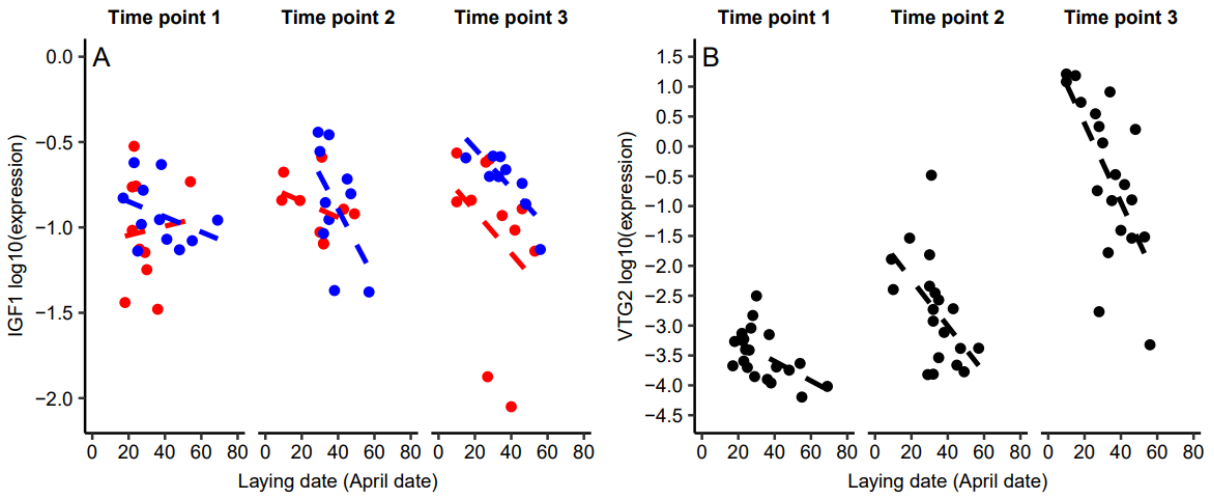


Figure 5.4. Normalized and subsequently log10 transformed mRNA levels of candidate genes in *liver* for three time points. Females differ in expression of *IGF1* (A) between generations (F_1 = red, F_2 = blue), but not *VTG2* (B, no distinction is made between generations, data shown in black). Early egg-laying females show increased expression compared to late egg-laying females in both *IGF1* and *VTG2*, independent of time point (Table 5.3). Regression lines are dashed, as not to imply significant interactions. Note that the scale of the y-axis differs per panel.

Table 5.3. The degree of variation per gene explained by egg-laying date, time point, their interaction and generation in liver. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. 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

Discussion

Gene expression dynamics within the HPGL axis have not been well studied in seasonally breeding females. Using a candidate gene approach, we set out to determine whether individual differences in egg-laying dates (obtained from the first breeding season) are reflected in differences in candidate gene expression levels, and if so, where these differences occur in the HPGL axis; upstream and/or downstream and when these differences can be picked up towards the expected egg-laying dates. We found significant differences in mRNA expression of candidate genes between generations in all three tissues examined. However, a correlation of candidate gene expression and egg-laying date (at the

three sampling time points) was found exclusively in ovary and liver, independent of generation. In particular, individual differences in timing of breeding in females are significantly reflected in mRNA expression for *IFRD1* and *VLDLR* in ovary and *IGF-1* in liver, and earlier breeding females show increased expression of *APOB*, *APOV1*, *BEST3* and *CTSEAL* over time in liver. These findings, 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.

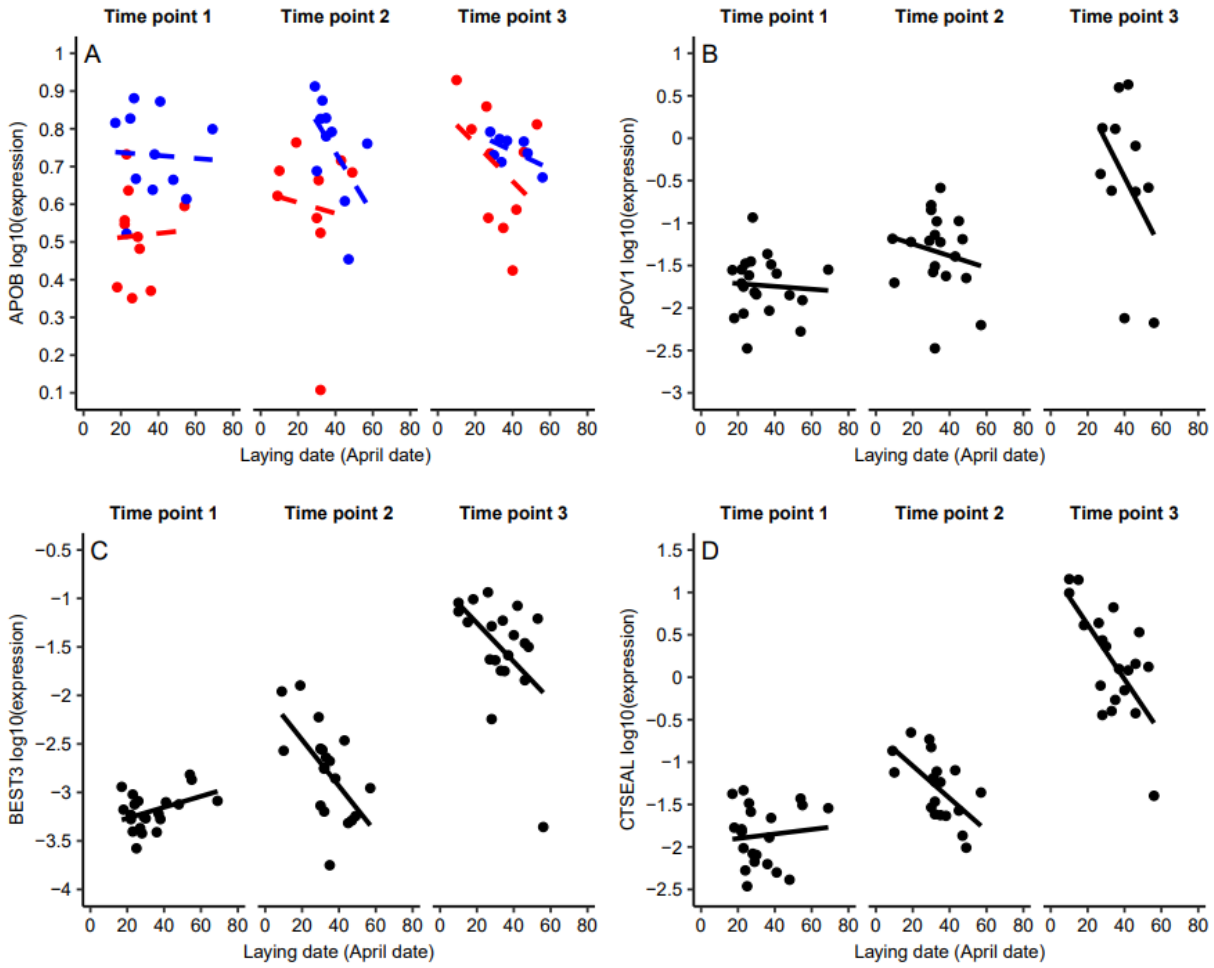


Figure 5.5. Normalized and subsequently log10 transformed mRNA levels of candidate genes in *liver* in three time points. Variation in egg-laying date, in interaction with time point, explains mRNA expression for *APOB* (A), *APOV1* (B), *BEST3* (C) and *CTSEAL* (D, Table 5.3). In addition, *APOB* expression 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.

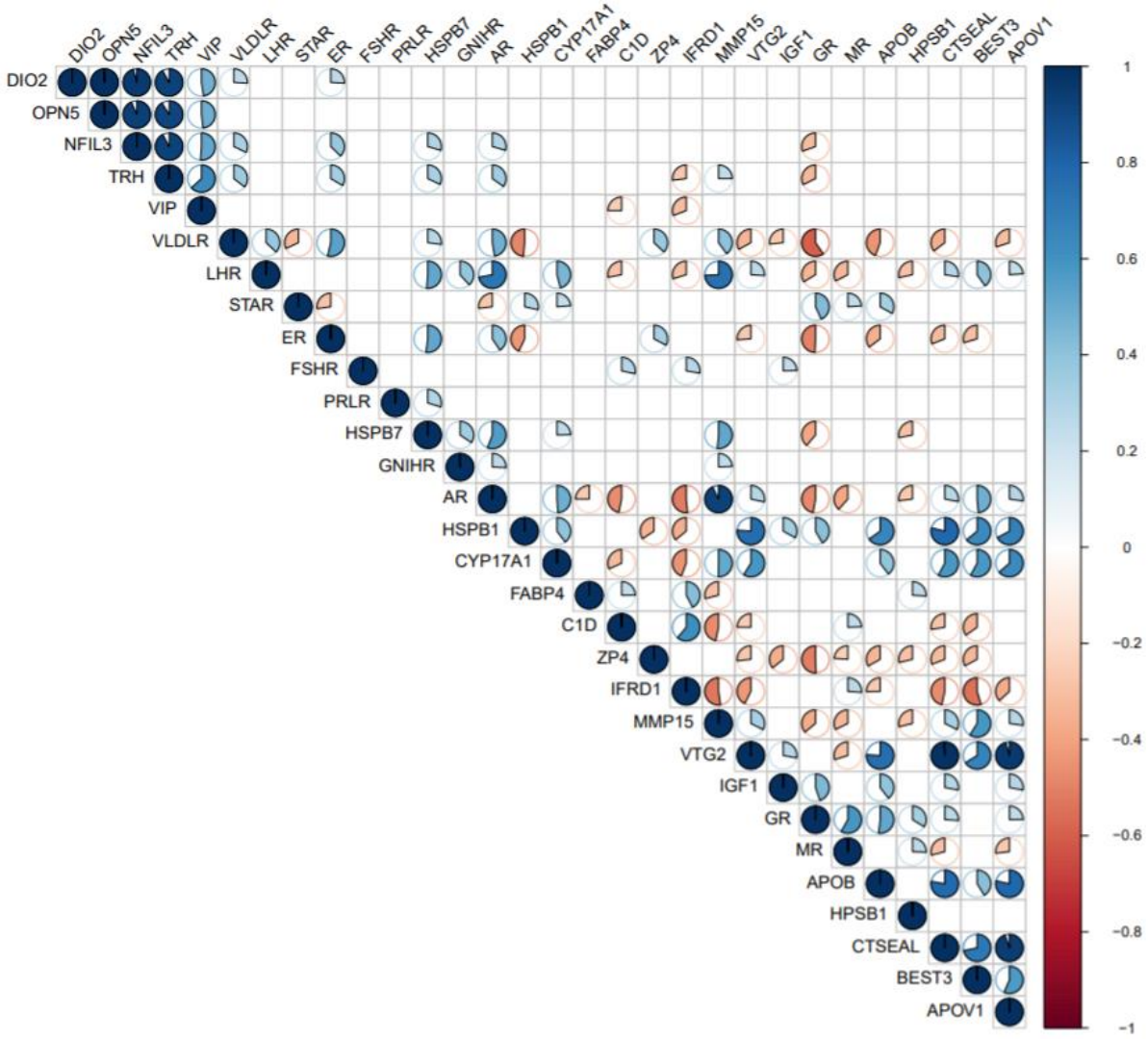


Figure 5.6. Correlation matrix of the pairwise correlations between all possible gene pairs. Positive correlations are displayed in blue and negative correlations in red. Colour intensity and the size of the circle are proportional to the correlation coefficients (see the colour legend on the right). Only the significant (p<0.05) correlations are shown. Genes *DIO2* 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.

Experimental limitations

We compared gene expression levels at different time points approaching egg-laying the first egg, but with different individuals per time point. The limitation here is, that an individual female could not be measured for 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 egg-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 (**Chapter 4**), and that several females had entered RYD or initiated egg-laying (Appendix 5.3), 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 (**Chapter 4**), but not in the ovary and liver in the same samples as used here of the F_2 generation females. 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 et al. 2012; Visser et al. 2009), but agrees with other studies where gonadal size is not affected by ambient temperature (Schaper 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 significant lower expression levels of *DIO2*, *NFIL3*, *OPN5*, and *TRH* in all three time points compared to the F_1 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 **Chapter 4**, but significantly for the early egg-laying females (Figure 5.3) . 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 et al. 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 et al. 1987) and highest in pre-laying females (i.e. time point 3) (Han et al. 2009).

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 (**Chapter 4**). *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 comparative 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 et al. 2012) and may play a role in processing egg yolk macromolecules (Bourin 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* (**Chapter 4**; Zheng et al., 2018).

We found expression of *IGF1* to reflect individual differences in egg-laying, with early egg-laying females showing higher *IGF1* expression compared to late egg-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 5.6), 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

Currently, one can only speculate on where the 'switch' that initiates egg-laying resides within the ovary and/or liver. A potential candidate is the 'competence' of the ovary to respond to gonadotropins via their receptors (Ball 2007; Caro et al. 2009; Johnson 2015a; Partecke, Van't Hof & Gwinner 2005; Schaper et al. 2012; Williams 2012). Further, in starlings, it has also been shown that sex steroid secretion can be regulated by local GnIH in the gonads (McGuire et al. 2011). As such, the gonadal GnIH system could be a potential mechanism in timing of breeding in females (McGuire et al. 2011; Needham et al. 2019). Another potential mechanism is the communication between the ovary and liver, where the E2-dependent shift in lipid metabolism or the up-regulation of VTG/VLDL-receptors could be candidates. These potential mechanisms, however, need to be regulated, and imply a more autonomous role for the ovary together with receiving signals that bypass the classic neuro-endocrine pathway. As such, the ovary and brain might act more as 'partners' (Ball 2007). For example, a study in Japanese

quail (*Coturnix japonica*) suggests that the ovary regulates her own functioning through its circadian clock, because the largest follicle, through production of circadian clock gene proteins, controls the LH surge that is essential for ovulation (Nakao et al. 2007).

Outlook

The exact downstream mechanisms that precede timing of breeding and how they are regulated remains to be determined. Nevertheless, we have shown that variation in mRNA expression levels of several candidate genes in ovary and liver, associated with reproductive functioning, explain variation in timing of breeding in these females. Our study confirms that shifting the focus a bit more towards *females* rather than males (Caro et al., 2009, Williams, 2012), and towards the ovary or follicles and liver rather than the hypothalamus in future experimental studies is highly important. Also, simultaneous examination of multiple, and preferably all, HPGL axis levels is of the essence in understanding mechanisms underlying timing of breeding. This way, we gain knowledge on the variation in the physiology underlying timing of avian breeding and what part of this variation is 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. 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).

Supplementary Information Chapter 5

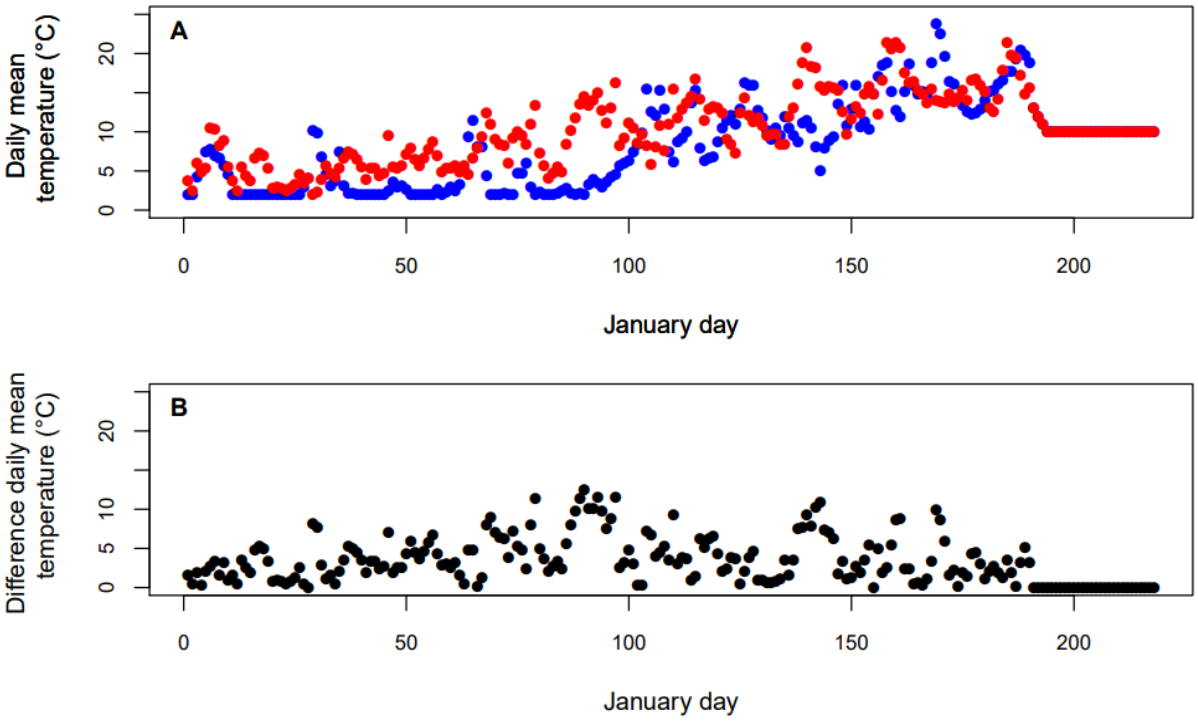


Figure S5.1. Daily mean realized temperatures (A) for the cold treatment (blue) and the warm treatment (red) and the difference in daily mean expected temperatures between the cold and warm treatment (right panel) shown from 1 January (January day 1) until 6 August (January day 218). The cold and warm treatments reflect daily mean realized temperatures from an extremely cold (2013) and warm (2014) spring respectively, in The Netherlands. Differences increase nearing the breeding season, but decrease again during and after (B). From 10 July (January day 191) onwards, birds received the same temperatures, indicated by the overlapping circles and flattening line, because they were transitioned to and entered the period with short days and low temperatures at 14 July (January day 195).

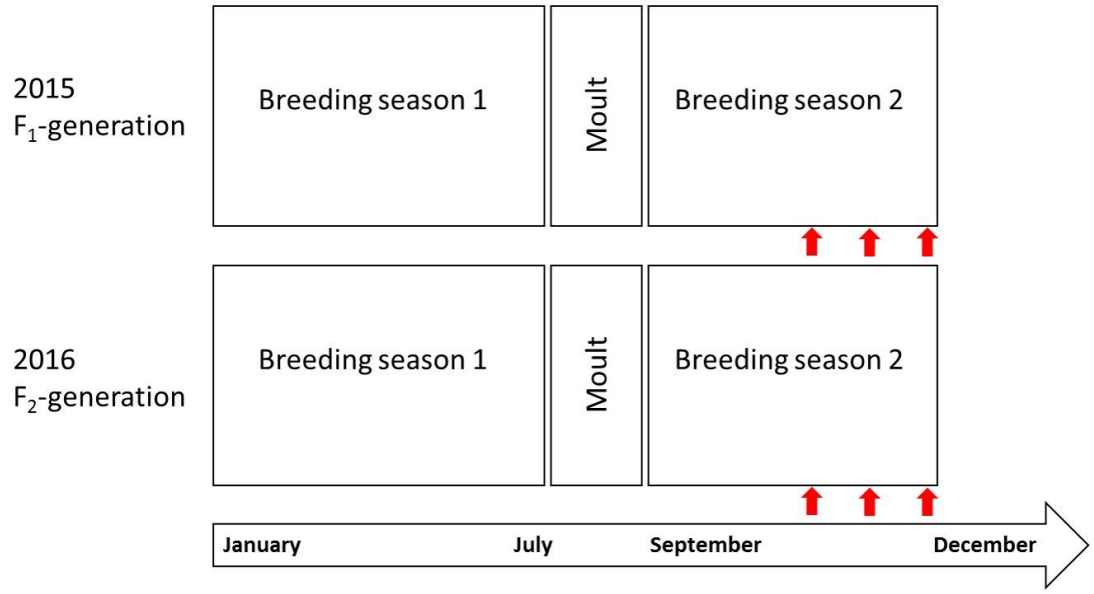


Figure S5.2. Representation of the two breeding seasons within one year, which are identical for the F_1 generation in 2015 and their offspring, the F_2 generation, in 2016. The red arrows indicate the three time points at which birds were sacrificed.

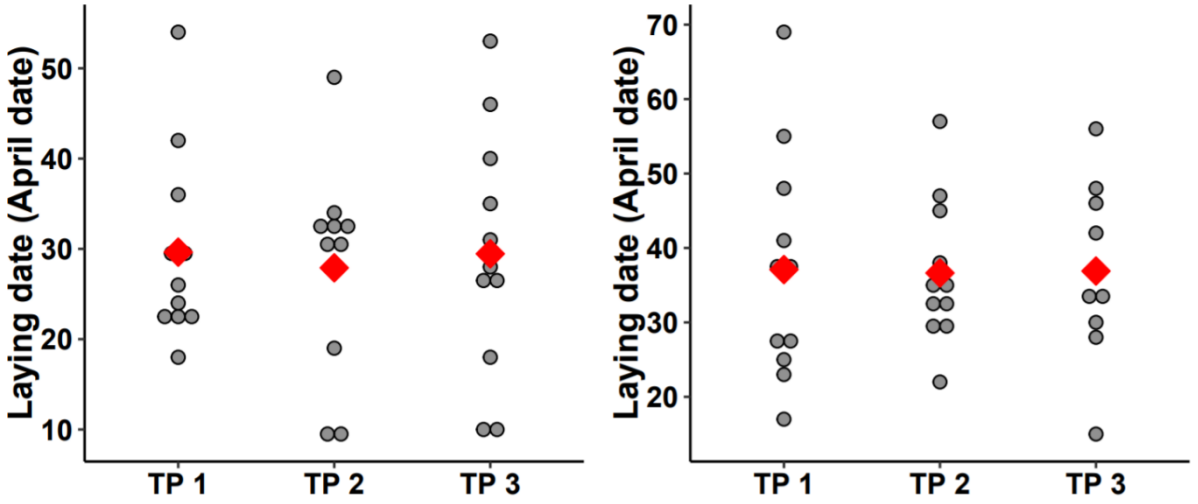


Figure S5.3. Distributions (grey dots) and means (red diamonds) of egg-laying dates (y-axis) in the groups for the three time points (x-axis) for 2015 (A) and 2016 (B). Egg-laying dates are shown as April dates (10 = 10 April, 40 = 10 May etc.). Means between groups within a year were not significantly different (2015; $p>0.730$ for all t-tests, 2016; $p>0.930$ for all t-tests).

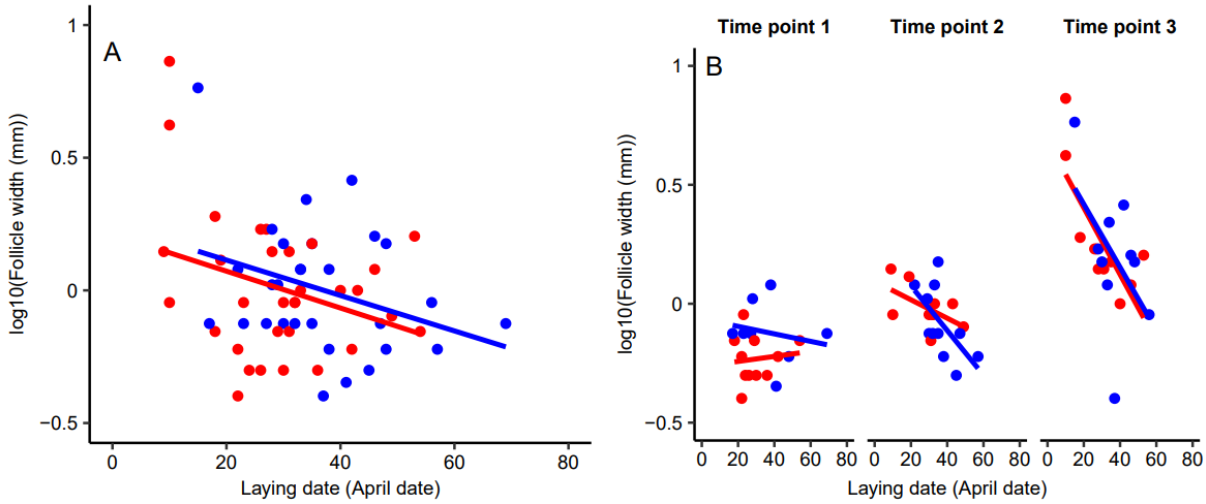


Figure S5.4. Relationships between egg-laying dates (x-axis) and log10 transformed widths of the largest follicle measured (y-axis) for all females (A), shown per generation (F_1 = red, F_2 = blue) and per time point, shown per generation (B). Every dot represents an individual female.

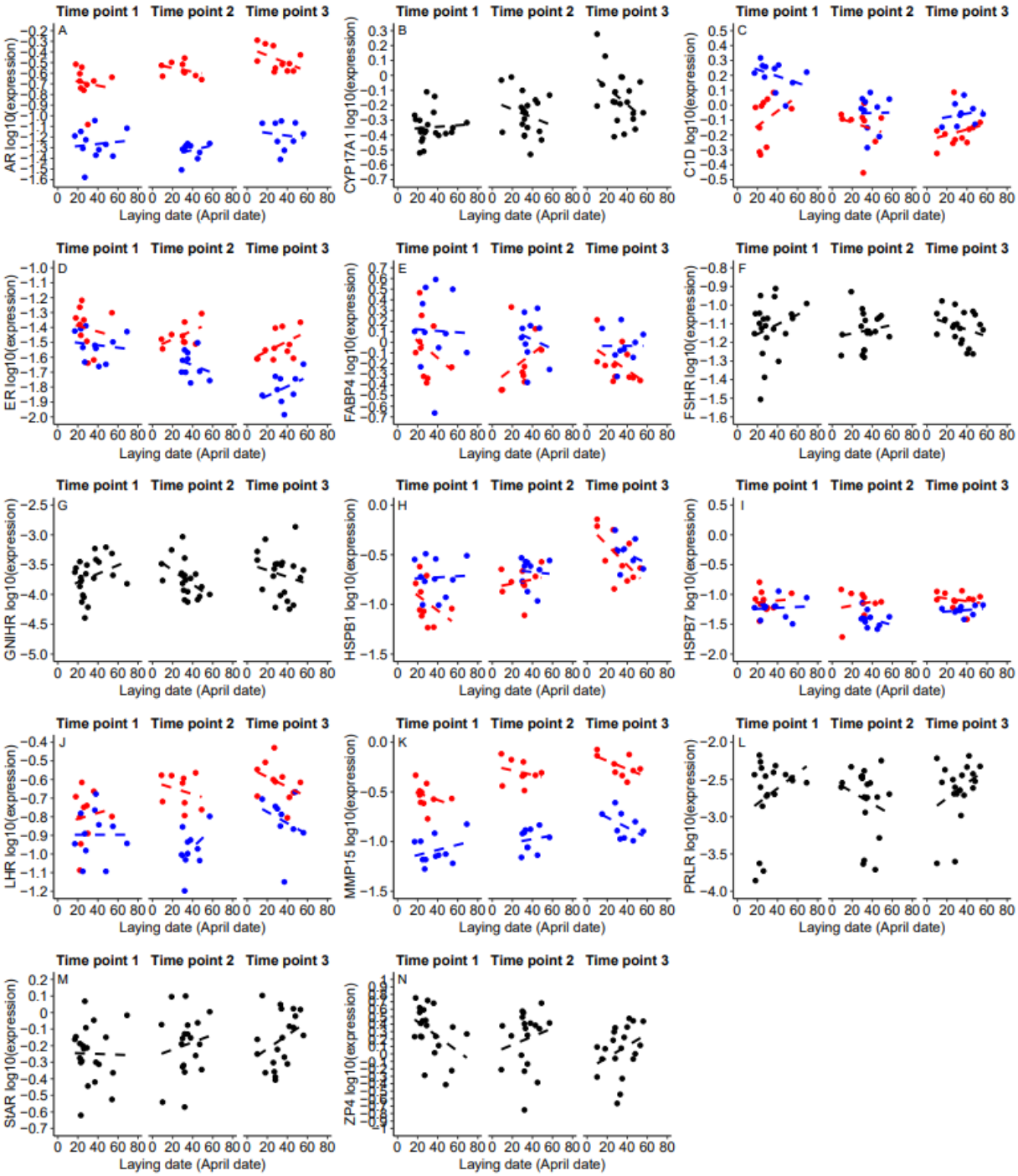


Figure S5.5. Normalized and subsequently Log10 transformed mRNA levels of candidate genes in the three time points in *ovary*. The expression of these genes did not explain variation in egg-laying dates, but main effects of generation (F_1 = red, F_2 = blue) were found. When no generation effect was found, we did not distinguish between generations (data shown in black). Regression lines are dashed, as not to imply significant interactions. Note that the scale on the y-axis differs per panel.

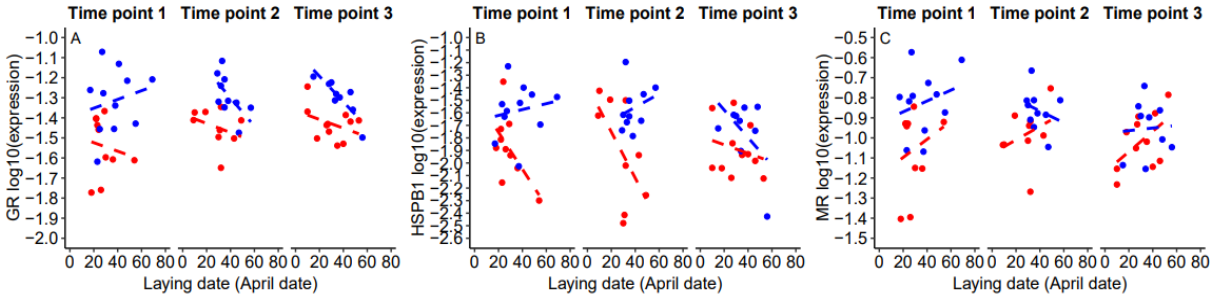


Figure S5.6. Normalized and subsequently log10 transformed mRNA levels of candidate genes in the three time points in liver. The expression of these genes (over time) did not explain variation in egg-laying dates, but main effects of generation (F_1 = red, F_2 = blue) were found. When no generation effect was found, we did not distinguish between generations (data shown in black). Note that the scale on the y-axis differs per panel.

Table S5.1. Lux measurements for 10 different locations in the climate controlled aviaries (n = 10). With the exception of the floor, all locations are regularly visited by the great tits. The average per climate controlled aviary was calculated to get an impression of average daily amount of lux.

	Aviary										
Location in aviary	3	5	6	10	11	15	20	24	29	31	***Average per lociation***
Perch left	508	499	744	530	496	571	569	522	426	528	539.3
Perch middle	208	340	348	345	356	356	325	324	323	345	327
Perch right	440	519	482	614	398	627	480	682	389	547	517.8
Tree	57	417	123	62	53	88	115	51	35	115	111.6
Feeding bowl	301	430	393	401	532	551	618	591	597	620	503.4
Water bowl	414	345	222	361	429	465	414	322	475	517	396.4
Nestbox 1	412	526	623	611	656	607	769	655	862	713	643.4
Nestbox 2	539	677	637	615	661	712	720	716	755	758	679
Nestbox 3	644	649	644	662	685	651	605	614	695	646	649.5
Floor	325	352	350	473	465	469	479	460	457	486	431.6
Average per aviary	384.8	475.4	456.6	467.4	473.1	509.7	509.4	493.7	501.4	527.5	
Average all aviaries											**479.9**

Table S5.2. List of all genes selected based on (1) literature and (2) the genome-wide study (**Chapter 4**) for which primers were built and tested. Primer length varied between 18-24 nucleotides and amplicon length was ideally as close to 100 base pairs as possible. Guanine-cytosine content was set at 55-60% and melting temperature at 55-65 °C. Gene symbol, annotated name (great tit genome), gene name and gene ID in NCBI are given.

Gene symbol	Annotated name	Gene name	Gene ID (NCBI)	Type
AANAT	AANAT	aralkylamine N-acetyltransferase	107212354	candidate
AEN	LOC107209579	apoptosis-enhancing nuclease	107209579	candidate
APOB	APOB	apolipoprotein B	107202091	candidate
APOV1	LOC107200088	apovitellenin-1-like	107200088	candidate
AR	AR	androgen receptor	107203969	candidate
B2M	B2M	beta-2-microglobulin	107209426	reference
BEST3	BEST3	bestrophin 3	107205207	candidate
BMAL1	ARNTL	aryl hydrocarbon receptor nuclear translocator like	107205705	candidate
BMP15	BMP15	bone morphogenetic protein 15	107203672	candidate
C1D	C1D	C1D nuclear receptor corepressor	107202412	candidate
CALM2	CALM2	calmodulin 2	107201314	candidate
FOS	FOS	Fos proto-oncogene, AP-1 transcription factor subunit	107206453	candidate
CLOCK	CLOCK	clock circadian regulator	107202959	candidate
COL18A1	COL18A1	collagen type XVIII alpha 1 chain	107207257	candidate
CRY1	CRY1	cryptochrome circadian regulator 1	107204925	candidate
CSNK1E	CSNK1E	casein kinase 1 epsilon	107204492	candidate
CTSEAL	LOC107205210	cathepsinE-A-like	107205210	candidate
CYP17A1	LOC107206790	cytochrome P450 17A1	107206790	candidate
DIO2	DIO2	iodothyronine deiodinase 2	107206229	candidate
DIO3	DIO3	iodothyronine deiodinase 3	107205941	candidate
DNMT1	DNMT1	DNA methyltransferase 1	107199207	candidate
DRD4	DRD4	dopamine receptor D4	107205240	candidate
EPB41L1	EPB41L1	erythrocyte membrane protein band 4.1 like 1	107213324	candidate
ER	ESR1	estrogen receptor 1	107201411	candidate
FABP4	LOC107200988	fatty acid-binding protein, adipocyte-like	107200988	candidate
FAS	FAS	Fas cell surface death receptor	107206851	candidate
FBXO32	FBXO32	F-box protein 32	107199682	candidate
FOXP2	FOXP2	forkhead box P2	107204827	candidate
FSHR	LOC107202460	follicle-stimulating hormone receptor	107202460	candidate
GABRG1	GABRG1	gamma-aminobutyric acid type A receptor gamma1 subunit	107203534	candidate
GAPDH	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	107211477	reference
GFRA1	GFRA1	GDNF family receptor alpha 1	107206580	candidate
GGH	GGH	gamma-glutamyl hydrolase	107200637	candidate
GnIH	NVPF	neuropeptide VF precursor	107198800	candidate
GnIHR	NPFFR1	neuropeptide FF receptor 1	107198800	candidate
GnRH	GnRH1	gonadtropin releasing hormone 1	107213905	candidate
GR	NR3C1	nuclear receptor subfamily 3 group C member 1	107210791	candidate
HMBS	HMBS	hydroxymethylbilane synthase	107214335	reference
HPRT	HPRT1	hypoxanthine phosphoribosyltransferase 1	107203956	reference
HSD3B2	HSD3B2; LOC107211459	3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase-like	107211459	candidate
HSPA2	HSPA2	heat shock protein family A (Hsp70) member 2	107205698	candidate
HSPB1	HSPB1	heat shock protein family B (small) member 1	107212733	candidate
HSPB11	HSPB11	heat shock protein family B (small) member 11	107208502	candidate
HSPB7	HSPB7	heat shock protein family B (small) member 7	107213476	candidate
IFRD1	IFRD1	interferon related developmental regulator 1	107204795	candidate
IGF1	IGF1	insulin like growth factor 1	107204610	candidate
ITGA3	ITGA3	integrin subunit alpha 3	107198661	candidate
LHR	LOC107201154	lutropin-choriogonadotropic hormone receptor	107201154	candidate
LRP1	LRP1	LDL receptor related protein 1	107198577	candidate
LY86	LY86	lymphocyte antigen 86	107200100	candidate
MAPK8IP1	MAPK8IP1	mitogen-activated protein kinase 8 interacting protein 1	107206189	candidate
MMP15	MMP15	matrix metallopeptidase 15	107209603	candidate
MR	NR3C2	nuclear receptor subfamily 3 group C member 2	107202999	candidate
MTNR1A	MTNR1A	melatonin receptor 1A	107203589	candidate
MTNR1B	MTNR1B	melatonin receptor 1B	107213109	candidate
MTTP	MTTP	microsomal triglyceride transfer protein	107203562	candidate
ncRNA	LOC107208304	uncharacterized LOC107208304	107208304	candidate
NF2	NF2	neurofibromin 2	107211754	candidate
NFIL3	NFIL3	nuclear factor, interleukin 3 regulated	107216383	candidate
NRF1	NRF1	nuclear respiratory factor 1	107205096	candidate
OPN4	OPN4	opsin 4	107206676	candidate
OPN5	OPN5	opsin 5	107202545	candidate
PER2	PER2	period circadian regulator 2	107208758	candidate
PER3	PER3	period circadian regulator 3	107213711	candidate
PGK1	PGK1	phosphoglycerate kinase 1	107203840	reference
PGR	PGR	progesterone receptor	107212560	candidate
PMM1	PMM1	phosphomannomutase 1	107204583	reference
PRKCA	PRKCA	protein kinase C alpha	107212334	reference
PRKCQ	PRKCQ	protein kinase C theta	107205082	reference

Table S5.2 continued

PRLR	PRLR	prolactin receptor	107216256	candidate
RPL13	RPL13	ribosomal protein L13	107209800	reference
RPL19	RPL19	ribosomal protein L19	107215058	reference
RPS28	RPS28	ribosomal protein S28	107215607	candidate
SCD1	SCD	stearoyl-CoA desaturase 1	107207092	candidate
SCYL1	SCYL1	SCY1 like pseudokinase 1	107199115	candidate
SDHA	SDHA	succinate dehydrogenase complex flavoprotein subunit A	107200805	reference
SERT	SLC64A	solute carrier family 6 member 4	107212972	candidate
SIRT1	SIRT1	sirtuin 1	107206664	candidate
SLC35F2	SLC35F2	solute carrier family 35 member F2	107210258	candidate
StAR	STAR	steroidogenic acute regulatory protein	107213856	candidate
TBP	TBP	TATA-box binding protein	107202134	candidate
TIMELESS	TIMELESS	timeless circadian regulator	107199211	candidate
TRH	TRH	thyrotropin releasing hormone	107210208	candidate
TRPA1	TRPA1	transient receptor potential cation channel subfamily A member	107214741	candidate
TRPM8	TRPM8	transient receptor potential cation channel subfamily M member	107207260	candidate
TSHB	TSHB	thyroid stimulating hormone beta	107214721	candidate
VIP	VIP	vasoactive intestinal peptide	107201107	candidate
VLDLR	VLDLR	very low density lipoprotein receptor	107198187	candidate
VTG1	VTG1	vitellogenin-1-like	107207949	candidate
VTG2	VTG2: LOC107208432	vitellogenin-2-like	107208432	candidate
YWHAZ	YWHAZ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase	107199995	reference
ZP4	ZP4	zona pellucida glycoprotein 4	107207025	candidate

Table S5.3. List of primer pairs for the final list of reference and candidate genes used in this study. In addition to the forward and reverse primer sequences, expected amplicon length in base pairs (bp) and target tissue(s) (H = hypothalamus, O = ovary, L = liver) is/are given. An '*' indicates the gene is selected based on results from the genome-wide study (**Chapter 4**).

Gene	Forward primer	Reverse primer	Amplicon length (bp)	Type	Organ(s)
SDHA	GGGCAATAACTCCACGGCAT	TTGTATGGCAGGTCTCTACGA	99	reference	H, O, L
PRKCA	TCCAGTGCCAAGTTTGCTGT	GGGATCATCAGTGTCGGGC	100	reference	H, O, L
RPL19	CTGCGGCAAGAAGAAGGTGT	TCAGCCCATCCTTGATCAGC	100	reference	H, O, L
B2M	AGAACACCCTCCACTGCTTC	TCCCCATAATTGACGTCGGG	99	reference	L
HPRT	TCTGTTCAGGGCTTTGATGT	GAAGGGCATGGGAGGACAC	99	reference	O
RPL13	TACTCCTTCAGCCTCTGCAC	ACAAGAAGTTTGCCCGGACT	99	reference	O
YWHAZ	GAGAAGTTCAGAGCCAGACCT	TGGAGCAATCACAACAGGCA	99	reference	O
DIO2	TCCACACTTGCCACCAACAT	CAAACTGGGAGGAGAAGCCC	99	candidate	H
*NFIL3**	ACCAAGTTTTGCGCACCATG	TCAGCTTTGCCTCACAAGCT	160	candidate	H
OPN5	GTGACTGTTCTGAATCCTGG	AACTGCGGTGATTGTGTTTT	99	candidate	H
TRH	CAACGGGGGACATCTCCTTC	GGACAGACTGAAGAATGAGGCT	100	candidate	H
VIP	AGGATTTAGCTCTTCCTGGC	CTTCACTGACAACTACAGCC	99	candidate	H
AR	TGGTCAGCAGGTTGGAGAAG	CAGCCCATCTTCCTCAACGT	99	candidate	O
*C1D**	GCTGGATTTGGTTTCGGTGT	GCATTGATTCCCTGGGTAGC	540	candidate	O
CYP17A1	AGGCATTTCTTCAGCAGGGC	AGGGTATTGTGGACACCGTG	99	candidate	O
ERα	TCCAGAGATTTCAAGGTGCTGG	TGATGAACCTTCAAGGGGAGG	99	candidate	O
*FABP4**	AGTCTCTTTCCCGTCCCACT	GGTCTCTTTCAAGCTGGGTGA	335	candidate	O
FSHR	CTGCGCTAACCCTTTCCTCT	CTGGGCTTGCATTTCACAGC	99	candidate	O
GnlHR	GTTTCGCTGCATTGTCCACC	CACATGATGAGCAGGGCCAG	98	candidate	O
HSPB1	CTGTCAGCATTCCATCGGGA	GGCTTCATCTCCAGGTGCTT	99	candidate	O
HSPB7	TCGTCAGCATCCCGTTATCC	GGCACCGTCATGAACACCTT	99	candidate	O
*IFRD1**	CCGAGCCCAGAATCTCCTTG	TTTGGACCTGAGCGCATGTA	78	candidate	O
LHR	ATGGCCGGCTTTGACTTTGA	ATCCCAGGATGTCTTCGCAG	99	candidate	O
*MMP15**	TCATCGATCTGCACCACCAC	TGAAGGTGAGGAGGAGGAGG	95	candidate	O
PRLR	AGGCTTTTCAGGAGGTAACCG	GTTCAGGTGCGTTGCATGTT	98	candidate	O
StAR	ATGCTCAGCAGCCAGGTG	AGGAGGGGTTCATAAGGGCT	99	candidate	O
VLDLR	GTGTGAAGAAGACATGTGCCG	CCATCCTCACAGTCCGGATC	99	candidate	O
*ZP4**	TCACCACTTGTCCTTCCAGC	CATCCTTGGCTGTGTCCTGT	457	candidate	O
APO-B	TCAGCCCAGAAACAAATCTG	GATGCTGTTCTGAAAACCCT	99	candidate	L
*APOV1**	CAACAGTCTCAGGGATGGCC	TCCTTAGCACCACTCTCCCT	126	candidate	L
*BEST3**	ACCTGGGGCTCGTGTTTATG	ACAAACCAGACCTGCCAACA	192	candidate	L
*CTSEAL**	TCGGTGAGTCGGTGTTTGAG	GATTGGCTCCTCTACCAGGC	148	candidate	L
GR	TCTCCATCCACAACCTCAGC	CACACGTCAGGACACCGTAG	99	candidate	L
HSPB1	CTGTCAGCATTCCATCGGGA	GGCTTCATCTCCAGGTGCTT	99	candidate	L
IGF1	ACCTTGGCCTGTGTTTGCTT	CCACACACGAACTGAAGAGC	99	candidate	L
MR	GGATGTCATTACGGGGTGGT	AGTCATTCCTTCCCGCACAC	99	candidate	L
VTG2	ATTTACGACACAGCCATCAG	CTTGGGAAAAGACCTTCCAG	99	candidate	L

Table S5.4. The degree of variation per gene explained by selection line, treatment and their interaction in hypothalamus, ovary and liver. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

		Selection line × treatment		**selection line**		**treatment**	
Organ	Gene	*F*	*P*	*F*	*P*	*F*	*P*
Hypothalamus	*DIO2*	$F_{(1,55)}$ = 3.0699	0.966	$F_{(1,57)}$ = 0.3739	0.970	$F_{(1,56)}$ = 0.0022	0.970
	NFIL3	$F_{(1,55)}$ = 1.0006	0.966	$F_{(1,56)}$ = 0.0210	0.970	$F_{(1,57)}$ = 0.1304	0.970
	OPN5	$F_{(1,55)}$ = 2.1557	0.966	$F_{(1,56)}$ = 0.0015	0.970	$F_{(1,57)}$ = 0.0015	0.970
	TRH	$F_{(1,55)}$ = 1.2803	0.966	$F_{(1,56)}$ = 0.1062	0.970	$F_{(1,57)}$ = 0.1495	0.970
	VIP	$F_{(1,55)}$ = 0.0018	0.970	$F_{(1,57)}$ = 0.3048	0.970	$F_{(1,57)}$ = 1.003	0.966
Ovary	*AR*	$F_{(1,56)}$ = 0.6161	0.981	$F_{(1,57)}$ = 0.0029	0.981	$F_{(1,58)}$ = 0.0720	0.981
	C1D	$F_{(1,56)}$ = 0.0006	0.981	$F_{(1,57)}$ = 0.3916	0.981	$F_{(1,58)}$ = 3.536	0.780
	CYP17A1	$F_{(1,56)}$ = 0.9695	0.981	$F_{(1,57)}$ = 0.0046	0.981	$F_{(1,58)}$ = 0.0619	0.981
	ER	$F_{(1,56)}$ = 0.2455	0.981	$F_{(1,58)}$ = 0.3076	0.981	$F_{(1,57)}$ = 0.0204	0.981
	FABP4	$F_{(1,56)}$ = 0.5195	0.981	$F_{(1,58)}$ = 0.2890	0.902	$F_{(1,57)}$ = 0.0009	0.981
	FSHR	$F_{(1,56)}$ = 0.2600	0.981	$F_{(1,57)}$ = 0.0063	0.981	$F_{(1,58)}$ = 0.1801	0.981
	GnIHR	$F_{(1,56)}$ = 0.0644	0.981	$F_{(1,57)}$ = 0.1950	0.981	$F_{(1,58)}$ = 2.0260	0.981
	HSPB1	$F_{(1,56)}$ = 1.3576	0.981	$F_{(1,58)}$ = 5.5870	0.464	$F_{(1,57)}$ = 0.1096	0.981
	HSPB7	$F_{(1,56)}$ = 0.5869	0.981	$F_{(1,58)}$ = 0.7245	0.981	$F_{(1,57)}$ = 0.1556	0.981
	IFRD1	$F_{(1,56)}$ = 0.8993	0.981	$F_{(1,57)}$ = 0.0269	0.981	$F_{(1,58)}$ = 0.2918	0.981
	LHR	$F_{(1,56)}$ = 0.0246	0.981	$F_{(1,57)}$ = 0.1145	0.981	$F_{(1,58)}$ = 0.1346	0.981
	MMP15	$F_{(1,56)}$ = 0.6821	0.981	$F_{(1,57)}$ = 0.0174	0.981	$F_{(1,58)}$ = 0.1616	0.981
	PRLR	$F_{(1,56)}$ = 0.0193	0.981	$F_{(1,57)}$ = 0.0059	0.981	$F_{(1,58)}$ = 0.6462	0.981
	STAR	$F_{(1,56)}$ = 1.1713	0.981	$F_{(1,57)}$ = 0.2546	0.981	$F_{(1,58)}$ = 5.0930	0.464
	VLDLR	$F_{(1,56)}$ = 0.8595	0.981	$F_{(1,58)}$ = 1.0430	0.981	$F_{(1,57)}$ = 0.0546	0.981
	ZP4	$F_{(1,56)}$ = 0.1191	0.981	$F_{(1,58)}$ = 10.750	0.096	$F_{(1,57)}$ = 0.1313	0.981
Liver	*APOB*	$F_{(1,56)}$ = 0.2151	0.954	$F_{(1,58)}$ = 1.0960	0.954	$F_{(1,57)}$ = 0.7889	0.954
	APOV1	$F_{(1,56)}$ = 0.0605	0.954	$F_{(1,58)}$ = 0.2397	0.954	$F_{(1,57)}$ = 0.0568	0.954
	BEST3	$F_{(1,56)}$ = 0.1894	0.954	$F_{(1,58)}$ = 0.2450	0.954	$F_{(1,57)}$ = 0.0370	0.954
	CTSEAL	$F_{(1,56)}$ = 0.1241	0.954	$F_{(1,58)}$ = 0.6397	0.954	$F_{(1,57)}$ = 0.0885	0.954
	GR	$F_{(1,56)}$ = 0.0005	0.983	$F_{(1,58)}$ = 0.6164	0.954	$F_{(1,57)}$ = 0.0046	0.982
	HPSB1	$F_{(1,56)}$ = 0.3781	0.954	$F_{(1,58)}$ = 0.9166	0.954	$F_{(1,57)}$ = 0.4920	0.954
	IGF1	$F_{(1,56)}$ = 0.1873	0.954	$F_{(1,57)}$ = 0.0585	0.954	$F_{(1,57)}$ = 1.9080	0.954
	MR	$F_{(1,56)}$ = 0.0159	0.972	$F_{(1,57)}$ = 1.9396	0.954	$F_{(1,58)}$ = 0.0382	0.954
	VTG2	$F_{(1,56)}$ = 0.1183	0.954	$F_{(1,58)}$ = 0.2248	0.954	$F_{(1,57)}$ = 0.2119	0.954

Table S5.5. Factor (gene) loadings per principal component from a PCA in hypothalamus. The last row indicated the percentage of variance explained by the PC.

	Principal component				
Gene	PC1	PC2	PC3	PC4	PC5
DIO2	-0.477	0.377	-0.151	0.138	0.767
NFIL3	-0.484	0.007	-0.748	0.018	-0.454
OPN5	-0.474	0.317	0.443	-0.647	-0.246
TRH	-0.488	-0.204	0.470	0.668	-0.231
VIP	-0.276	-0.846	-0.021	-0.341	0.302
Variance explained (%)	67.2	19.6	6.6	3.6	3.0

Table S5.6. Factor (gene) loadings per principal component from a PCA in ovary. The last row indicated the percentage of variance explained by the PC.

	Principal component															
Gene	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
AR	-0.437	0.090	0.017	-0.027	0.072	0.038	0.043	0.027	0.008	-0.010	-0.080	0.022	0.317	-0.028	-0.371	-0.737
C1D	0.264	0.138	0.391	0.093	-0.257	0.090	0.019	-0.341	-0.022	0.570	0.104	0.101	0.454	0.005	-0.069	0.048
CYP17A1	-0.224	-0.361	0.007	-0.160	-0.077	0.207	-0.225	0.145	-0.426	0.471	-0.135	0.300	-0.369	0.088	-0.131	0.007
ER	-0.184	0.381	0.247	0.083	0.218	-0.030	-0.219	0.267	0.166	0.261	-0.536	-0.168	-0.050	-0.361	0.144	0.147
FABP4	0.207	-0.032	0.315	-0.075	0.607	-0.062	0.202	0.475	-0.280	0.043	0.302	0.058	0.183	0.063	-0.015	0.018
FSHR	0.103	-0.052	0.370	-0.469	-0.480	-0.040	0.098	0.417	0.235	-0.221	-0.208	0.180	0.079	0.166	0.029	-0.023
GnIHR	-0.179	-0.118	0.384	0.165	-0.085	-0.606	-0.065	-0.242	-0.470	-0.254	-0.200	-0.074	0.025	0.087	0.030	0.053
HSPB1	-0.025	-0.475	0.081	-0.052	0.009	0.012	0.466	-0.031	0.132	0.271	-0.158	-0.639	-0.075	0.047	0.080	-0.090
HSPB7	-0.290	-0.002	0.410	0.091	0.239	-0.038	-0.091	-0.204	0.501	0.061	0.238	0.161	-0.340	0.411	0.090	-0.037
IFRD1	0.318	0.293	0.223	-0.127	-0.072	0.034	-0.119	-0.051	-0.089	-0.067	0.131	-0.363	-0.469	-0.058	-0.575	-0.093
LHR	-0.368	-0.045	0.234	-0.129	-0.220	0.074	0.070	0.007	-0.059	-0.061	0.526	-0.046	-0.086	-0.630	0.203	0.036
MMP15	-0.435	0.005	-0.046	-0.082	0.017	0.097	0.089	0.028	0.059	-0.064	0.037	-0.118	0.270	0.145	-0.512	0.636
PRLR	0.031	-0.066	0.267	0.619	-0.102	0.503	0.311	0.085	-0.088	-0.286	-0.184	0.165	-0.115	-0.050	-0.082	0.015
STAR	0.096	-0.306	0.221	-0.251	0.241	0.449	-0.492	-0.265	-0.030	-0.308	-0.116	-0.186	0.238	-0.040	0.102	-0.008
VLDLR	-0.212	0.411	-0.021	0.018	-0.203	0.277	-0.071	0.155	-0.328	0.013	0.151	-0.373	0.036	0.475	0.373	-0.051
ZP4	-0.060	0.315	0.020	-0.456	0.220	0.159	0.499	-0.431	-0.190	-0.063	-0.238	0.215	-0.148	-0.025	0.115	0.044
Variance explained (%)	29.8	18.3	10.6	8.1	6.1	5.5	4.3	4.0	3.1	2.6	2.1	1.8	1.5	1.2	0.7	0.1

Table S5.7. Factor (gene) loadings per principal component from a PCA in liver. The last row indicated the percentage of variance explained by the PC.

	Principal component								
Gene	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
APOB	0.319	0.349	-0.067	0.210	-0.740	0.392	0.152	-0.068	0.022
APOV1	0.463	-0.029	-0.065	-0.175	-0.007	-0.396	0.518	0.542	0.176
BEST3	0.437	-0.170	0.255	0.040	0.272	0.410	-0.280	0.029	0.625
CTSEAL	0.464	-0.150	0.107	0.022	0.103	0.171	-0.322	0.285	-0.725
GR	0.187	0.539	0.174	0.099	-0.062	-0.581	-0.528	-0.035	0.112
HPSB1	-0.014	0.416	0.209	-0.843	0.072	0.237	0.050	-0.064	-0.071
IGF1	0.194	0.251	-0.875	-0.042	0.313	0.112	-0.119	-0.082	0.033
MR	-0.034	0.518	0.261	0.447	0.504	0.180	0.400	0.029	-0.129
VTG2	0.452	-0.177	0.086	-0.045	0.091	-0.231	0.266	-0.779	-0.129
Variance explained (%)	46.0	27.7	9.3	8.1	4.3	2.0	1.4	0.9	0.3

Table S5.8. The degree of variation within a principal component explained by egg-laying date, time point, their interaction and generation in hypothalamus. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

	Laying date × time point		Laying date		Time point		Generation	
Principal component	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
PC1	$F_{(2,52)} = 1.1528$	0.718	$F_{(1,56)} = 0.7338$	0.718	$F_{(2,54)} = 0.4565$	0.881	$F_{(1,57)} = 143.00$	**<0.0001**
PC2	$F_{(2,52)} = 0.0698$	0.933	$F_{(1,56)} = 0.1445$	0.881	$F_{(2,54)} = 0.3855$	0.881	$F_{(1,57)} = 1.8390$	0.603
PC3	$F_{(2,52)} = 0.1322$	0.528	$F_{(1,54)} = 0.0760$	0.894	$F_{(2,55)} = 1.0254$	0.718	$F_{(1,57)} = 1.3810$	0.613
PC4	$F_{(2,52)} = 0.8049$	0.894	$F_{(1,54)} = 1.4685$	0.613	$F_{(2,55)} = 4.8538$	0.060	$F_{(1,57)} = 7.1088$	0.060
PC5	$F_{(2,52)} = 0.7780$	0.775	$F_{(1,57)} = 8.4750$	0.050	$F_{(2,55)} = 0.5182$	0.881	$F_{(1,54)} = 0.0162$	0.933

Table S5.9. The degree of variation within a principal component explained by egg-laying date, time point, their interaction and generation in ovary. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

	Laydate × timepoint		Laydate		Time point		Generation	
Principal compone	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
PC1	$F_{(2,51)} = 1.9658$	0.644	$F_{(1,53)} = 0.7393$	0.901	$F_{(2,53)} = 14.684$	**<0.0001**	$F_{(1,54)} = 269.57$	**<0.0001**
PC2	$F_{(2,51)} = 4.2002$	0.213	$F_{(1,56)} = 0.6377$	0.913	$F_{(2,55)} = 22.010$	**<0.0001**	$F_{(1,51)} = 26.319$	**<0.0001**
PC3	$F_{(2,51)} = 1.0069$	0.901	$F_{(1,53)} = 0.0065$	0.957	$F_{(2,55)} = 3.7032$	0.283	$F_{(1,54)} = 0.4503$	0.953
PC4	$F_{(2,51)} = 1.3022$	0.884	$F_{(1,56)} = 0.0540$	0.957	$F_{(2,54)} = 0.0699$	0.957	$F_{(1,53)} = 0.0008$	0.978
PC5	$F_{(2,51)} = 2.4382$	0.564	$F_{(1,55)} = 0.0953$	0.957	$F_{(2,53)} = 0.4335$	0.953	$F_{(1,56)} = 0.8978$	0.901
PC6	$F_{(2,51)} = 0.5321$	0.953	$F_{(1,56)} = 0.9532$	0.901	$F_{(2,53)} = 0.6314$	0.953	$F_{(1,55)} = 0.7767$	0.901
PC7	$F_{(2,51)} = 1.0095$	0.901	$F_{(1,56)} = 1.0920$	0.884	$F_{(2,54)} = 0.9027$	0.909	$F_{(1,53)} = 0.0977$	0.957
PC8	$F_{(2,51)} = 2.5602$	0.564	$F_{(1,56)} = 0.3730$	0.953	$F_{(2,53)} = 0.0607$	0.957	$F_{(1,55)} = 0.0196$	0.957
PC9	$F_{(2,51)} = 0.1278$	0.957	$F_{(1,54)} = 0.2249$	0.953	$F_{(2,55)} = 1.3410$	0.884	$F_{(1,53)} = 0.0480$	0.957
PC10	$F_{(2,51)} = 0.4154$	0.953	$F_{(1,56)} = 1.0760$	0.884	$F_{(2,53)} = 0.7675$	0.953	$F_{(1,55)} = 1.2181$	0.884
PC11	$F_{(2,51)} = 0.0596$	0.957	$F_{(1,54)} = 0.0621$	0.957	$F_{(2,55)} = 2.0170$	0.644	$F_{(1,55)} = 1.2814$	0.957
PC12	$F_{(2,51)} = 0.4006$	0.953	$F_{(1,56)} = 3.0690$	0.564	$F_{(2,54)} = 2.4852$	0.564	$F_{(1,53)} = 0.0467$	0.957
PC13	$F_{(2,51)} = 0.3132$	0.957	$F_{(1,56)} = 1.7110$	0.784	$F_{(2,53)} = 0.4412$	0.953	$F_{(1,55)} = 0.3250$	0.953
PC14	$F_{(2,51)} = 0.4064$	0.953	$F_{(1,56)} = 2.3400$	0.644	$F_{(2,53)} = 0.1350$	0.957	$F_{(1,55)} = 0.3624$	0.953
PC15	$F_{(2,51)} = 0.7338$	0.953	$F_{(1,54)} = 2.4270$	0.644	$F_{(2,55)} = 0.3813$	0.953	$F_{(1,53)} = 0.0092$	0.957
PC16	$F_{(2,51)} = 0.3651$	0.953	$F_{(1,54)} = 1.1732$	0.884	$F_{(2,55)} = 4.4320$	0.205	$F_{(1,53)} = 0.0234$	0.957

Table S5.10. The degree of variation within a principal component explained by egg-laying date, time point, their interaction and generation in liver. All factors were fixed effects in the model. The F-statistic with the degrees of freedom ($F_{(df, ndf)}$) and Benjamini and Hochberg adjusted *p*-values are given. Bold *p*-values indicate significance.

	Laydate × timepoint		Laydate		Time point		Generation	
Principal component	*F*	*P*	*F*	*P*	*F*	*P*	*F*	*P*
PC1	$F_{(2,53)} = 11.019$	**0.001**	$F_{(1,53)} = 21.616$	**<0.0001**	$F_{(2,53)} = 102.02$	**<0.0001**	$F_{(1,53)} = 7.2676$	0.054
PC2	$F_{(2,53)} = 1.2076$	0.553	$F_{(1,55)} = 0.8176$	0.567	$F_{(2,56)} = 2.8158$	0.262	$F_{(1,58)} = 38.090$	**<0.0001**
PC3	$F_{(2,53)} = 0.0237$	0.977	$F_{(1,58)} = 3.1780$	0.262	$F_{(2,56)} = 1.5340$	0.426	$F_{(1,55)} = 0.1767$	0.839
PC4	$F_{(2,53)} = 0.0274$	0.977	$F_{(1,58)} = 5.5790$	0.113	$F_{(2,55)} = 0.0638$	0.977	$F_{(1,57)} = 2.1747$	0.329
PC5	$F_{(2,53)} = 2.3870$	0.285	$F_{(1,58)} = 2.2050$	0.329	$F_{(2,55)} = 0.6746$	0.712	$F_{(1,57)} = 1.7167$	0.413
PC6	$F_{(2,53)} = 0.5691$	0.733	$F_{(1,57)} = 0.3336$	0.733	$F_{(2,55)} = 0.0396$	0.977	$F_{(1,58)} = 1.5080$	0.426
PC7	$F_{(2,53)} = 5.6007$	0.043	$F_{(1,54)} = 0.0449$	0.962	$F_{(2,54)} = 2.3713$	0.285	$F_{(1,53)} = 0.7384$	0.567
PC8	$F_{(2,53)} = 2.2476$	0.298	$F_{(1,58)} = 5.0030$	0.131	$F_{(2,56)} = 1.0108$	0.567	$F_{(1,55)} = 0.1306$	0.863
PC9	$F_{(2,53)} = 1.0068$	0.567	$F_{(1,55)} = 0.0339$	0.962	$F_{(2,56)} = 0.9542$	0.567	$F_{(1,58)} = 3.3140$	0.262

Appendix 5.1. explaining variation in reproductive phenotypes

Differences in egg-laying dates and largest follicle widths between selection lines and treatments were tested by performing ANOVA, with egg-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.

Egg-laying dates (first breeding season) ranged from April date 9 (9 April) to April date 69 (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 egg-laying dates. We did, however, find an effect of generation ($F_{2,96} = 3.38$, $p = 0.04$); a difference in mean egg-laying date of ~7.5 days between the F_1 and F_2 generation ($t_{61.99} = -2.50$, $p = 0.02$) (Figure A5.1.1).

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 A5.1.2).

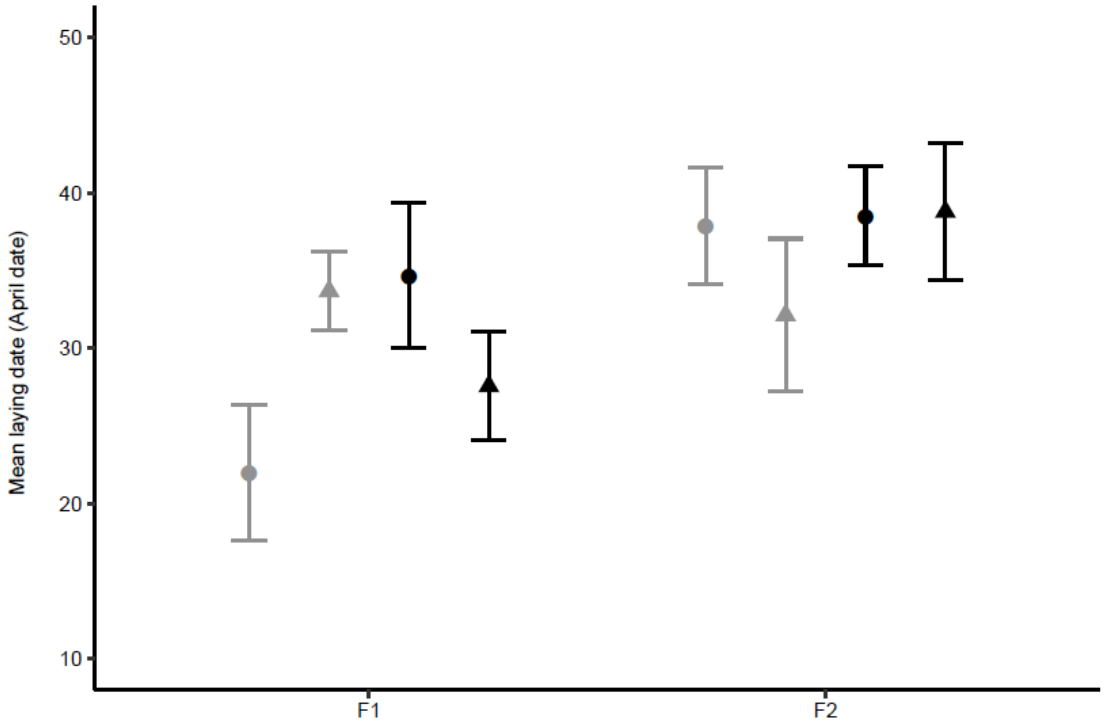


Figure A5.1.1. Mean egg-laying dates (mean ± SEM) in April dates (y-axis, 10 = 10 April, 40 = 10 May etc.) for the F_1 and F_2 generation (x-axis) shown for females based on their selection line × treatment groups. No significant differences were found in mean egg-laying date between early (dark grey) 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.

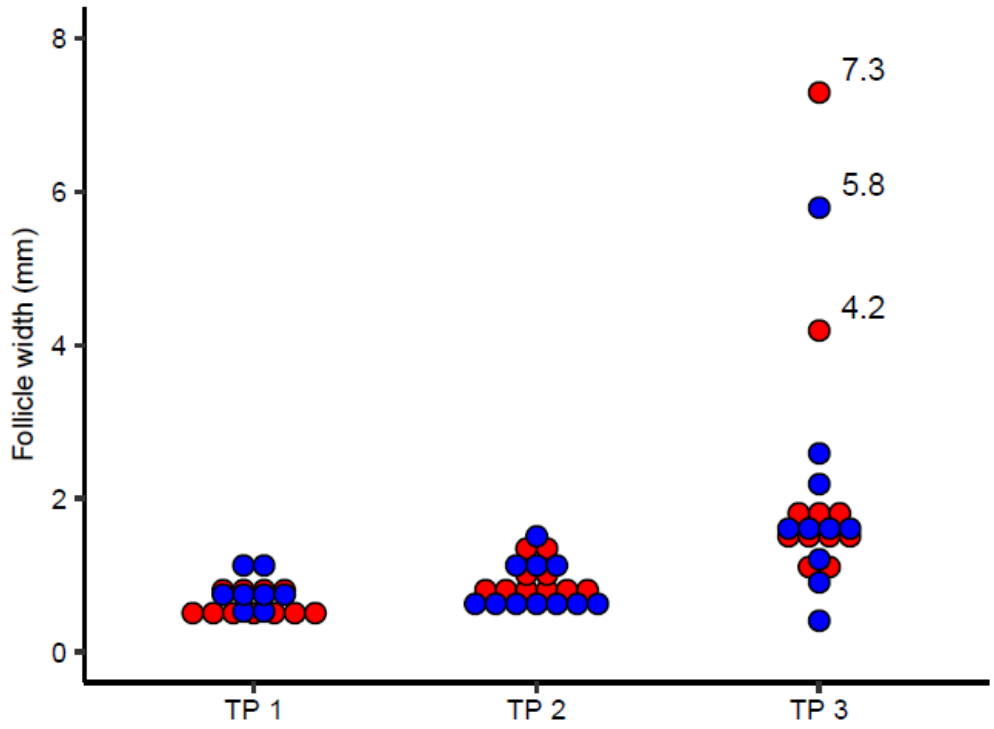


Figure A5.1.2. Follicle widths of the largest follicles for the three different time points (TP 1 = time point 1, TP 2 = time point 2, TP 3 = time point 3) from the females of the F_1 (red) and F_2 (blue) generation.

Appendix 5.2. 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 et al. 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 A5.2.1) 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 A5.2.1) 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} = 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, $V_{B2M\text{-}PRKCA\text{-}RPL19\text{-}SDHA\text{-}PRKC}$ < 0.15 and together with the high R^2 (Figure A5.2.1) found between NF_3 and NF_4, adding a fifth reference gene is not necessary for accurate normalization.

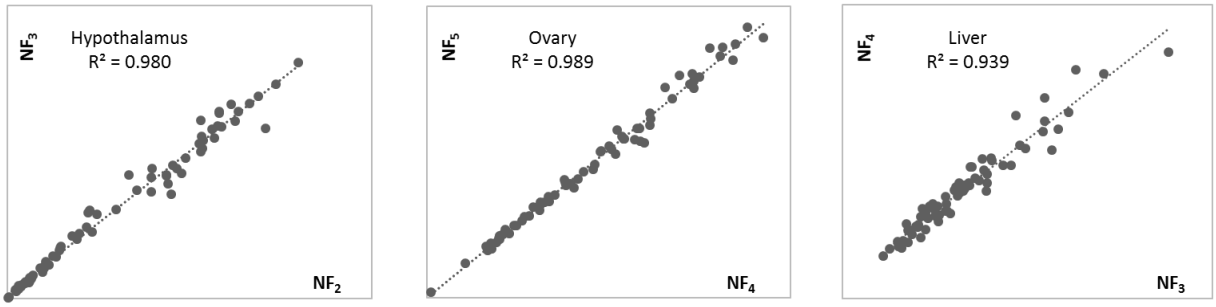


Figure A5.2.1. Scatterplots for hypothalamus, ovary and liver of normalization factors before (x-axis) and after (y-axis) the addition of a reference gene. The high R^2 indicates that the inclusion of a second, fourth and third reference gene is not necessary for hypothalamus, ovary and liver respectively. But, because V was still slightly above 0.15 for hypothalamus and ovary and V < 0.15 after adding the fourth reference gene in liver, we did add the extra reference gene.

Appendix 5.3. Validation for linking breeding season 1 (spring) to breeding season 2 (autumn)

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 A5.3.1), 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 et al. 1974; Walzem 1996). We found increasing follicle widths over time (Figure S5.4), with three females carrying follicles similar to F3-F1 approximate sizes (Table A5.1) 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 egg-laying dates and follicle widths, especially in time point 3 (Figures S5.6).We are therefore confident that the mRNA expression levels from the second breeding season are representative of the phenotypes (i.e. egg-laying dates) recorded and assume these breeding seasons to be similar.

Table A5.3.1. Calculation of approximate size for the F5-F1 follicles based on the traditional hierarchical model of follicle development (Astheimer & 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 (mm^3)
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

CHAPTER 6

Exploring temporal changes in DNA methylation and RNA expression in a small songbird - correlations within and between tissues

Irene Verhagen*, Melanie Lindner*, Veronika N. Laine, Heidi M. Viitaniemi, Marcel E. Visser, Arild Husby & Kees van Oers

**contributed equally to the manuscript*

ABSTRACT

Temporally expressed traits are affected by environmental conditions that need to be translated into gene transcription, for which DNA methylation is a likely key regulator. To understand this transcriptional role of DNA methylation we need within-individual assessment of both methylation and expression changes over time. Within-individual repeated sampling of tissues where essential transcription takes place such as brain and ovary is, however, challenging. Here, we determine to what extend methylation and expression patterns in more repeatedly accessible tissue, i.e. blood, reflects patterns in tissues unable to sample repeatedly in the great tits (Parus major). For this, 18 females were sacrificed at three time points throughout the pre-laying and egg-laying period and their blood, liver, hypothalamus, and ovary were sampled. We assessed both DNA methylation and expression profiles and found a positive correlation between changes in CpG site methylation in RBCs and liver. For CpG sites around the TSS of a gene in the ovary, an increase in methylation in RBCs over time was associated with a decrease in the expression of the associated gene, while no such association was present for CpG sites within the gene body or in the 10k up- or downstream region. This study provides unique insights into tissue-specific and tissue-general temporal changes in DNA methylation and how such changes relate to changes in expression patterns. As such, tissue-general temporal changes open up the possibility to monitor how environmental conditions affect temporally expressed traits within individuals, even in natural populations.

Submitted to BMC Genomics

Introduction

Temporally regulated traits, like aging-related traits or seasonal timing, change in value throughout an individual's life-time or change repeatedly throughout specific time periods of the year. For example, every spring a seasonally breeding female responds to increasing photoperiod and temperature to gradually switch from an inactive state to an active reproductive state, but the time of this transition depends on the environmental conditions, i.e. the trait is phenotypically plastic (Pigliucci 2001). It, however, remains poorly understood which mechanism underlies the translation of the environmental conditions to those within-individual changes in trait value over time, i.e. that lead to trait plasticity.

Epigenetic modifications, like DNA methylation, are known to modulate the expression of phenotypes through an interaction with proteins required for the initiation of gene transcription (Bird 2002). DNA methylation can be highly dynamic in response to environmental signals (Bind et al. 2012; Liu et al. 2003; Stevenson & Prendergast 2013; Viitaniemi et al. 2019) and hence constitutes a candidate for the regulation of transcriptional mechanisms that shape temporally regulated traits (Sepers et al. 2019). Indeed, changes in DNA methylation were found as a common factor for aging in mammals with a striking tissue-specificity for age related DNA methylation changes (Maegawa et al. 2010; Slieker et al. 2018). Also, in chicken DNA methylation regulator genes were tissue-specifically responsive to acute and chronic stress in the form of food deprivation, and hepatic glucocorticoid receptors (GR) were suggested to play a critical role in regulating the early-life nutritional stress response of birds (Kang et al. 2017). Furthermore, DNA methylation was found to regulate seasonally expressed traits like hibernation of thirteen-lined ground squirrels (*Ictidomys tridecemlineatus*, Alvarado et al. 2015), photoperiodic timing in insect parasitoid (*Nasonia vitripennis*, Pegoraro et al. 2016), flowering time in plants (Cortijo et al. 2014; Wilschut et al. 2016), and timing of reproduction in Siberian hamsters (*Phodopus sungorus*, Stevenson & Pendergast 2013). The latter study demonstrated that short day length induced a temporal decrease in DNA methylation levels within the promoter region of type III deiodinase (DIO3), a gene involved in the photoperiodic regulation of reproduction, and furthermore established a causal link between reduced DIO3 promoter methylation and gonadal regression via increased transcription of DIO3 (Stevenson & Pendergast 2013).

Most studies on associations between temporal changes in DNA methylation and traits are, however, based on between- individual samples, which is complicated by the large between individual variation in trait- and methylation values. Therefore, to understand to what extent changes in DNA methylation shape changes in individual trait values, within-individual measures of methylation values are essential. Depending on the trait of interest, it might not be feasible to repeatedly sample tissues of biological relevance within the same individual. For example, avian timing of breeding requires crucial physiological processes like vitellogenesis, yolk deposition, follicle growth, and oviduct development. These processes are regulated by a neuroendocrine cascade (the hypothalamic-pituitary-gonadal-liver (HPGL) axis), which is triggered by environmental information that is received, translated and transduced from the brain (Williams 2012). Understanding how transcriptional mechanisms in inaccessible tissues such as liver, ovary, and hypothalamus

that underlie the HPGL axis are regulated throughout the breeding season would give new insights on how females time their reproduction. However, sampling within-individual changes in DNA methylation in such inaccessible tissues is impossible as it requires sacrificing each sampled individual. It also does not allow for measuring the final trait value such as timing of egg-laying.

An accessible tissue that allows for repeated within-individual sampling is blood. Interestingly, avian blood, in contrast to mammalian blood, contains nucleated red blood cells (RBCs), so that more than 90% of the DNA isolated from avian blood originates from erythrocytes (Verhulst et al. 2016). Therefore, only a small amount of avian blood (<10µl) is required to isolate sufficient genomic DNA (~1ug) to determine genome-wide DNA methylation profiles via reduced representation bisulfite sequencing (RRBS) (Mäkinen et al. 2019; Meissner et al. 2005). This low amount of blood required opens up the possibility to examine within-individual short-term changes in DNA methylation. Indeed, repeated blood sampling of great tit (*Parus major*) females revealed within-individual changes in RBC methylation levels throughout the breeding season that correlated to a female's reproductive state (Lindner et al. *in prep.*; Viitaniemi et al. 2019). It is, however, currently unclear how representative changes in RBC methylation are for methylation in the organs involved, and how relevant they are for expression patterns, in shaping avian timing of breeding.

Previously, strong correlations have been found between absolute blood methylation levels and absolute methylation levels in inaccessible tissues. Therefore, DNA methylation in blood is proposed as a biomarker for DNA methylation in other tissues (Derks et al. 2016; McKay et al. 2011). It is, however, currently unknown to what extend changes in RBC DNA methylation reflect changes in DNA methylation in other tissues (i.e. tissue-general changes). To address, which seasonal changes in DNA methylation are tissue-general or not (i.e. tissue-specific) and how tissue-general changes relate to changes in gene expression, we used 18 great tit females that were divided into three sampling groups and sampled for RBCs, liver, hypothalamus, and ovary tissue at three time points throughout the pre-laying and egg-laying period. We used the collected tissues to assess DNA methylation levels (RBCs, liver) together with candidate gene (liver, using individual qPCR data) and genome-wide (hypothalamus, ovary and liver, using RNA-seq data of pooled individuals) expression profiles. We examined to what extent i) changes in DNA methylation in RBCs coincide to changes in DNA methylation in liver, ii) changes in DNA methylation in liver correlate to changes in gene expression in liver, and iii) changes in DNA methylation in RBCs correlate to changes in gene expression in liver, ovary and hypothalamus.

Material and methods

Experimental setup

For a detailed description of the experimental setup and sampling see **Chapter 5**. In short, 36 great tit pairs (18 early pairs and 18 late pairs) originating from the F_2 generation of lines artificially selected for early and late timing of breeding (for details see Gienapp et al. 2019, **Chapter 2**), were housed in 36 climate-controlled aviaries (2m × 2m × 2.25m) at the Netherlands Institute of Ecology (NIOO-KNAW). Birds were subjected to a photoperiod mimicking the natural photoperiod and two contrasting temperature environments representing the cold 2013 spring and the warm 2014 spring in the Netherlands. Temperatures changed every hour to follow the observed hourly temperatures in these years as closely as possible. The combination of selection line and temperature environment resulted in four groups: 'early-warm', 'early-cold', 'late-warm' and 'late-cold'. Birds were fed ad libitum and had water available for drinking and bathing (Visser et al 2011).

Where great tits normally only have one reproductive season per year, the pairs included in this study were induced to breed twice (see **Chapter 5** for details). In short, in the first breeding event from January until July, individuals were blood sampled bi-weekly for other studies (for details see Mäkinen et al. 2019; Viitaniemi et al. 2019), and egg-laying dates were obtained. Then, birds went through a period of short-day length and low temperatures to make them photoreceptive and temperature sensitive again. Subsequently, birds received the same temperatures and photoperiods as in the first breeding event to induce a second breeding event (September until November). The 36 pairs were divided into three groups, based on the females' egg-laying dates in the first breeding season, and sacrificed at three time points throughout this second breeding event.

Tissue collection and preparation

Based on the reproductive behavior from the first breeding event, three sampling time points throughout the second breeding season were chosen,: (1) October 7 (resembling March 7) when photoperiod exceeded 11hrs, which is necessary to initiate gonadal maturation (Silverin et al. 1993), (2) October 28 (resembling March 30) when nest building occurred in the first breeding season, but prior to egg-laying and (3) November 18 (resembling April 20) when about 25% of the females in 2015 had initiated egg-laying in the first breeding event. Per time point we sacrificed both sexes of one group (n = 12 pairs), but we focus on females in this study (see **Chapter 5** for details). Briefly, birds were caught per pair from their aviary and anaesthetized deeply with Isoflurane during which a blood sample (300 µl) was taken, and stored in either RNAlater (transferred to -80°C) or Queens buffer for analysis (see '*Reduced representation bisulfite sequencing (RRBS)*' below). Tissues, including brain, ovary and liver were dissected and stored in -80°C until further processing. At a later stage, the hypothalamus, being the center for integration, transduction and translation of environmental cues, was isolated from the rest of the brain and, until further

processing, stored in -80°C. The samples used in this study are from the early selection line females in the second (autumn) breeding season only (n=18, with 6 females per sampling time point), because during this breeding event blood, hypothalamus, ovary and liver where collected as opposed to the first breeding event.

Reduced representation bisulfite sequencing (RRBS)

To acquire the reduced representation libraries, the preparation protocol according to manufacturer's protocol (Illumina) was used with some changes. Briefly, samples were digested using the restriction enzyme MspI and the resulting DNA fragments of various size were subsequently bisulfite treated, which converts un-methylated cytosine bases into uracil bases, whereas methylated cytosine bases are resistant to the treatment. Fragmented and bi-sulfite treated DNA was then end-repaired with DNA polymerase I and A-overhangs were added to the 3' ends of each fragment for adapter ligation. Individual sample libraries were barcoded using standard Illumina adapters. Libraries were purified, size selected with Ampure XP beads (Beckman Coulter) and concentrations were determined by quantitative polymerase chain reaction (qPCR). This selection yielded a fragment size range of approximately 30-180 base pairs, with a mean of 85. Six libraries were pooled into the same sequencing lane (Table S6.1). Each pool was sequenced 100bp single end (Table S6.1) on a HiSeq2500 sequencer with a HiSeq SBS sequencing kit version 4 (Illumina). Sequencing was conducted in two separate HiSeq runs to yield enough coverage per sample. An internal positive control (PhiX) was used to obtain reliable sequence generation in the sequencing processing and the phiX reads and adapters were removed before data analysis. Library preparation and sequencing were performed at the SciLife Lab, Uppsala University, Sweden.

Sequence read quality and alignment

Sequencing read quality was investigated with the FastQC 0.11.7 quality control tool. Low quality bases as well as illumina adapter contamination resulting from read-through of short fragments were trimmed using Trim Galore! v0.4.4 with default parameters under the –rrbs mode. This mode disregards the first five bases in the 5' to reduce calling of false positive methylation as a result of bisulfite treatment. Each sample's reads from both of the sequencing runs were combined together for alignment. Trimmed sequencing reads were aligned against a bisulfite converted version of the *Parus major* reference genome v1.1[**] using Bismark 0.19.1 aligner in *rrbs* mode. We used all assembled chromosomes as well as all scaffolds of the reference genome. After alignment and CpG site calling we selected the sites with a minimum coverage of 10x across all samples within a tissue (RBCs, hypothalamus, liver and ovary) for further analyses. We calculated the methylation proportion for a site in the respective sample as the proportion of methylated counts relative to the total read counts. As we were interested in sites that change over time, we

[**] https://www.ncbi.nlm.nih.gov/assembly/GCF_001522545.2

excluded all sites that showed a methylation proportion of either zero or 1 across all samples from downstream analyses.

Gene annotation

To better understand the potential relevance of a GpC site for expression of the local gene, we annotated CpG sites, using R packages 'GenomicFeatures' (Lawrence et al. 2013) and 'rtracklayer' (Lawrence et al. 2009), in respect to different genomic locations: TSS (300bp upstream - 50bp downstream of the annotated gene start site), promoter region (2000bp upstream - 200bp downstream of the annotated gene start), gene body (exons and introns), and 10k up- and downstream regions (10kb regions adjacent to the gene body, respectively). Each identified CpG site was assigned to above specified genomic locations (and the gene annotated to that region) with BEDtools v.2.26.0. (Quinlan & Hall 2010). Earlier studies in our species have shown that only methylation levels around TSS and within promoter regions associate with RNA expression (Laine et al. 2016, Derks et al. 2016). CpG sites within the TSS or promoter region of annotated genes were used for (i) RBC methylation and liver methylation correlation as well as (ii) liver methylation and candidate gene expression (qPCR, see below) correlation. CpG sites within the TSS, gene body, and 10k up-/downstream regions were used for (iii) correlation of liver methylation to liver gene expression and RBC methylation to expression in liver, hypothalamus, or ovary (Figure 6.1). Here gene expression levels were obtained by means of RNA-sequencing (RNA-seq, see below).

RNA extraction, real-time quantitative polymerase chain reaction and sequencing

From the same females of which we acquired DNA methylation patterns, we used qPCR and RNA-seq data, which was generated by two other studies; individual candidate gene expression levels (qPCR, **Chapter 5**) and genome-wide expression levels of 12 pools of three females (**Chapter 4**). In short, RNA was isolated from hypothalamus, ovary and liver by Trizol extraction and reverse transcribed into cDNA (see **Chapter 5** for details).

qPCR. Briefly, primer pairs were built based on the *Parus major* reference genome v1.1. and *Parus major* annotation release 101[††] and checked for specificity using a BLAST search. Efficiency of each primer pair was determined by a 5-point standard curve of cDNA samples. Relative transcript levels were measured within hypothalamus, ovary and liver for a set of candidate genes in individual females by real-time qPCR using the SYBR Green method followed by fluorescence measurements and analyses to obtain cycle thresholds. Expression levels of the candidate genes were normalized against reference genes. The combination and number of reference genes differs per organ and can be found elsewhere (**Chapter 5**).

[††] https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Parus_major/101/

RNA sequencing. RNA of three females per line × treatment × time point group was pooled, which resulted in a total of 12 pools of which 6 (with 3 females each) are represented by early selection line females (see **Chapter 4** for details) and used in this study. Briefly, libraries were made using the Illumina TruSeq strand-specific mRNA method (Illumina, San Diego, CA, USA) and one lane of Illumina HiSeq 2500 (single-end 50bp) for 12 pools. Reads were filtered for low quality. Subsequently, trimmed reads were mapped to the *Parus major* reference genome v1.1, after which transcripts were assembled based on the *Parus major* annotation release 101. Unique reads that mapped to transcripts were counted.

Statistical analysis

All statistics and plotting was performed in R (Version 3.5.2). An overview of how the different data sets and tissues are linked is provided in Figure 6.1.

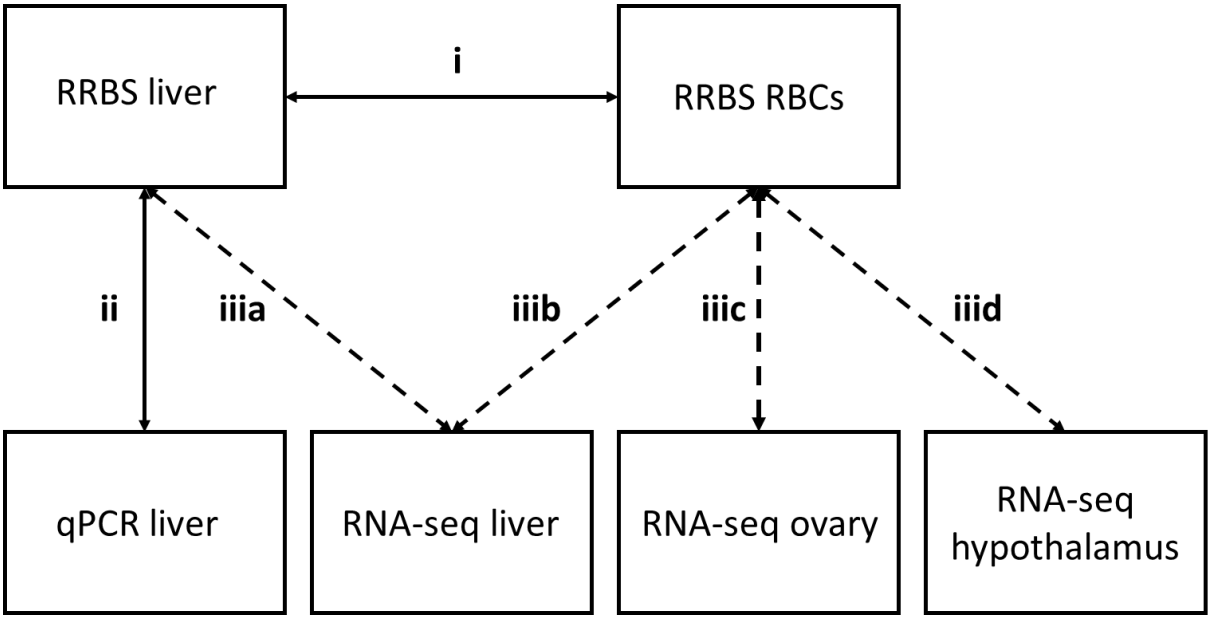


Figure 6.1. Overview of the data used in this study and how they were linked. Solid lines refer to associations in which only data from individual female great tits was used, while dashed lines refer to associations in which both individual (RRBS) and pooled (RNA-seq) data was used. Number-character combinations indicate the aims (see '*Introduction*') of the study and the methods used (see '*Methods*' for details).

(i) Blood methylation and liver methylation correlation analysis (in promoter and TSS regions)

We tested the difference in mean and variance in methylation between samples with a Kruskall-Wallis test ($p<0.05$) and Fligner Killeen test ($p<0.05$), respectively (Table S6.2). Subsequently, we standardized methylation proportions for liver and blood separately, by calculating z-scores, i.e. subtract the mean of all CpG site methylation levels from each individual CpG site methylation level, divided by the standard deviation of all CpG site

methylation levels. Sites present in both tissues were isolated from the dataset after which we calculated the mean methylation level (i.e. mean z-score) per site per time point for both liver and blood. Subsequently, we were able to calculate the change in methylation per site between time points within these tissues. Next, we conducted a differential methylation analysis ('methylKit' package, Akalin et al. 2012) on the common sites in order to find differentially methylated sites (DMS) between time point 1 and 2 ($\Delta_{1,2}$), and time point 2 and 3 ($\Delta_{2,3}$) in either blood or liver (tissue-specific change) or in both tissues (tissue-general change). We considered a site significantly differentially methylated between time points with a difference in methylation percentage of ≥15% and a q-value ≤0.01. We used the Pearson's correlations coefficient (*r*), as it measures linear trends, to evaluate the relationship between DNA methylation in blood and liver. For sites that were situated in the promoter and TSS regions of genes in both RBCs and liver we went through the same process.

Additionally, a gene ontology (GO) analysis was performed on the genes associated with the tissue-specific and tissue-general changing DMS to explore which functional groups (GO terms) are over-represented (Gaudet & Dessimoz 2017; Primmer et al. 2013) and possibly linked to timing of breeding. The DMS could be associated to 3350 unique great tit genes (NCBI Parus major genome version 1.1). GO analysis was performed using Cytoscape plugin ClueGo 2.5.4 (Bindea et al. 2009). Using kappa statistics, ClueGo constructs and compares networks of GO terms. A two-sided hypergeometric test (Rivals et al. 2007) was applied with GO term fusion, Kappa score and network specificity were kept at default values. The GO tree interval was set at 3 (min. level) and 10 (max. level), the GO term/pathway selection at 5% and false discovery correction was performed using the Benjamini-Hochberg step-down method. Both the human (23.06.2019) and chicken (23.06.2019) gene ontologies and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (Kanehisa & Goto 2000) database were used.

(ii) Liver methylation and liver candidate gene expression correlation analysis within individuals

We selected the methylation levels for sites in both TSS and promoter regions within genes in liver that are either key to reproductive functioning (i.e. in relation timing of breeding) or reference genes (i.e. to normalize qPCR expression data) (Tables S6.3 and S6.4) and of which there is also qPCR gene expression data available (**Chapter 5**) (Tables S6.5). For five candidate genes we found CpG sites with 10x coverage across all samples from the total gene set analyzed in liver (nine genes of interest and four reference genes, **Chapter 5**) in order to evaluate the association between DNA methylation and RNA expression in TSS regions: beta-2-microglobulin (*B2M*), glucocorticoid receptor (*GR*), heat shock protein family B (small) member 1 (*HSPB1*), mineralocorticoid receptor (*MR*) and protein 2 kinase C alpha (*PRKCA*). In addition to these genes, ribosomal protein 19 (*RPL19*) and succinate dehydrogenase complex flavoprotein subunit A (*SDHA*) could be evaluated for promoter regions. Per gene, we calculated $\Delta_{1,2}$ and $\Delta_{2,3}$ for both expression and methylation levels. For example, from the methylation level of an individual female in time point 2, methylation

levels of all females in time point 1 (n=6) were subtracted. Subsequently, these six values were used to calculate the average change in methylation per female in time point 1 across all females from time point 2, and vice versa (see Figure S1 for a visualization). The same process was repeated for expression levels. Pearson's correlations were used to evaluate relations between the average change in expression and methylation levels. P-values were adjusted for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini & Hochberg 1995).

(iii) Correlation of change in DNA methylation and gene expression using RRBS and RNAseq data

Here, we used RRBS data of individual females (n=6 females per time point) and RNA-seq data of pools of the same females (n=2 pools per time point, n=3 females per pool) to relate changes in CpG site methylation to changes in expression of the associated gene. We examined how (a) the change in liver methylation related to the change in liver gene expression and how the change in RBC methylation related to change in (b) liver, (c) ovary, and (d) hypothalamus gene expression (Figure 6.1). Data sets for each comparison (a-d) were trimmed such that CpG sites without an associated gene in the RNA-seq data were excluded from the RRBS data and genes without an associated CpG site in the RRBS data were excluded from the RNA-seq data. A gene and a CpG site were defined as 'associated' if a CpG sites was located within the TSS, promoter regions, gene body, 10k up- or downstream region of the gene. To prevent analyzing a CpG site or a gene within the same tissue more than once, we merged the trimmed RRBS data sets used for comparison b-d (i.e. all comparisons including RBC methylation) into one RBC RRBS dataset excluding CpG site duplicates and merged the RNA-seq data sets for comparison a and b (i.e. all comparisons including liver gene expression) into one liver RNA-seq dataset excluding gene duplicates. We only used CpG sites and genes that showed a significant time effect throughout the three sampling time points to examine how change in methylation related to change in gene expression. CpG sites with significant time effect were identified using a differential methylation analysis and genes with significant time effect were identified using a differential gene expression analysis performed in **Chapter 4**.

Differential methylation analyses were performed for 297,916 CpG sites in RBCs and 529,717 CpG sites in liver using package 'methylKit' (Akalin et al. 2012). We tested whether the full model, including time point and temperature environment as fixed effect, explained the methylation profile of a site better than the null model only including the temperature environment as fixed effect. We considered a time effect to be significant for sites with q-value ≤0.01. Differential expression analysis is described in detail in **Chapter 4**. In short, main effect models for time point and selection line were tested using the standard DeSeq2 protocol (Love et al. 2014) and a likelihood ratio test such that the main effect models were compared to a model excluding the main effect. Models were performed separately for each tissue. Genes present in the trimmed RNA-seq data sets for liver, hypothalamus, and ovary tissue with an adjusted p <0.05 when testing the main effect model for a time point effect and with an adjusted p>0.05 when testing the main effect model for a selection line effect

(as data from both selection lines were included when testing the main effect models) based on **Chapter 4** were classified as significantly differentially expressed genes.

To examine the association between DNA methylation change and change in gene expression in tissue comparisons a-d, we quantified $\Delta_{1,2}$ (first to second sampling time contrast) and $\Delta_{2,3}$ (second to third sampling time contrast) for both the methylation level of CpG sites and gene expression levels. We quantified the change in methylation by calculating the average methylation levels (i.e. methylation proportion × 100) across females for all three time points and then calculated the difference between the respective time points. We quantified the change in gene expression by calculating the $\log_2$Foldchange for the first and second time contrast using DeSeq2. Here we only used the differentially expressed genes identified in the full time point model from **Chapter 4** where the reported log2Fold changes between time points 1 and 3. We furthermore trimmed the data sets by excluding CpG sites with a change in methylation level<5% and genes with a change in $\log_2$Foldchange<0.5 between the respective time points. To better understand the effect of the genomic location on the relationship between changes in DNA methylation and changes in gene expression, we differentiated between genomic locations (i.e. TSS, promoter regions, gene body and 10k up- and downstream). For each combination of comparison (a-d), time contrast (first and second), and genomic location we plotted the $\log_2$Foldchange against the change in methylation level. There are four possible quadrants of association between $\log_2$Foldchange and change in methylation level: hypo-methylation and increased gene expression (Q1), hyper-methylation and increased gene expression (Q2), hyper-methylation and decreased gene expression (Q3), and hypo-methylation and decreased gene expression (Q4). While Q1 and Q3 would relate to changes in the predicted directions (based on the expectation that methylation and expression are negatively correlated), Q2 and Q4 would relate to changes opposite to the predicted directions. For the correlation between RBC methylation and gene expression in ovary, we tested whether associations within the TSS were more often within quadrants Q1 or Q3 than within quadrants Q2 or Q4 using a Fisher's exact test in which we compared the proportion of associations within quadrants Q1 or Q3 between the TSS and 10k downstream region. We used the 10k downstream region as a control region for CpG sites randomly distributed across Q1-Q4 as we do not expect any structural relationship between methylation and gene expression in this region (Laine et al. 2016; Derks et al. 2016). We did not use the 10k upstream region or gene body as control regions as the 10k upstream region overlaps with the promoter region (in which we would rather expect a relationship between methylation and gene expression) and at least parts of the gene body are hypothesized to show a relationship between methylation and gene expression.

Results

(i) Correlation between common DMS in RBCs and liver.

There were 302,647 sites in both RBCs and liver for which all females had ≥10x coverage (Table S6.6), of which 2377 and 3934 sites showed a differential $\Delta_{1,2}$ (Table S6.7) or $\Delta_{2,3}$ (Table S6.8) respectively, in either RBCs and liver or both RBCs and liver. Both DMS changing in a tissue-specific way and DMS changing in a tissue-general way show a strong correlation between both time point 1 and 2 ($r = 0.77$, df = 2375, $p < 0.0001$, Figure 6.2A) and time point 2 and 3 ($r = 0.75$, df = 3932, $p < 0.0001$, Figure 6.2B).

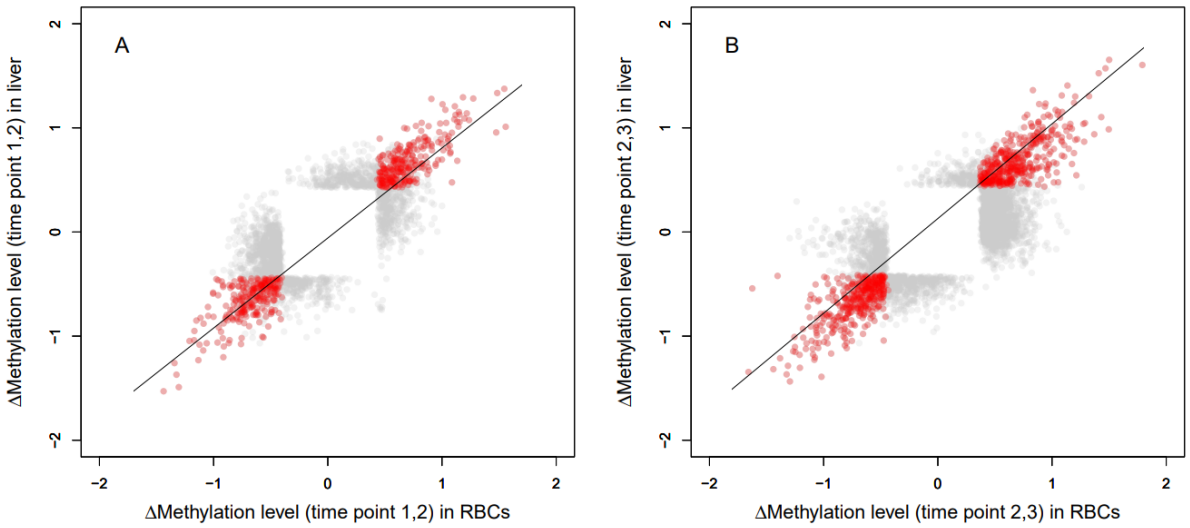


Figure 6.2. Correlation between RBC and liver of sites that show a differential change in methylation for $\Delta_{1,2}$ (A) and $\Delta_{2,3}$ (B). We depict sites that significantly change in methylation in both tissues (tissue-general) in red or in one of the tissues (tissue-specific) in grey. We applied transparency because of the high number of overlapping data points. Line is the regression line.

Correlation between DMS in RBCs and liver in promoter and TSS regions specifically

Within promoter regions, there were 108298 sites present in both RBCs and liver, of which 221 and 457 sites were differentially methylated (in one or both tissues) for $\Delta_{1,2}$ and $\Delta_{2,3}$, respectively. DMS that changed both tissue-specific and tissue-general show a strong correlation between both time point 1 and 2 ($r = 0.74$, n = 219, $p < 0.0001$, Figure 6.3A) and time point 2 and 3 ($r = 0.70$, df = 455, $p < 0.0001$, Figure 6.3B).

Within TSS, there were 41591 sites in both RBCs and liver, of which 24 and 65 sites were differentially methylated (in one or both tissues) for $\Delta_{1,2}$ and $\Delta_{2,3}$, respectively. DMS that changes both tissue-specific and tissue-general show a strong correlation between both time point 1 and 2 ($r = 0.71$, df = 22 $p = 0.0001$, Figure 6.3A) and time point 2 and 3 ($r = 0.62$, n = 63, $p < 0.0001$, Figure 6.3B).

Overall, the number of DMS that changed between time points was higher in red blood cells compared to liver or both tissues. Also, the number of DMS is higher for $\Delta_{2,3}$ compared to $\Delta_{1,2}$ (Table S6.9).

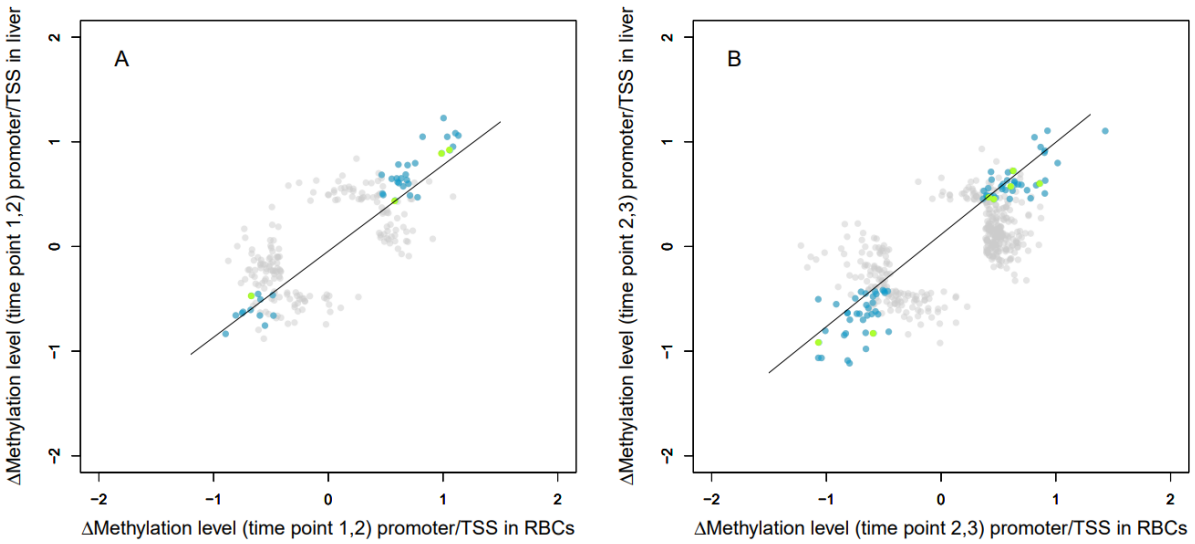


Figure 6.3. Correlation between those sites in promoter and TSS regions of RBC and liver that showed a differential change in methylation for $\Delta_{1,2}$ (A) and $\Delta_{2,3}$ (B). Sites that change significantly in methylation in both tissues (tissue-general change) in promoter and TSS regions are shown in blue and green, respectively. Sites that change significantly in methylation in one of the tissues (tissue-specific change), independent of gene region, are shown in grey. We applied transparency because of the high number of overlapping data points. Line is the regression line.

GO analyses

The tissue-specific and tissue-general changing DMS found in gene body + 10k up- and downstream (Figure 6.2), promoter and TSS (Figure 6.3) covered a total of 3350 unique great tit genes (Table S6.10). DMS that change in a tissue-general manner covered 1153 unique great tit genes (Table S6.10), whereas tissue-specifically changing DMS covered 2352 and 1408 unique great tit genes in red blood cells and liver, respectively, independent of gene region (Table S6.10). Using the chicken GO database, we found one and five significant GO terms associated with the genes associated with DMS that change in a tissue-general and tissue-specific manner, respectively (Table S6.11). With the human GO database, we found two and 36 significant GO terms associated with the genes related to DMS that change in a tissue-general manner and tissue specific manner, respectively (Table S6.11). One significant KEGG term was found using the human KEGG database (Table S6.11). Overall, the GO terms and KEGG term related to a wide range of functions (Table S6.11).

(ii) Correlation between change in methylation and change in candidate gene expression in liver

Per gene, the number of sites ranged between 3-15 and 6-54 with ≥10x coverage, in TSS and promoter regions respectively (Table S6.12). No significant correlations were found within genes between the change in DNA methylation and change in RNA expression between time points for both TSS and promoter regions (Table S6.13, Figures S6.2-S6.5).

(iii) Correlation between change in methylation and genome-wide change in gene expression

We identified 243 CpG sites in liver (Table S6.14) and 2256 CpG sites in RBCs (Table S6.15) with a significant change in methylation level over time points and 63 genes in hypothalamus (Table S6.16), 1073 genes in ovary (Table S6.17) and 143 genes in liver (Table S6.18) with a significant change in gene expression over time points. How changes in gene expression were associated to changes in methylation is displayed in Figure 6.4 and Figures S6.6-S6.13 for all tissue comparisons; (a) change in liver methylation related to the change in liver gene expression and change in RBC methylation related to (b) change in liver, (c) ovary, and (d) hypothalamus gene expression. Associations within the gene body, 10k up- or downstream region, and promoter region were located within all four quadrants (Q1-Q4) without an enrichment for either combination of quadrants (i.e. Q1 and Q3 or Q2 and Q4, Figure 6.4 and Figures S6.6-S6.13) irrespective of the tissue comparison (a-d). Although the number of associations between genes and CpG sites was limited (max. four associations per tissue comparison), associations within the TSS were exclusively located within Q1 and Q3 for tissue comparison (a) change in liver methylation related to the change in liver gene expression and change in RBC methylation related to change in gene expression (b) in liver and (d) in hypothalamus (Figures S6.6-S6.13). For tissue comparison (c) change in RBC methylation and change in gene expression in ovary (Figure 6.4), associations in Q1 or Q3 were overrepresented for the change of sites within the TSS of a gen between time point 2 and 3 when compared to sites within the 10k downstream region (Fisher's Exact Test: p = 0. 0.0012, Figure 6.4B). This trend was a non-significant for the change between time point 1 and 2 (Fisher's Exact Test: p = 0.11, Figure 6.4A). Genes, number of associated CpG sites, and number of association within quadrants Q1 or Q3 and in quadrants Q2 or Q4 are listed for each combination of comparison (a-d), time contrast (first and second), and genomic location in Tables S6.19-S6.26.

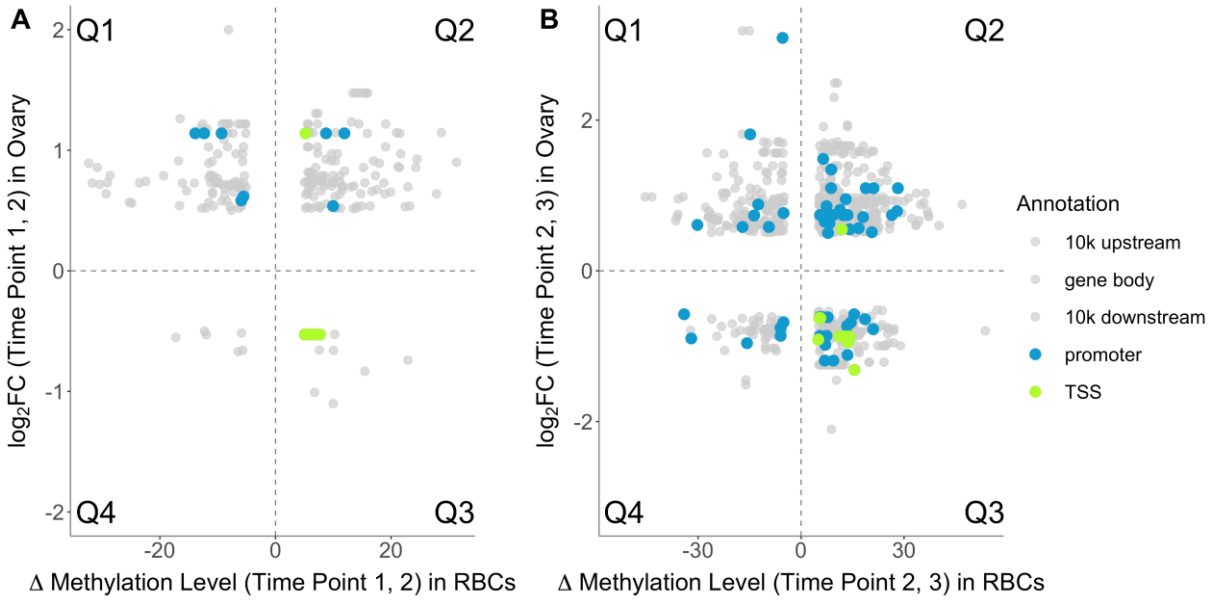


Figure 6.4. Log_2 Foldchange (log_2 FC) of genes in ovary in relation to change in methylation level of a CpG site in red blood cells within the TSS (green), promoter region (blue) or 10k up- and downstream region and gene body (all grey) of that gene for $\Delta_{1,2}$ (A) and $\Delta_{2,3}$ (B). The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled using Q1-Q4. Transparency is applied to the grey data points such that the area of overlap between plots appears darker.

Discussion

Evidence whether blood-derived measurements of DNA methylation could be used as a proxy for other tissues is growing but still incomplete, especially in an ecological context. We investigated whether DNA methylation changes over time points were either tissue-specific or tissue-general and whether they associated with expression changes in the target tissue. We found a strong positive correlation between changes in methylation of differentially methylated sites (DMS) in RBCs and changes of these DMS in liver, both when using sites that are situated throughout the genome and when restricting to sites situated within the promoter region or the TSS of annotated genes.

For single candidate genes, we found no correlations between the change in CpG site methylation and the change in RNA expression between time points. Genome-wide, we found a high correlation in TSS between CpG site hyper-methylation and a decrease in the associated gene in the ovary. In addition, as expected, no association was found between changes in CpG site methylation and expression within the gene body or in the 10k up- or downstream region of the respective gene, irrespective of which tissues were compared when investigating the sites that changed significantly over time in both methylation and the associated gene's expression.

Here, we discuss four distinguishable groups of DMS regarding methylation change between time points and their potential association with gene expression. The first two groups are DMS (1) showing a *tissue-specific change* (Figures 6.2-6.4) that are expressed in a tissue-specific manner or (2) in a tissue-general manner. These groups cannot be used as biomarkers (for temporally expressed traits), because of their tissue-specific change in

methylation and/or expression. Although there is a growing body of studies investigating tissue-specific methylation, these studies are mostly in relation to aging and diseases (e.g. Lokk et al. 2014; Wang et al. 2018; Xie et al. 2018; Zhu et al. 2018). Here, tissue-specific methylation will not be informative in relation to temporally expressed traits. Further, these studies often do not elucidate the mechanism(s) by which methylation changes are induced or its functional consequence. It is likely that the (de)methylation mechanism underlying these tissue-specific changes are also tissue-specific. There is some evidence that methylation patterns in tissues are more similar when these tissues are derived from for example the same germ layer (Smith et al. 2015) and that the rate of cell division contributes to tissue-specific methylation profiles (Dmitrijeva et al. 2018). However, whether this relates to tissues-specific changes, remains to be established.

The other two groups are DMS (3) showing a *tissue-general change* (Figures 6.2-6.4) that are expressed in a tissue-specific or (4) a tissue-general manner. Both groups can potentially be used as biomarkers (for temporally expressed traits), because they change in a similar way across tissues (or at least here, in RBCs and liver) and extrapolation from one tissue to another may be possible. However, for the fourth group the association between the methylation change and expression change should be established in the target tissues in order for the methylation patterns to be predictive of the expression changes. In group 3, the environment could for example cause the release of hormones with system-wide effects, which may have common effects on DNA methylation across tissues, but that differ in magnitude (Klengel & Binder 2015). An example of such a common effect is the activation of the glucocorticoid receptor (GR) gene. When stress activates the activation of the hypothalamic-pituitary-adrenal axis, cortisol is globally increased. Although GR binding sites show tissue-specificity, their activation is shown across tissues (John et al. 2011). As such, activation of GR may lead to epigenetic changes across tissues, as shown in both humans and rodents (Ewald et al. 2014; Zhang et al. 2013). In case of group 4, there could be a general increase in body-wide DNMT activity, catalysing DNA methylation and preserving methylation after cell division in a tissue-general manner.

Further, it is important to realize that certain tissues, like the brain, liver and ovary, play key roles in traits such as timing of breeding and stress responsiveness, and could have very specific signalling pathways, whereas others are common across tissues (Klenger & Binder 2015). Additionally, in complex tissues, epigenetic mechanisms also differ according to tissue regions, sub-tissue regions, and cell types, as shown previously in human brain (Davies et al. 2012; Smith et al. 2015). Thus, even though methylation changes in RBCs could potentially predict a part of the methylation change in other tissues, results from epigenetic studies in peripheral blood have to be interpreted with great care with regard to their reflection of epigenetic patterns in highly heterogeneous tissues.

Exploring whether the genes covered by the DMS that show either a tissue-specific or tissue-general change in the different gene regions would associate with certain functional groups or GO terms (related to timing of breeding), resulted in several GO terms related to a wide range of biological processes. However, for most of the genes associated to tissue-general changing DMS and the smaller gene regions, no GO terms and pathways were

found. Here, it is likely that the number of associated genes, which is limited especially for the promoter and TSS region, resulted in a power issue. We found no GO term clearly pointing towards timing of breeding though we like to point out that this does not mean that the GO terms found do not have a function in relation to timing of breeding. In addition, the chicken does not breed seasonally, let alone the human. As such, the onthologies for chicken and human, might, not include GO terms of functional relevance for seasonally regulated reproduction. Further, the GO databases do not store tissue-specific information, which could hamper the interpretation of the results found and as such the tissue specificity of these predictions cannot be validated.

We also investigated whether RBC methylation changes could predict candidate gene or genome-wide expression changes in other tissues. For a set of candidate genes, no correlations were found between the change in CpG site methylation and the change in RNA expression between time points. The genes we were able to analyze, irrespective of whether they were used as a reference gene (*PRCKA*, *RPL19*, *SDHA*) or genes of interest (*HSPB1*, *GR*, *MR*) were expressed very stable over time (**Chapter 5**). As such, it might not be surprising to not find a correlation between the change in methylation and expression for these specific genes. Previous within-individual studies in great tits have shown a negative association between TSS methylation in RBCs and associated gene expression in the brain (Derks et al. 2016; Laine et al. 2016) and that hypo-methylation at the TSS is enriched in genes with functional classes that relate directly to processes specific to each tissue type (Derks et al. 2016). Genome-wide, we find a similar trend, in which CpG site hyper-methylation within the TSS in RBCs was predominantly associated with a decrease in the expression of the respective gene, most pronouncedly in the ovary. As predicted, no specific trend was found in the other genomic locations, including the promoter. This poses the question about how to define the region surrounding the base pair at which gene transcription is initiated to best study how DNA methylation affects gene expression. We hypothesize that sites in the TSS could be hypo-methylated and change in a tissue-general manner, but are functional in only the tissues where the activated processes are performed (i.e. group 3, see above) and inactive in the tissues where they are not. For example, genes could be activated in ovary, but inactivated by regulatory mechanisms other than DNA methylation in RBCs (Campbell & Langlois 2018; Kassam et al. 2019; Miragaia et al. 2019).

We emphasize that the time points and tissues in this study were chosen in relation to timing of breeding to explore its underlying molecular mechanisms elsewhere (**Chapter 4** and **5**; Viitaniemi et al. 2019; Lindner et al. *in prep.*). RBCs are likely to have a limited biological function with regard to complex traits like timing of breeding. These biological functions will be in tissues within the HPGL axis, which regulates gonadal function and ultimately egg-laying. Recent studies in great tits, found temporal variation in genome-wide DNA methylation in RBCs collected throughout the breeding season (Viitaniemi et al. 2019) and a correlation between changes in DNA methylation levels and a female's reproductive stage (Lindner et al. *in prep.*). The CpG sites in these studies that show a time, treatment or reproductive stage-specific responses in DNA methylation are of interest for understanding to what extent DNA methylation acts as a mechanism that translates environmental signals

into a phenotypic response, e.g. timing of breeding. However, the link between methylation changes and changes in gene expression across tissues not clear. The overall strong correlation between methylation change in RBCs and liver needs to be interpreted carefully as this is not the same as showing that RBCs can be used as a proxy for methylation patterns in other tissues. This, because DMS underlying this association, include both DMS that change in a tissue-specific and a tissue-general manner (Figures 6.2 and 6.3), indicating both common and unique epigenetic alterations within tissues that likely reflect differential functions. However, despite that many DMS are tissue-specific and cannot be used as biomarkers, there is a potential for methylation patterns in RBCs to be informative for a proportion of the (seasonal) change in methylation patterns in liver.

Although we sampled tissues from individuals at three different time points, these are not *within*-individual repeated measures as opposed to another study in the same birds using repeatedly RBC sampling (Viitaniemi et al. 2019). It is not possible to repeatedly sample inaccessible tissues like the brain, and it is highly challenging for liver, and thus we used a *between*-individual approach as a proxy of *within*-individual sampling. However, the fact that we do find a strong correlation between changes in methylation levels in RBCs and liver and an association between RBC methylation in the TSS and gene expression in ovary, shows that these effects are strong enough to also be found with a between-individual approach.

Unfortunately, we were not able test the associations between gene expression and methylation changes in other candidate genes that are known to be key in reproductive functioning, as CpG sites within those genes did not have ≥10x coverage for all samples in the RRBS data. This also resulted in a limited number of sites for the correlation analyses between CpG site methylation and single or genome-wide gene expression, which could have resulted in not finding correlations, especially within TSS regions, as they are relatively small. Similarly, the RRBS data is based on individuals, whereas the RNA-seq data originates from pooled samples (**Chapter 4**). We used genes identified as differently expressed over time in hypothalamus, ovary and liver from the study that used individuals from both selection lines (n = 12 pools, **Chapter 4**). The number of pools for the e*arly* selection line used in the current study was limited (n=6), and hence, only genes with a time effect and no line effect were chosen.

We calculated the change in methylation level based on samples of individuals and the $\log_2$Fold-change in expression level based on pooled samples. As described in **Chapter 4**, most differentially expressed genes over time were found in the ovary, while numbers of differentially expressed genes were lower in hypothalamus and liver. As such, it is likely that we only had enough power to detect a significant association between the change in RBC methylation and change in ovary gene expression. For the other tissue comparisons (especially liver-liver, RBCs-liver, and RBCs-hypothalamus) there is a too low number of data points (0-4 CpG sites) within the TSS to draw any conclusions, because of the low number of CpG sites differentially methylated over time within the TSS in comparison to the other genomic locations (gene body, 10k upstream and downstream region).

Sites within the TSS show a lower methylation level in general than sites within other genomic locations and even small methylation levels (about 20%) within the TSS are associated to downregulation of the associated gene which is not found for sites within other genomic locations (Laine et al. 2016). Thus, analysing sites within the TSS for differential methylation, together with sites in the other genomic locations might cause a biological relevant change in methylation within the TSS to appear as statistically insignificant based on the high number of tests performed with sites in genomic locations that show changes in higher magnitudes.

We investigated whether methylation changes were tissue-specific or tissue-general and whether such methylation changes were associated with expression changes within and between tissues. Tissue-general changes in DNA methylation are potentially informative for changes in gene expression in inaccessible tissues. To predict the influence of environmental factors on target tissues using blood-derived changes in methylation measurements (in natural populations), and whether these changes in methylation are followed by changes in transcription and subsequent phenotypic responses, further genome-wide studies are warranted to investigate the relationship between temporal changes in tissue- and blood derived CpG site methylation patterns and the relationship between temporal changes in methylation and expression within and between tissues. More specifically, we need to determine (1) whether sites that change in a tissue-specific or tissue-general manner are also methylated as such, (2) whether tissue-specific and tissue-general change in methylation results in a similar change in gene expression. When these links are established, subsequently, the tissue-general change in DNA methylation could be potentially highly valuable in the search for genes and their networks underlying temporally expressed traits. Nevertheless, this study provides unique insights into temporal changes in methylation across tissues and how such changes relate to changes in expression patterns. It furthermore highlights the importance for distinguishing between tissue-specific and tissue-general DNA methylation changes, where the latter are informative for changes in gene expression in inaccessible tissues. As such, tissue-general changes open up the possibility to monitor within individuals how environmental conditions affect temporally expressed traits, even in natural populations.

Acknowledgements

We thank Christa Mateman and colleagues in the molecular lab at NIOO-KNAW, Fleur Gawehns-Bruning for bioinformatics support and the members of the Ecological Epigenetics Research Theme at NIOO-KNAW for discussing the results. This study was supported by an ERC Advanced Grant (339092 – E-Response to MEV) and by funding from the Research Council of Norway (239974) and Centre of Excellence (223257) awarded to AH. Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

Supplementary Information Chapter 6

Supplementary tables not presented here

Table S6.1. Details on library preparation and sequencing.

Table S6.3. Methylation levels per CpG site for females in time point 2 and 3 within the genes of which qPCR data is available.

Table S6.4. Methylation levels per sites for females in time point 2 and 3 within the genes of which qPCR data is available.

Table S6.7. CpG sites that showed a differential change between time point 1 and 2 in either RBCs, liver or in both tissues.

Table S6.8. CpG sites that showed a differential change between time point 2 and 3 in either RBCs, liver or in both tissues.

Table S6.10. Genes covered by tissue-general and tissue-general changing DMS, or both.

Table S6.11. GO terms found for genes associated with tissue-general and tissue-specific changing DMS, subdivided over gene regions and time points.

Table S6.14. CpG sites in liver with a significant change in methylation across time points.

Table S6.15. CpG sites in RBCs with a significant change in methylation across time points.

Table S6.16. Genes that significantly change across time points in hypothalamus.

Table S6.17. Genes that significantly change across time points in ovary.

Table S6.18. Genes that significantly change across time points in liver.

Table S6.19. Number of sites in liver changing in methylation in different gene regions between time point 1 and 2 within genes associated with either Q1 and Q3 or Q2 and Q4 in liver.

Table S6.20. Number of sites in liver changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in liver.

Table S6.21. Number of sites in RBC changing in methylation in different gene regions between time point 1 and 2 within genes associated with either Q1 and Q3 or Q2 and Q4 in hypothalamus.

Table S6.22. Number of sites in RBC changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in hypothalamus.

Table S6.23. Number of sites in RBC changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in ovary.

Table S6.24. Number of sites in RBC changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in ovary.

Table S6.25. Number of sites in RBC changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in liver.

Table S6.26. Number of sites in RBC changing in methylation in different gene regions between time point 2 and 3 within genes associated with either Q1 and Q3 or Q2 and Q4 in liver.

These tables and figures will become available online after publication

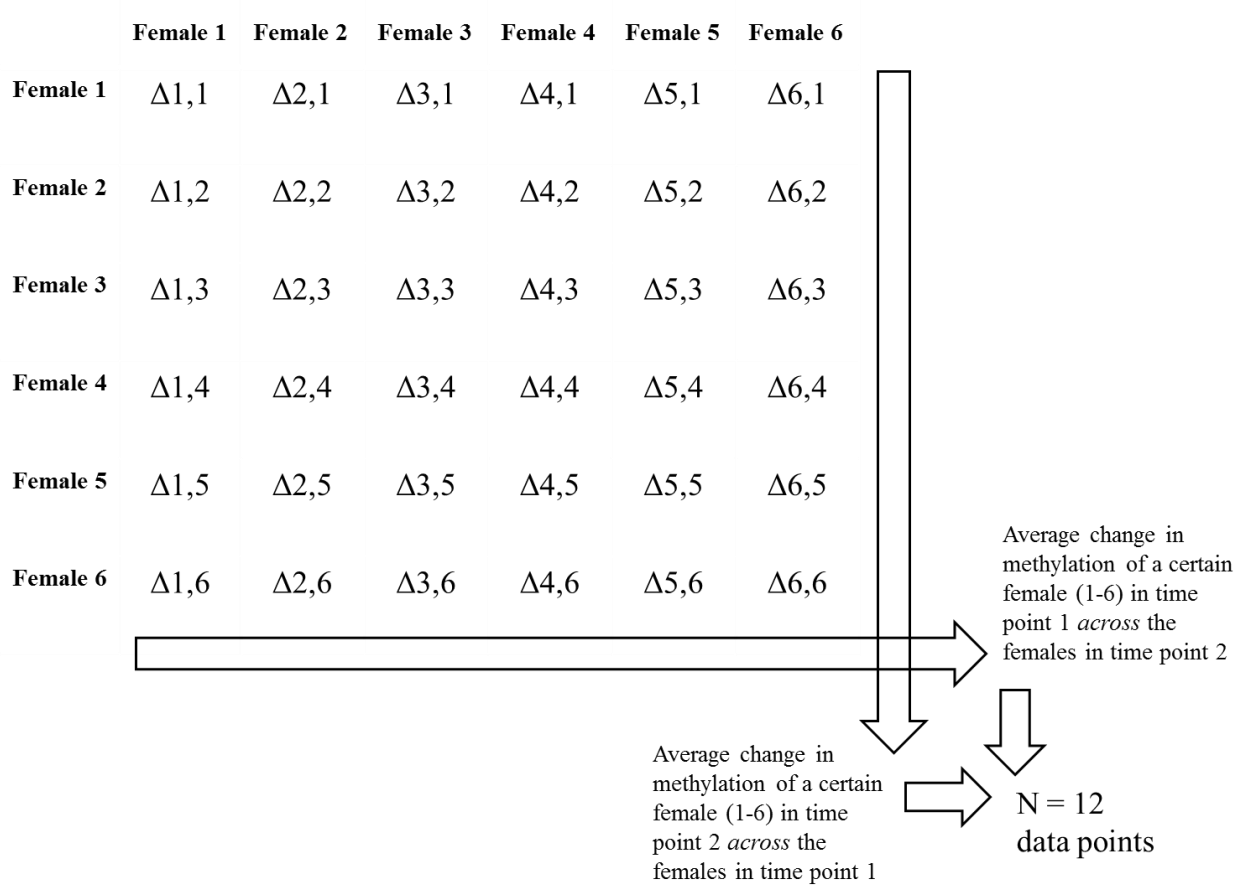


Figure S6.1. Methodology to calculate, per candidate gene, the change in methylation per site between time points by subtracting the CpG site methylation level of a female in, for example, time point 2 with all females in time point 1. Subsequently the average change per female in time point 1 across all females from time point 2 is calculated, and vice versa. This procedure applies also to the change between time point 2 and 3 and expression levels (see '*Materials and methods*').

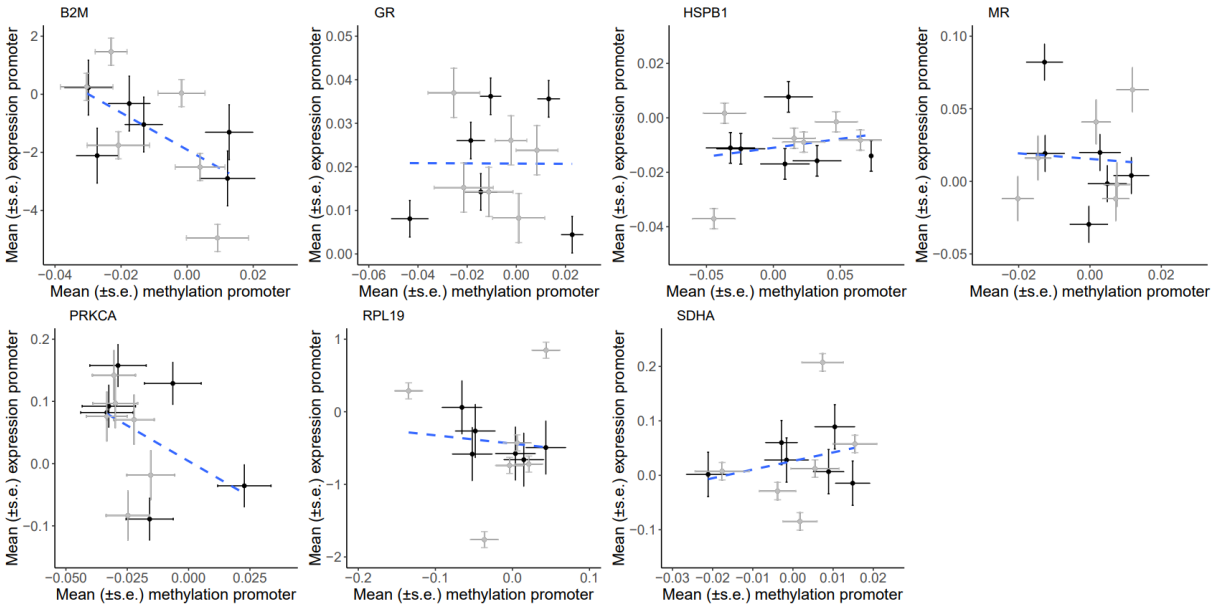


Figure S6.2. Mean (±s.e.) difference in both DNA methylation in promoter regions and RNA expression per female in time point 1 (in grey) across all females in time point 2, and vice versa (in black) for the candidate genes.

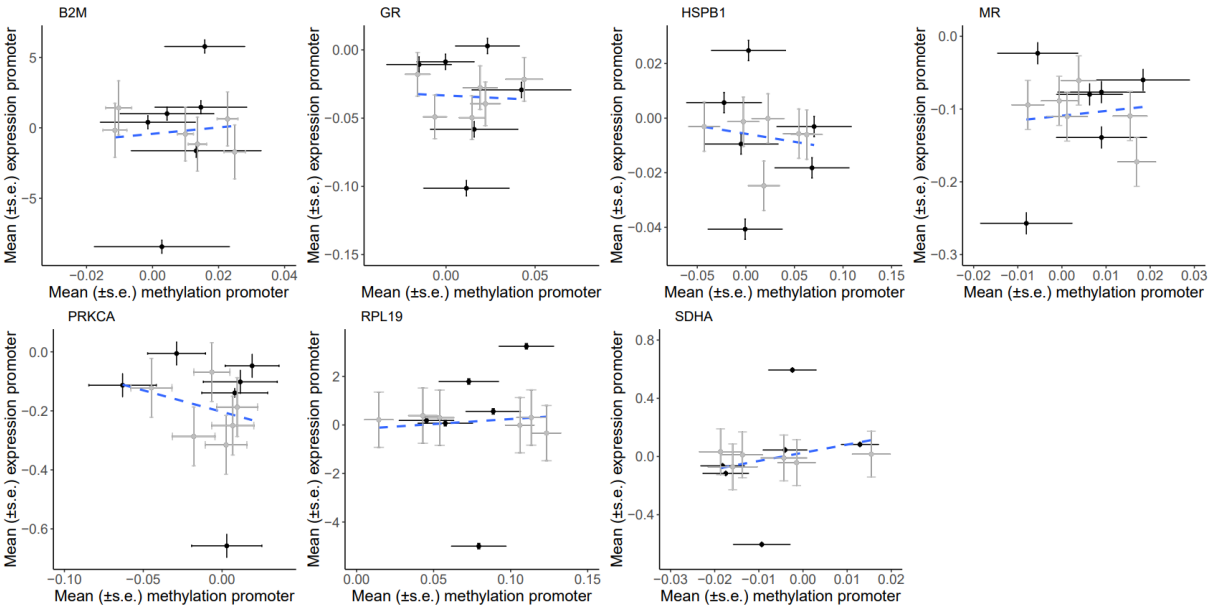


Figure S6.3. Mean (±s.e.) difference in both DNA methylation in promoter regions and RNA expression per female in time point 2 (in grey) across all females in time point 3, and vice versa (in black) for the candidate genes.

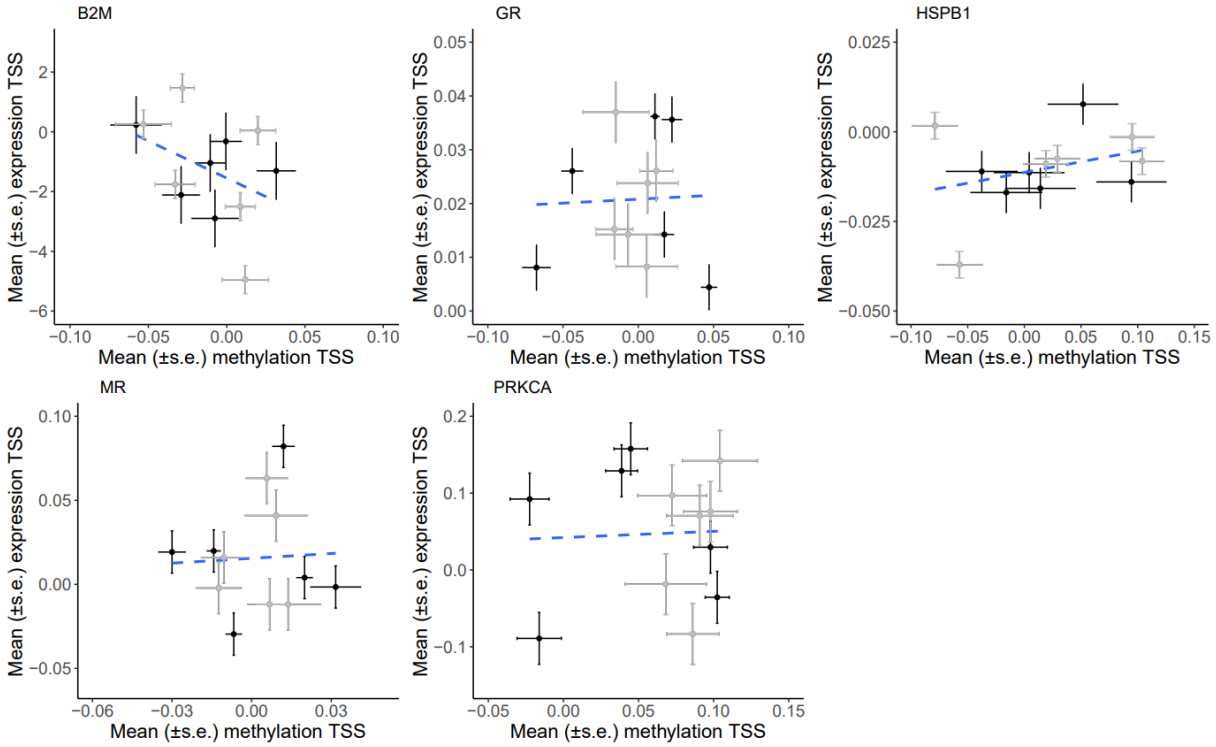


Figure S6.4. Mean (±s.e.) difference in both DNA methylation in TSS and RNA methylation per female in time point 1 (in grey) across all females in time point 2, and vice versa (in black) for the candidate genes.

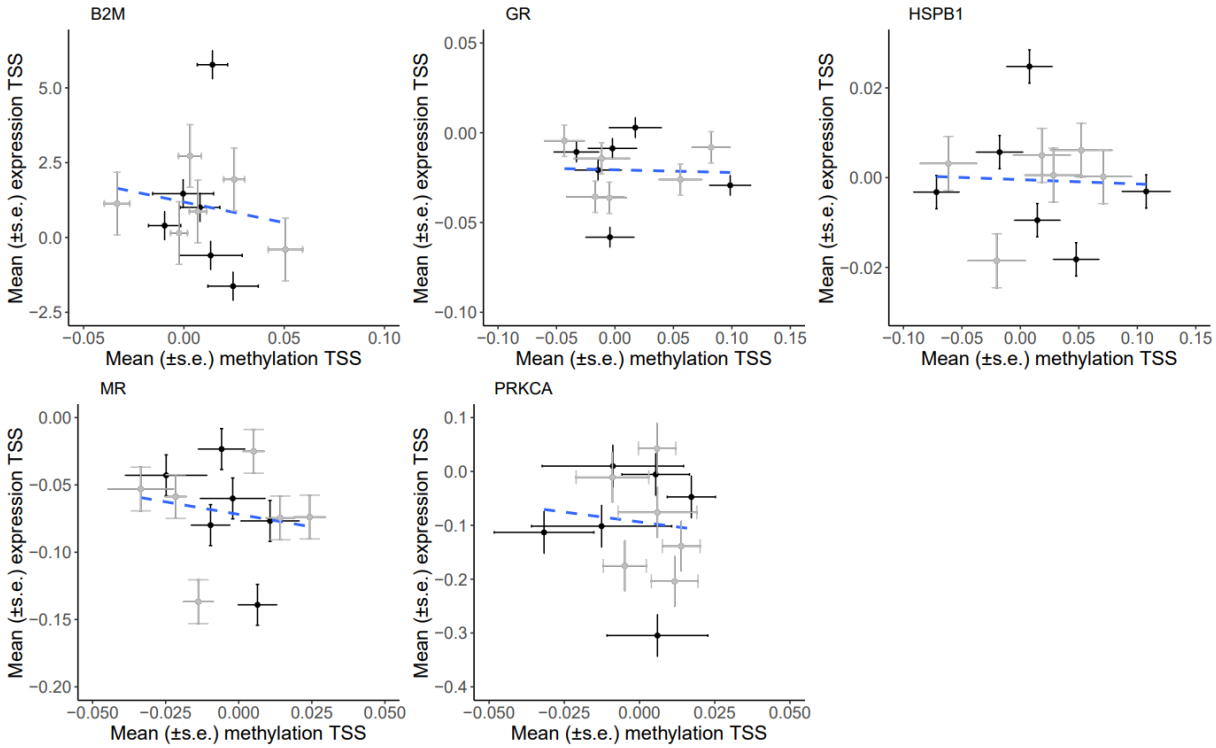


Figure S6.5. Mean (±s.e.) difference in both DNA methylation in TSS and RNA methylation per female in time point 2 (in grey) across all females in time point 3, and vice versa (in black) for the candidate genes.

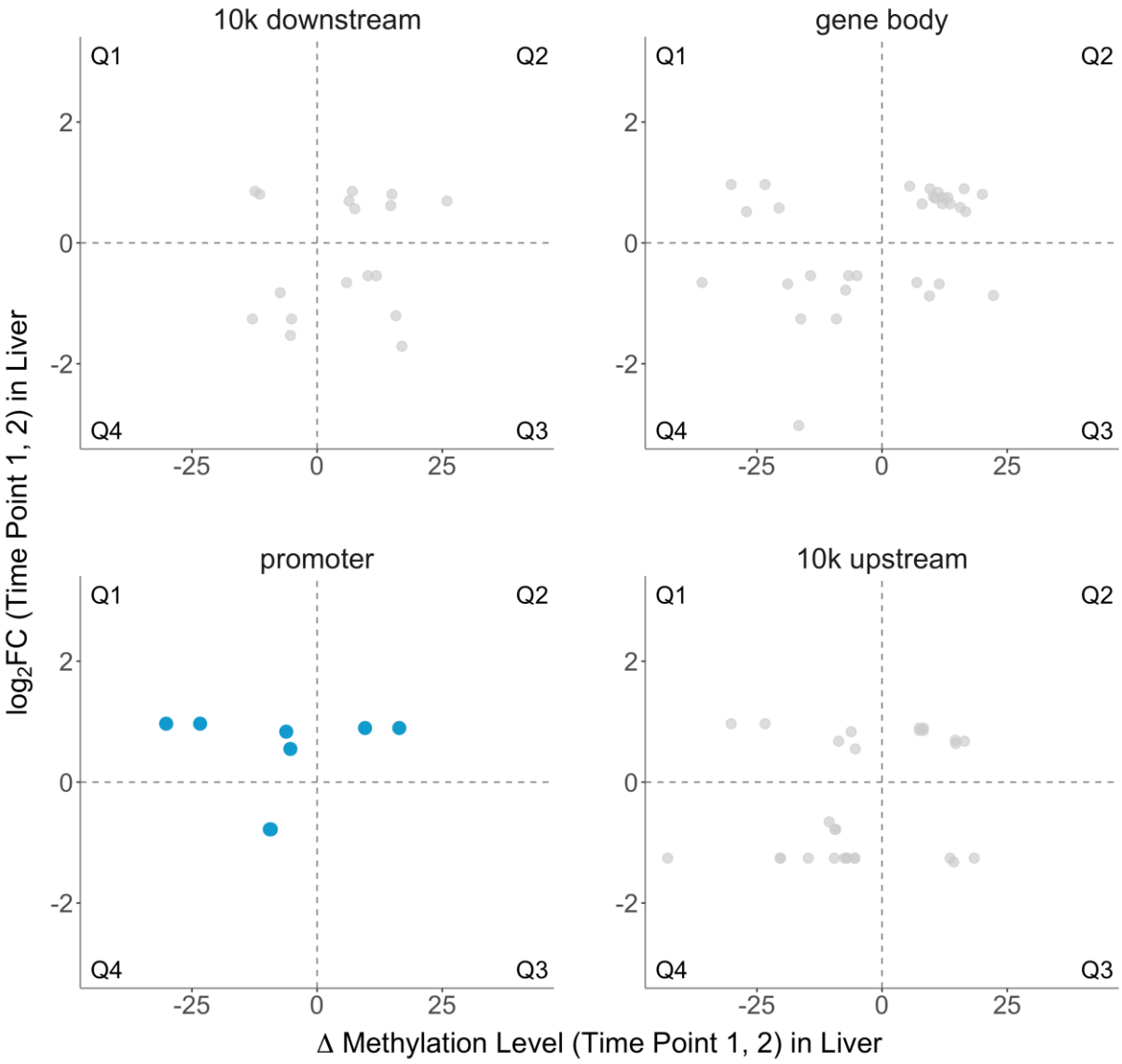


Figure S6.6. Log_2 fold change of genes in *liver* in relation to change in methylation level of a CpG site in *liver* within the 10k downstream region, gene body, promoter region, and 10k upstream region that gene for $\Delta_{1,2}$. Within the transcription start site (TSS) we did not find a significant change CpG site methylation located within a gene with significant change in expression. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

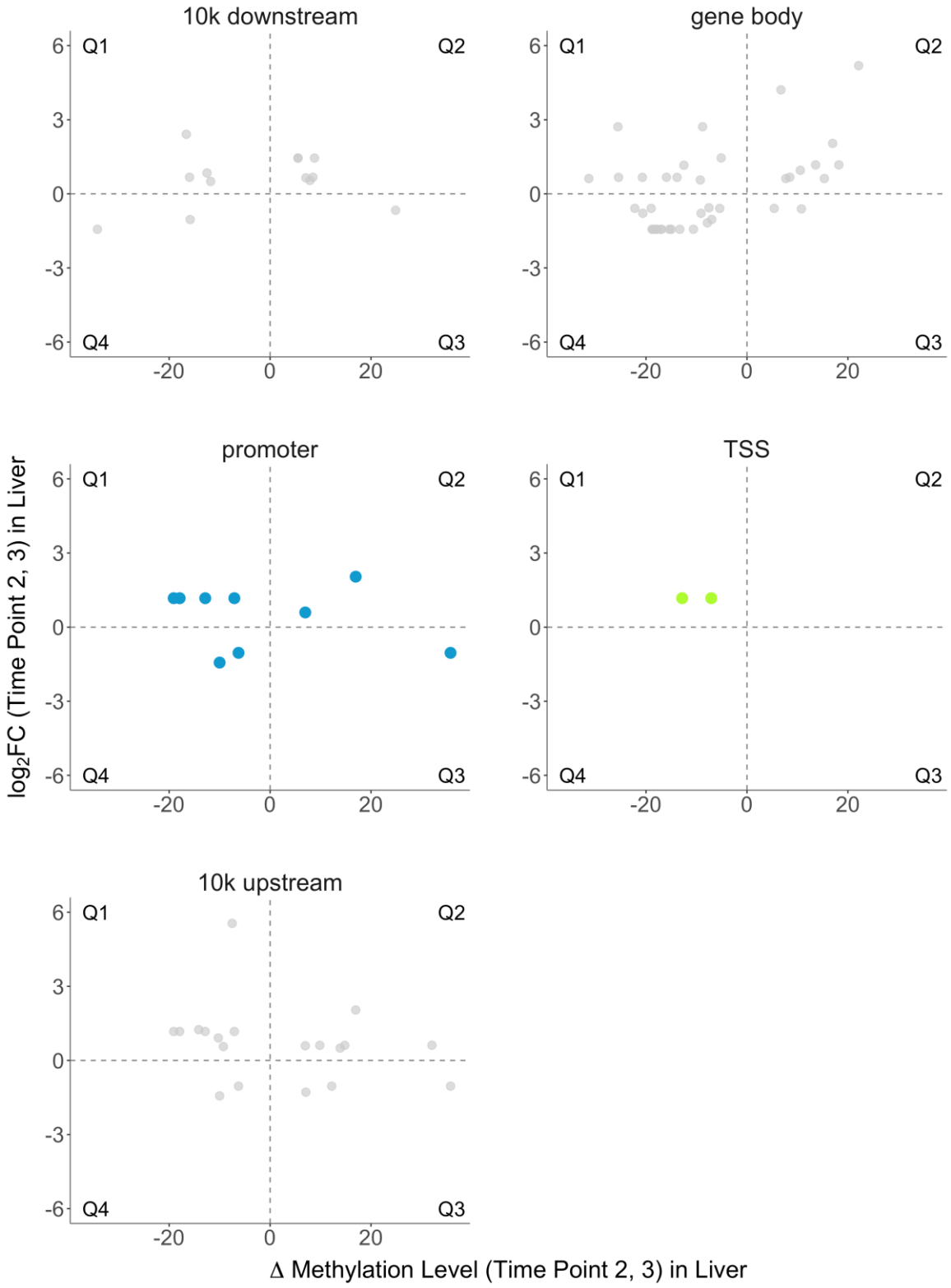


Figure S6.7. Log_2 fold change of genes in *liver* in relation to change in methylation level of a CpG site in *liver* within the 10k downstream region, gene body, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $\Delta_{2,3}$. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such hat the area of overlap of between data points appears darker.

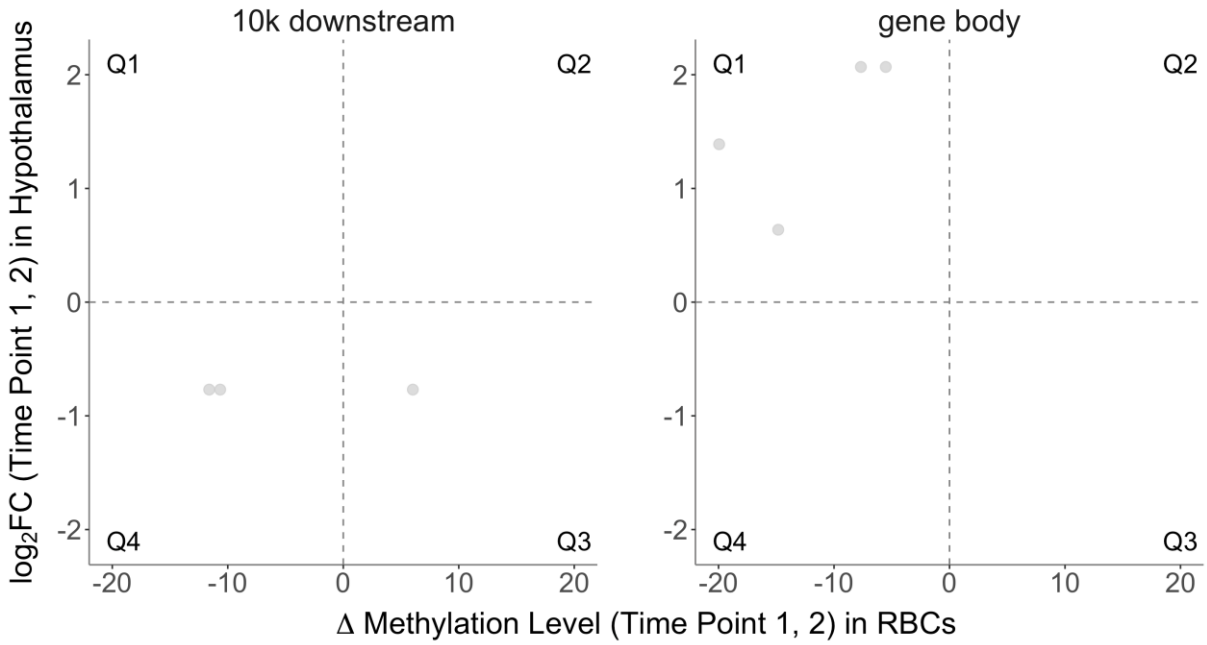


Figure S6.8. Log_2 fold change of genes in *hypothalamus* in relation to change in methylation level of a CpG site in *red blood cells* within the 10k downstream region and gene body of that gene $\Delta_{1,2}$. Within the 10k upstream region, promoter region, and transcription start site (TSS) we did not find a significant change CpG site methylation located within a gene with significant change in expression. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

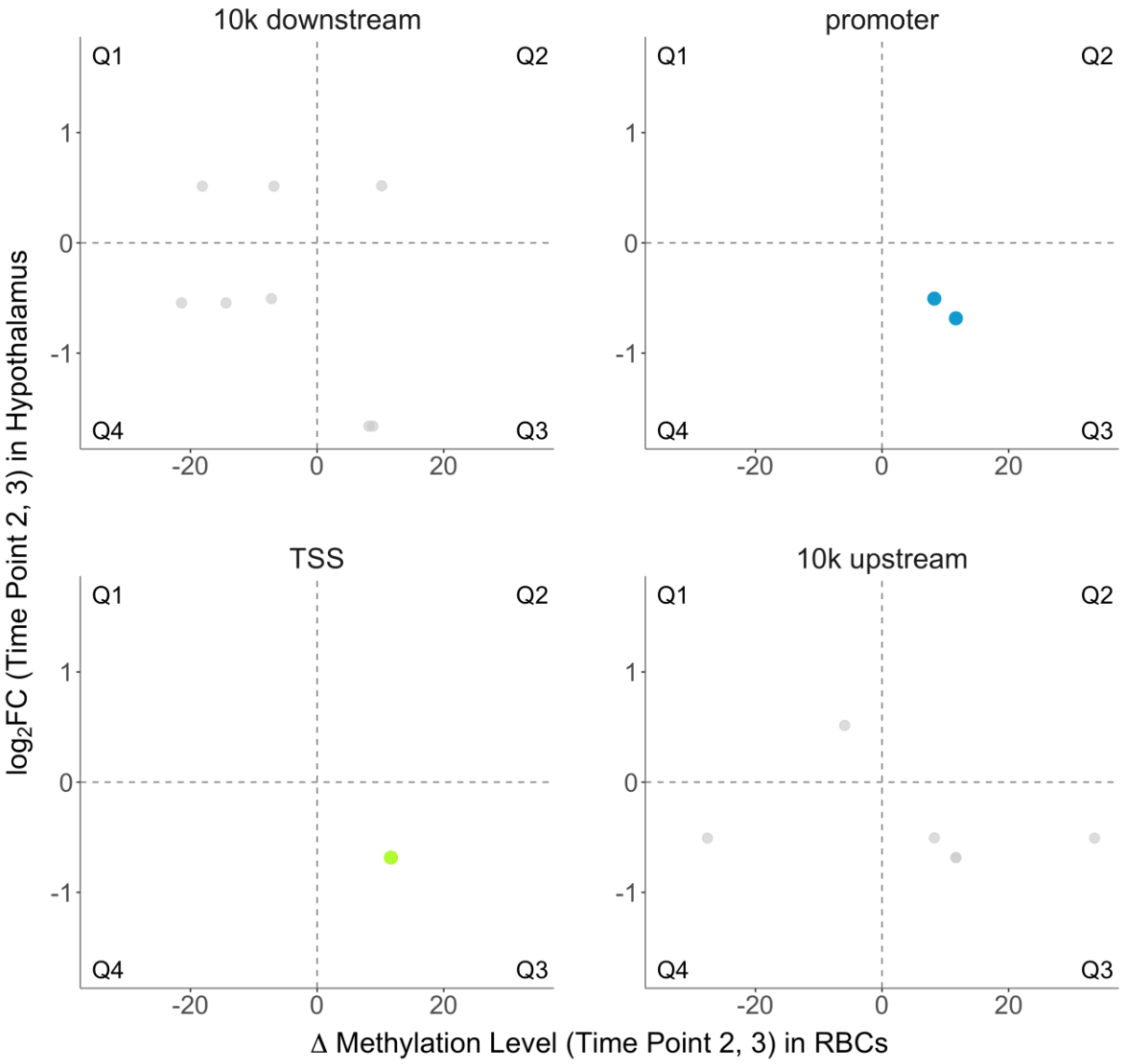


Figure S6.9. Log_2 fold change of genes in *hypothalamus* in relation to change in methylation level of a CpG site in *red blood cells* within the 10k downstream region, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $_{2,3}$. Within the gene body we did not find a significant change CpG site methylation located within a gene with significant change in expression. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

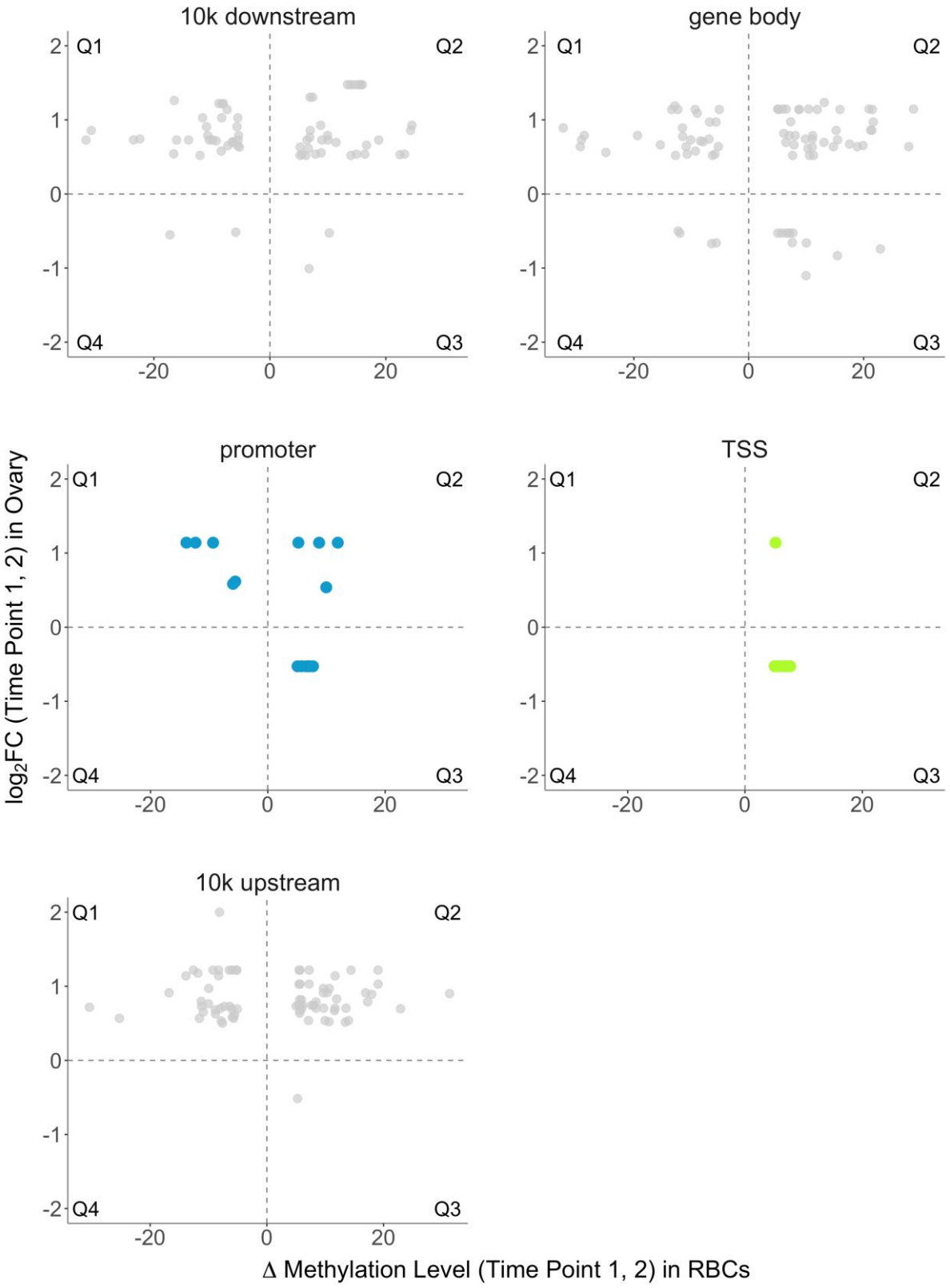


Figure S6.10. Log_2 fold change of genes in *ovary* in relation to change in methylation level of a CpG site in *red blood cells* within the 10k downstream region, gene body, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $\Delta_{1,2}$. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

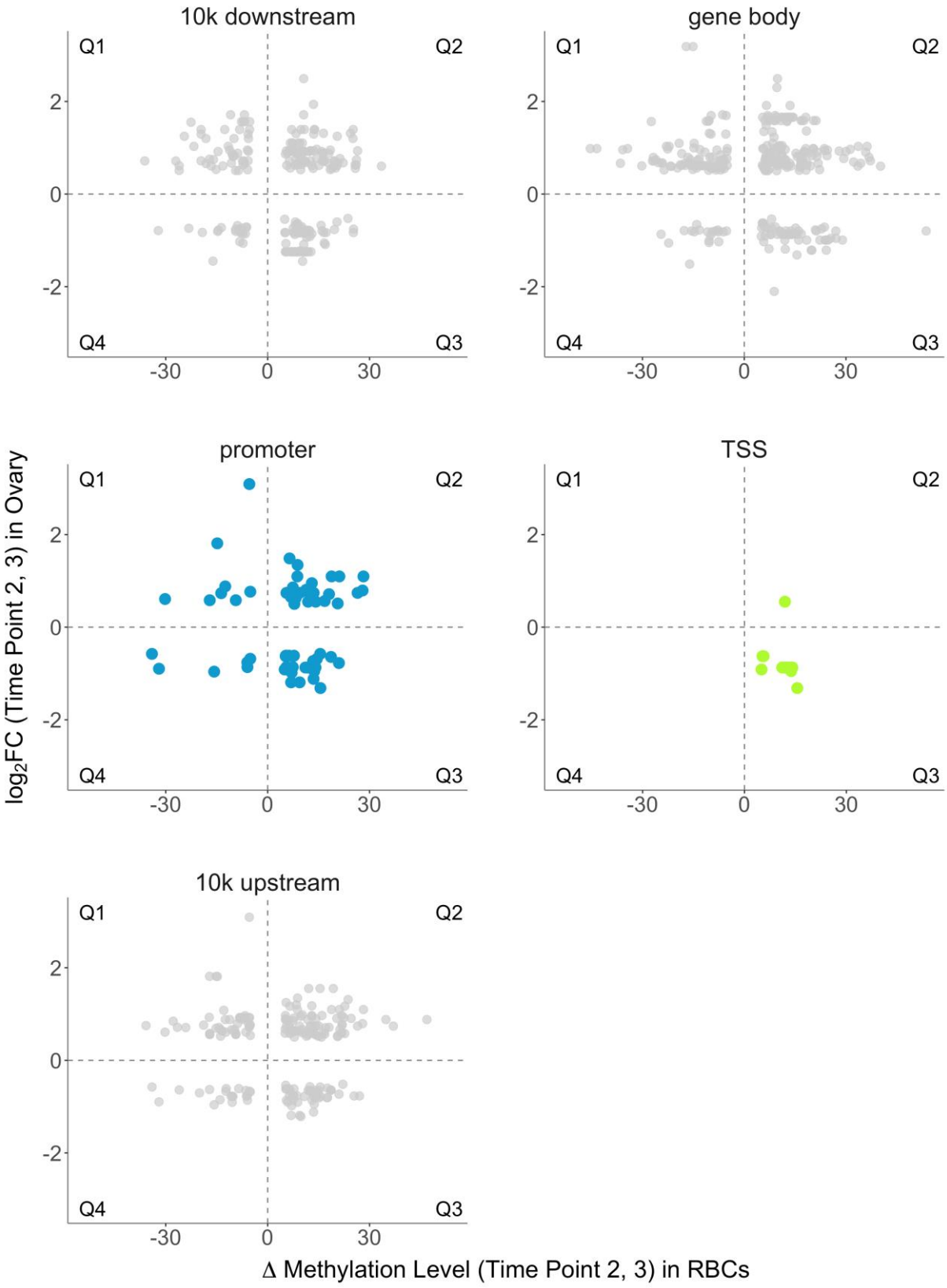


Figure S6.11. Log_2 fold change of genes in *ovary* in relation to change in methylation level of a CpG site in *red blood cells* within the 10k downstream region, gene body, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $\Delta_{2,3}$. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

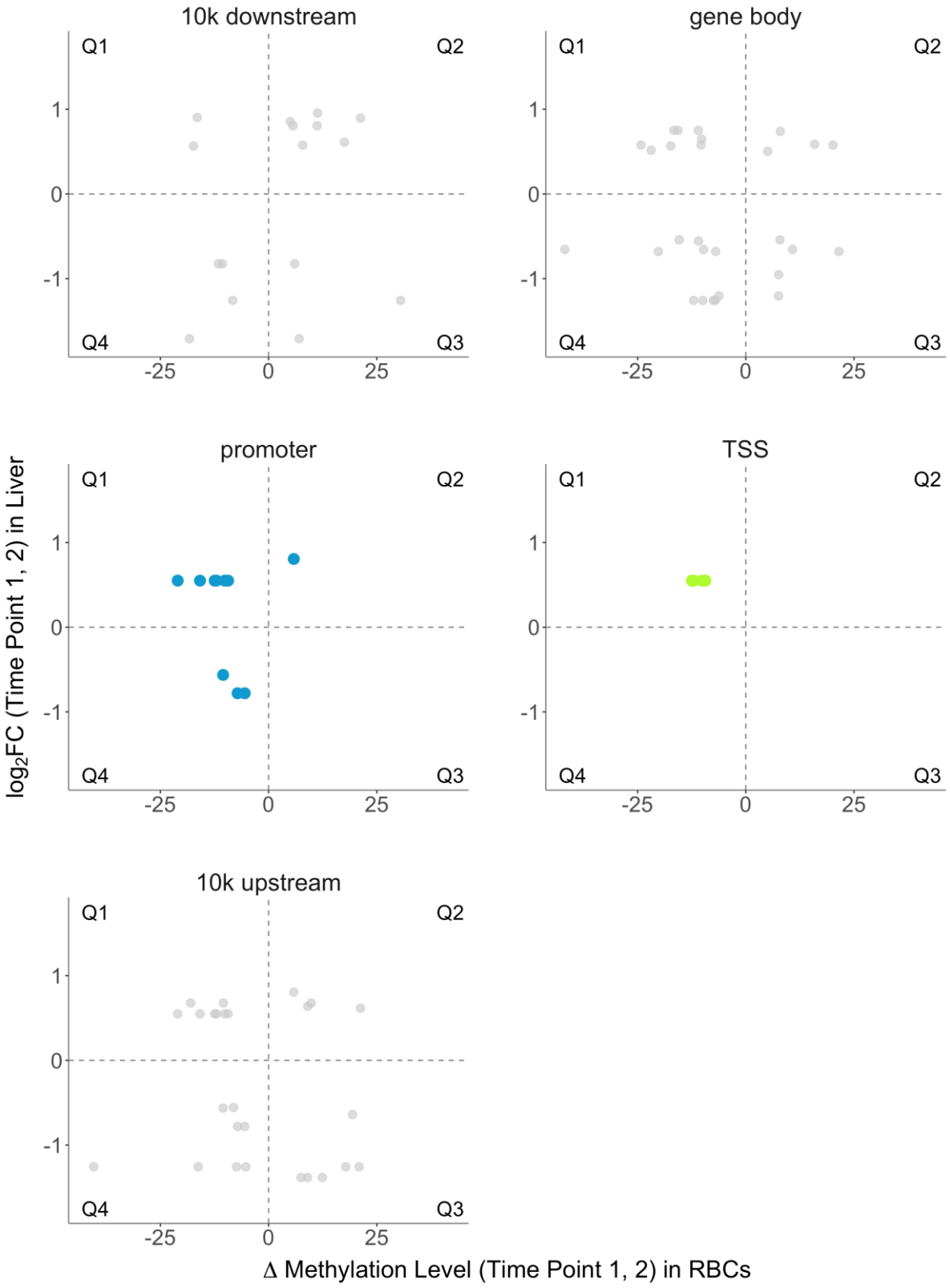


Figure S6.12. Log_2 fold change of genes in *liver* in relation to change in methylation level of a CpG site in *red blood cells* within 10k downstream region, gene body, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $\Delta_{1,2}$. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data point appears darker.

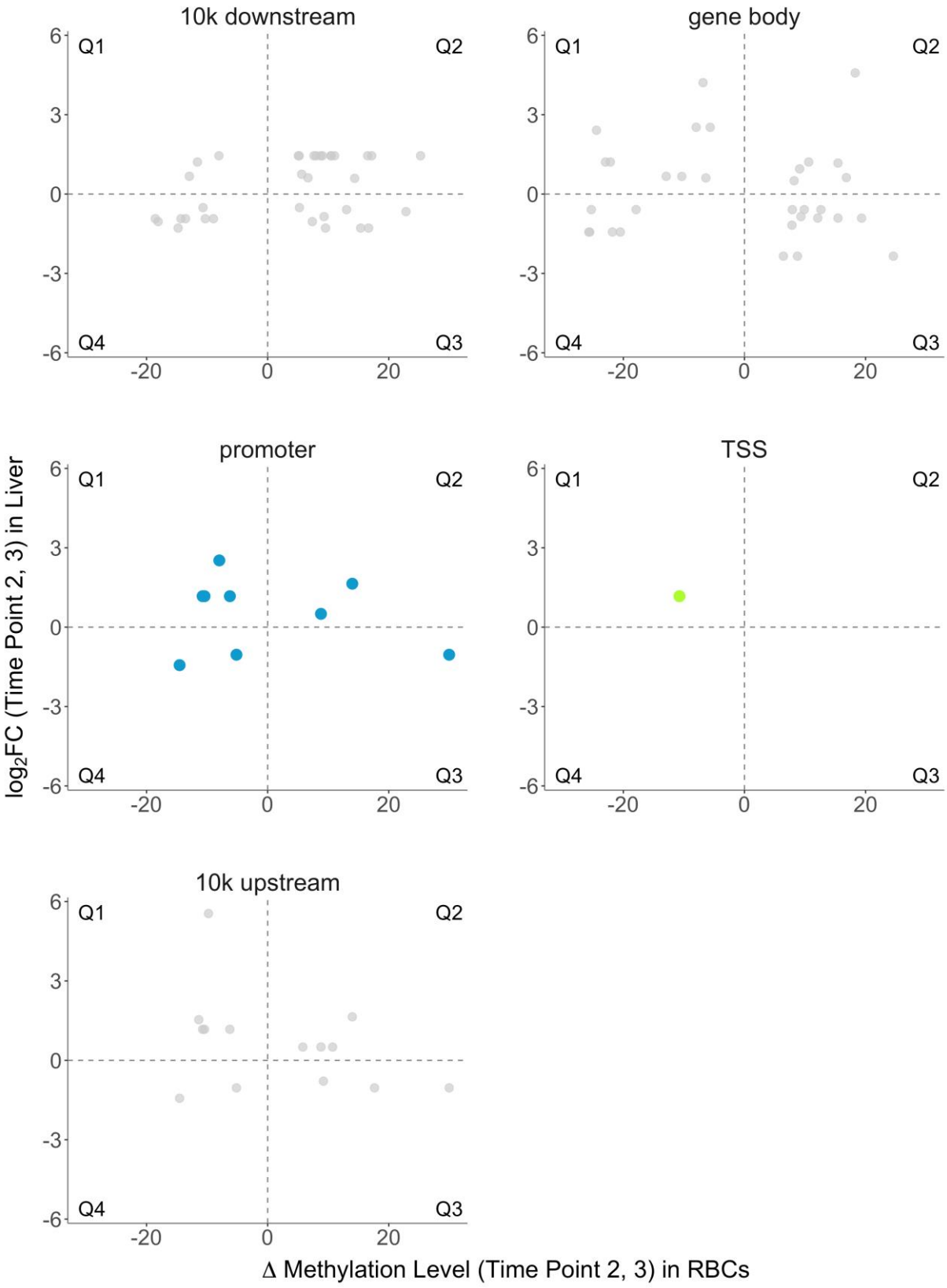


Figure S6.13. Log_2 fold change of genes in *liver* in relation to change in methylation level of a CpG site in *red blood cells* within 10k downstream region, gene body, promoter region, transcription start site (TSS), and 10k upstream region of that gene for $\Delta_{2,3}$. The four quadrants (see '*Materials and methods*') are separated by dotted lines and labeled as 'Q1-Q4'. Transparency is applied to the grey data points such that the area of overlap of between data points appears darker.

Table S6.2. Females differed significantly in mean and variance in methylation across all sites in both RBC and liver, as well as when considering only the promoter regions or TSS.

			MEAN			VARIANCE		
tissue	**genomic region**	**number of sites**	χ^2	**df**	***p***	χ^2	**df**	***p***
RBC	all	6227928	20225	17	<0.0001	63540	17	<0.0001
RBC	promoter	151172	5107.1	17	<0.0001	22446	17	<0.0001
RBC	TSS	55678	3298.7	17	<0.0001	1438	17	<0.0001
liver	all	11387232	29903	17	<0.0001	112490	17	<0.0001
liver	promoter	295315	5320.5	17	<0.0001	1963.2	17	<0.0001
liver	TSS	132013	4106.7	17	<0.0001	1600.7	17	<0.0001

Table S6.5. Expression levels of the candidate genes from **Chapter 5** for which all females show ≥10x coverage for associated CpG sites.

female	**time point**	***B2M***	***HSPB1***	***GR***	***MR***	***PRKCA***	***RPL19***	***SDHA***
BD_27012	two	4.454892	0.021146	0.076496	0.216365	0.305986	2.569587	0.285495
BD_27013	three	10.29401	0.046728	0.00289	0.010862	0.019043	5.658511	0.901514
BD_27027	three	3.918943	0.01873	0.040288	0.107062	0.333674	2.578464	0.296584
BD_27037	two	3.647832	0.018211	0.066338	0.153466	0.434066	2.251657	0.280483
BD_27043	one	4.266094	0.03013	0.045828	0.108997	0.247819	3.154391	0.29986
BD_27046	two	4.71937	0.015255	0.075912	0.138261	0.368687	2.174176	0.26434
BD_27068	one	4.479799	0.020247	0.037262	0.133865	0.181602	3.135847	0.391981
BD_27075	three	2.891987	0.018875	0.06393	0.073183	0.21061	4.206576	0.390298
BD_27082	one	9.46722	0.05902	0.052825	0.161808	0.253025	4.17667	0.09995
BD_27087	three	5.983461	0.012471	0.052362	0.070057	0.27637	2.48746	0.243109
BD_27098	three	5.525009	0.027642	0.050346	0.12657	0.3182	2.975243	0.191181
BD_27113	two	5.441153	0.020814	0.054549	0.104654	0.187335	2.895266	0.338845
BD_27116	one	6.276651	0.023411	0.034996	0.152126	0.407066	1.567358	0.249712
BD_27132	one	7.026777	0.029443	0.024081	0.086812	0.22673	2.127333	0.295053
BD_27154	one	3.055021	0.030883	0.046809	0.161874	0.342207	2.844035	0.336328
BD_27160	two	2.864376	0.039892	0.044726	0.154134	0.405356	2.341219	0.367868
BD_27166	three	4.915855	0.003745	0.031778	0.089883	0.222085	2.601215	0.352131
BD_27194	two	5.990645	0.01641	0.048387	0.132656	0.240776	2.259123	0.306883

Table S6.9. Number of DMS in the set of sites common between liver and RBCs, in promoters and TSS, that showed a differential change between time point 1 and 2, and time point 2 and 3.

	Tissue	$\Delta_{1,2}$	$\Delta_{2,3}$
common DMS	red blood cells	1665	3081
	liver	1249	1706
	both	537	853
DMS in promoters	red blood cells	140	348
	liver	119	186
	both	38	77
DMS in TSS	red blood cells	14	48
	liver	14	24
	both	4	7

Table S6.10. Number of sites within candidate genes.

gene	TSS 1,2	TSS 2,3	Prom 1,2	Prom 2,3
SDHA	n.a.	n.a.	31	32
MR	3	3	51	54
B2M	13	15	19	26
GR	5	4	19	21
PRKCA	6	10	24	25
HSPB1	10	10	16	16
RPL19	n.a.	n.a.	6	6

Table S6.11. Correlations between the change in candidate gene expresssion and change in methylation for both promoter and TSS regions. P-values were corrected (p-adjusted) with the Benjamini-Hochberg procedure.

region $\Delta_{\text{time points}}$	gene	*r*	df	*p*	*p-adjusted*
TSS $\Delta_{1,2}$	*B2M*	-0.41	10	0.191	0.77
	GR	0.04	10	0.905	0.905
	HSPB1	0.32	10	0.308	0.905
	MR	0.05	10	0.876	0.905
	PRKCA	0.04	10	0.898	0.770
TSS $\Delta_{2,3}$	*B2M*	-0.15	10	0.641	0.904
	GR	-0.04	10	0.904	0.904
	HSPB1	-0.04	10	0.894	0.904
	MR	-0.18	10	0.587	0.904
	PRKCA	-0.10	10	0.756	0.904
Promoter $\Delta_{1,2}$	*B2M*	-0.62	10	**0.033**	0.231
	GR	0.00	10	0.990	0.981
	HSPB1	0.23	10	0.469	0.990
	MR	-0.07	10	0.841	0.578
	PRKCA	-0.43	10	0.165	0.821
	SDHA	0.25	10	0.425	0.821
	RPL19	-0.09	10	0.773	0.981
Promoter $\Delta_{2,3}$	*B2M*	0.08	10	0.802	1.000
	GR	-0.04	10	0.892	1.000
	HSPB1	-0.14	10	0.672	1.000
	MR	0.10	10	0.758	1.000
	PRKCA	-0.21	10	0.509	1.000
	SDHA	0.25	10	0.438	1.000
	RPL19	0.07	10	0.819	1.000

CHAPTER 7

General discussion

Globally, natural systems are being affected by regional climate changes, particularly by temperature increases (IPCC 2013). Species can respond to rapidly changing environments through dispersal to more suitable habitats and changes in seasonal timing, or phenology through plasticity (Gienapp et al. 2008; Hoffmann & Sgrò 2011). Ecologists are very well aware about the ultimate reasons for this plasticity in phenology, such as breeding. For many species, the timing of life-history events has major fitness consequences, as the window with optimal conditions, to for example successfully raise offspring, is often short. However, plastic responses in phenology to these changing environments, will be insufficient in the long run and genetic adaptations through micro-evolution are necessary (Hoffmann & Sgrò 2011; Merilä & Hendry 2014; Visser 2008). As such, before it is possible to correctly predict phenological responses of species to environmental change, we need to understand evolutionary processes in the wild. Therefore, we need to (1) identify the proximate drivers of species' phenology, together with how they influence the physiological mechanisms underlying phenology and (2) we need to identify the components of the physiology underlying phenology which can be affected by selection, which ultimately will lead to a better understanding of a species' adaptive potential.

For example, proximate approaches to understand timing of breeding decisions, to date, have focused generally on reproductive physiology (e.g. measurements of circulating hormones), which have resulted in extensive knowledge on the hypothalamic-pituitary-gonadal (HPGL) axis functioning (see **Box 2** in **Chapter 1**), but not specifically on the mechanisms that regulate a females' 'decision' to initiate egg-laying. The recent revolution in genomic techniques, however, provides possibilities to bridge the gap between genetic and phenotypic variation (Bengston et al. 2018; Cheviron et al. 2008; Fidler et al. 2007; Slate et al. 2010).

In this thesis, in which I collaborated with researchers from different fields, we explore the mechanisms underlying avian seasonal timing of breeding. First, great tits were genomically selected for early and late egg-laying, in order to compare females that are early and late in timing of breeding (**Chapter 2**). Using these genetically early and late birds in experiments under controlled conditions allowed us to determine (1) whether temperature has a direct effect on egg-laying (**Chapter 3**), (2) whether we could identify specific genes or gene networks that are involved in timing of breeding (**Chapter 4**) and (3) whether individual differences in timing of breeding in females are reflected in differences in their molecular physiology and if so, where in the HPGL axis these differences reside (**Chapter 5**). Finally, in **Chapter 6**, we set out to explore whether changes in methylation over time are similar, and whether changes in methylation reflect changes in gene expression, within and between tissues. Here, in **Chapter 7**, I will discuss the results found and methods used within this project and relate them to other studies. Also, I will provide some ideas for possible future studies.

Genomic selection on timing of breeding (in a natural population) is possible...

Genomic selection, i.e. on genomically estimated breeding values (GEBV, Gienapp et al. 2019), is in general more accurate compared to selection on the phenotype, and should thus result in faster phenotypic responses across generations (Meuwissen et al. 2016). In **Chapter 2**, we describe a large-scale selection experiment that aimed to select genomically for early and late egg-laying in great tits, and so create an *early* and a *late* selection line. Both genetic and phenotypic responses were found to this artificial selection. GEBVs and fixation indexes (F_{ST}), a measure of genetic differentiation (Holsinger and Weir 2009), diverged between selection lines across generations. Further, we realized an average divergence of ~6 days across generations between the early and late selection line, which is in line with the standard deviation of the within year variation in the Hoge Veluwe population (Gienapp et al. 2006) and with the difference found between early and late F_4-generation selection line females (though, with a total n = 5), which recruited in this population (Ramakers et al. *unpublished manuscript*). In **Chapter 1** we found no significant generation × line-interaction, because quite some within-generation phenotypic variation within the early and late selection line remained, but in the F_3 generation the difference in average egg-laying dates (~10 days) between selection lines did differ significantly. Although previously no association was found between SNPs and egg-laying date (Gienapp et al. 2017), the heritability of egg-laying date is low (h^2 ~ 0.2, Gienapp et al. 2019), the accuracy of GEBVs for egg-laying date (~0.2, Gienapp et al. 2019) is lower compared to what is found in domesticated animal and plant species (>0.6, Hayes et al. 2009; Lin et al. 2014; VanRaden et al. 2009), and the strength of genomic selection applied was moderate (Kingsolver 2001), there were clear genomic and phenotypic responses to bi-directional artificial selection in only three generations, making the selection lines successful and genomic selection a potentially powerful tool within the field of evolutionary ecology.

… but it requires the proper (correlations between) cues

However, the divergence in egg-laying dates between the selection lines was only found in the pairs housed in *outdoor* aviaries (i.e. the pairs that had the most extreme GEBVs and produced the next generation, **Chapter 2**), but not in the pairs housed in *climate-controlled* aviaries (**Chapter 5**). In addition, average egg-laying dates across generations were not affected by the contrasting temperature environments provided in the climate-controlled aviaries (**Chapter 5**), as shown previously (Visser et al. 2011), but in contrast to Visser et al. (2009). Where Visser et al. (2011) used a constant difference of 4°C between the two temperature treatments, the temperature treatments in Visser et al. (2009) mimicked a very cold and very warm year in The Netherlands. Interestingly, we also used temperature treatments that mimicked a very cold and very warm year. This highlights the importance of environmental cues, and especially temperature, and how complex their influence is on timing of breeding. So what explains that we did find a diverging response to selection between the selection lines in the outdoor aviaries, but not in the climate controlled aviaries? And why is there no effect of temperature on the average egg-laying date across generations

in the climate-controlled aviaries? We touch upon this in the discussion of **Chapter 2**, but there could be several reasons. For example, birds in the climate-controlled aviaries have less extreme GEBVs, but I am convinced that that is not the (only) reason. Because cues are so important in timing of breeding, the reduced environmental variability in the climate-controlled aviaries is a much likelier candidate. Even though photoperiod follows the natural cycle, the light strength is significantly lower compared to natural conditions (**Chapter 5**) and thus the photoperiod in controlled conditions could be perceived as short days, resulting in a disrupted correlation between, for example, photoperiod and temperature (Bentley et al. 1998). Further, cues other than photoperiod and temperature, such as for example social cues (Perfito et al. 2015, Bourret et al. 2015), are lacking in the controlled conditions, in order for females to properly time their breeding (Lambrechts et al. 1999). This is also shown in **Chapter 3**, where females in the climate-controlled aviaries appear less sensitive to the average environment (i.e. plasticity in egg-laying) compared to wild conspecifics, indicating that temperature is very likely not the only driver affecting egg-laying and subsequent plasticity. Nonetheless, controlled conditions have proven to be invaluable in identifying the cues that have an indirect (e.g. through food phenology) or direct effect on timing of breeding (Visser et al. 2009, Schaper et al. 2011), which is challenging, if not impossible, in natural conditions.

Timing of breeding is causally affected by temperature… but how?

We were able to determine that temperature has a direct effect on timing of breeding (**Chapter 3**). Some might now think; "didn't we know this already"? Yes, a previous study, also regarding timing of breeding in great tits housed in climate-controlled aviaries and subjected to contrasting temperature environments, showed this direct effect of temperature (Visser et al. 2009). However, as opposed to the *between*-individual approach used in that study, we used a *within*-individual approach in **Chapter 3**, from which we derived stronger conclusions. We show that the selection line females bred significantly earlier in a warmer environment, both in outdoor aviaries and climate-controlled aviaries, as has also been shown for natural populations (Dunn 2004; Nussey et al. 2007). This finding is important in order to predict the maximum rate of temperature increase organisms can cope with (i.e. by being plastic in their timing) or adapt to and it could further advance our understanding of the mechanisms underlying breeding decisions. However, when looking at their reaction norm plasticity, no differences were found between the early and late selection line females, though this could very well be due to the moderate strength of genomic selection (**Chapter 2**). The elevation did differ between the selection lines. Therefore, climate change will lead to a phenotypic change in the average environment, but likely not in phenotypic sensitivity to the environment, as also shown in the wild long-term study population (Ramakers et al. 2019).

Which information from temperature profiles is used by birds to time their breeding, is poorly understood. Though, there is evidence that great tits use a seasonal increase in temperature, rather than variation in mean temperature and daily temperature, to time their

egg-laying (Schaper et al. 2012). In **Chapter 3**, we also found an increase in temperature in the 11 days prior to egg-laying. In birds, temperature is perceived through thermo-receptors, which belong to the transient receptor potential-ion channel family (Dhaka et al. 2006; McKemy et al. 2002). With many of these being activated at relatively high temperatures, the receptors that are activated within a more 'natural' range could be potential candidates to perceive seasonal changes in temperature (Caro et al. 2013). However, work on avian thermo-receptors in relation to temperature dependent timing of breeding is non-existing and it therefore remains unknown if, or how, birds might use these receptors to detect and integrate temperature information over long periods of time. Thermal information, must then be integrated into the brain and translated into a physiological response that underlies seasonal timing, but *how* this occurs also remains a mystery (Caro et al. 2013). In **Chapter 4**, we show evidence of temperature affecting gene expression in the brain, with several genes involved in circadian rhythms differentially expressed. This could be an indication that temperature influences the clock (Lehmann et al. 2012). Though, we need to keep in mind that this does not necessarily mean that the circadian clock, influenced by temperature or not, affects seasonal timing (**Chapter 2**), because to date it remains unclear whether there is a link between the circadian and circannual clock (Agarwal et al. 2017; Budki et al. 2014; Gwinner 1986; Myung et al. 2015).

Timing of breeding is regulated downstream in the HPGL axis

Though egg-laying dates were not affected by temperature treatment (**Chapter 5**, but note that this was an across-female comparison), they are directly affected by temperature (**Chapter 3**). Thermal information needs to be perceived and transduced into a signal via the brain, because it is very unlikely that the downstream tissues *directly* perceive supplementary cues and respond to it. The hypothalamus has long been seen as the final integration point of cues and to guide top-down control of ovarian function (Dawson 2008; Tsutsui et al. 2012). Interestingly, a number of studies did find variation in timing, which could not be explained by variation in reproductive physiology (e.g. LH concentrations) in the pathway *before* the level of the ovary in the HPGL axis (Partecke et al. 2005; Schaper et al. 2012). A study in Japanese quail (*Coturnix japonica*) showed that the largest follicle in the ovary controls, through its own production of circadian clock gene proteins, the LH surge to induce ovulation (Nakao et al. 2007). This provides a different view on the relationship between the ovary and brain, away from the classic top-down regulation of the ovary, where the ovary and brain are acting more as 'balancing partners instead of the ovary being a passive recipient of instructions from the brain' (Ball 2007). It follows, that there is possibly a more independent role for the downstream parts of the HPGL axis, in which the ovary seemingly changes its 'competence' to respond to circulating gonadotrophins (Ball 2007; Caro et al. 2009; Lambrechts & Visser 1999; Partecke et al. 2005; Schaper et al. 2012; Williams 2012). Indeed, we find that variation in genes associated with reproductive functioning in both ovary and liver, but not hypothalamus, explain variation in egg-laying in **Chapter 5**, as did other recent studies (Bergeon Burns et al. 2014; McGuire et al. 2011; Needham et al. 2019; Perfito et al. 2015). The mechanistic underpinnings, however,

together with mechanistic links (e.g. is follicle selection linked to ovulation?) have yet to be determined (see '*Where to further look in the future?*').

Methylate the way

Viitaniemi et al. (2019) showed that in *early* selection line females from the F_2 generation ~9% of the CpG sites with ≥10x coverage in all samples showed large between-individual time-dependent changes in methylation levels, suggesting that these could potentially be involved in the regulation of seasonal timing of breeding. Finding this large between-individual variation in methylation over time, together with within-gene heterogeneity in methylation, highlights the dynamicity of DNA methylation. Caution is thus required, especially in studies which use a sample from a single time point and/or a few targeted CpG sites (e.g. Derks et al. 2016; Laine et al. 2016; Saino et al. 2017; Verhulst et al. 2016). Another study, also using the selection line females, found a correlation between the changes in DNA methylation levels and a female's reproductive stage (Lindner et al. *in prep*). However, until the link between observed methylation changes and transcriptomic expression within and between tissues is tested, predictions about gene regulation underlying timing of breeding is tentative (Viitaniemi et al. 2019). To understand the transcriptional role of methylation, it is necessary to assess within-individual methylation and expression states over time. In **Chapter 6** we assessed to what extend methylation patterns in red blood cells were reflected with those in liver. We found a strong positive correlation between changes of CpG site methylation in red blood cells and liver throughout the whole genome as well as within the promoter region and transcription start site (TSS). Further, CpG site methylation within the TSS showed a trend towards a negative association with the expression of the respective gene, most pronouncedly in the ovary, while, as a null model, the changes in CpG site methylation located within the gene body and the 10k up- or downstream region did not show specific associations to changes in expression of the respective genes in any of the tissue comparisons. In addition, we assessed whether temporal changes in CpG site methylation associate with changes in candidate-gene (**Chapter 5**) and genome-wide gene expression (**Chapter 4**) within and between tissues. We looked at CpG sites, instead of regions or islands, as methylation changes at the site-level could already result in changes in phenotype (Bentz et al. 2016; Leenen, Muller & Turner 2016; Riyahi et al. 2015). These findings are exciting, despite the *between*-individual approach, because they indicate that red blood cells have the potential to 1) be informative of at least part of the (seasonal) change in methylation patterns in liver and 2) give insight about the temporal expression dynamics within the TSS in other tissues. Linking the genes to the sites that show a change in methylation in either a tissue-specific or tissue-general way could further provide information about possible gene networks and their function and if there are promising genes in relation to timing of breeding. Though future studies are warranted, the findings in **Chapter 6** are insightful, valuable and promising regarding insights into the relationship between temporal changes in methylation across tissues, how such changes relate to expression patterns and ultimately the phenotype. The potential for red blood cells as a surrogate for other tissues are not only important in studies using a

within-individual approach, but especially also for future studies in ecological epigenetics in natural populations.

The endo-phenotypes we did not explore, but should be in future studies

Besides the transcriptome (**Chapters 1**, **4** and **5**) and epigenome (**Chapter 6**) there are also the proteome, metabolome and microbiome. Furthermore, these so-called endo-phenotypes are influenced by the environment, interact with each other and contribute to shaping the phenotype (te Pas et al. 2017). As such, a detailed understanding of complex traits, requires knowledge of all the functional levels (i.e. the (epi)genome and endo-phenotypes).

Proteome

Gene activity can be determined by transcriptional expression (mRNA levels, **Chapter 4** and **5**) and protein levels. Changes in complex traits may be caused by biological information that is not identifiable at the transcriptomic level, such as for example differential protein levels where mRNA levels are similar (Doolittle et al. 1990; Stylianou et al. 2008) and protein function caused by changes in the amino acid sequence (Theuns et al. 2000). Proteomics, i.e. the study of proteins encoded by the genome, offers a platform to analyse many proteins (expression proteomics) and protein networks and signalling pathways (interaction and function proteomics), which are more direct determinants that underlie complex biological functions (Grant & Blackstock 2001; Vogel & Marcotte 2012). Expression proteomics is analogous to differential gene expression, but transcript abundance only partially predicts protein abundance (e.g. Greenbaum et al. 2003; Guo et al. 2008; Picotti et al. 2013; de Sousa Abreu et al. 2009; Wilhelm et al. 2014). The proteome is dynamic, and characterized by cell type and time-dependent expression patterns, as it varies with cellular location, over time and in response to the environment. Therefore, proteomic analyses provide information on the proteins and their networks in several ways. Firstly, through comparisons, such as for example comparing pre-laying and egg-laying stage or high and low egg producing individuals in poultry to screen for markers to improve egg-production (Kuo et al. 2005; Luan et al. 2017). An exciting study in broilers and laying hens in which genome-wide gene expression and proteomic data were compared shed new light on the potential function of adipose tissue in relation to reproduction (Börnelov et al. 2018). In this study reproduction related proteins (e.g. zona pellucida proteins, see **Chapters 2** and **4**, and *APOV1*, *VTG2*, see **Chapters 4** and **5**) were found to be expressed in fat, which suggests more direct crosstalk between the adipose tissue and the reproductive system (see '*Metabolome*' below) and possible metabolic regulation of fertility. Secondly, proteins provide information through seasonal changes, as shown for example in the dormancy-active growth transition in poplar (*Populus nigra*) roots (Trupiano et al. 2013), carbon metabolism in the cambial zone of eucalyptus (*Eucalyptus grandis,* Budzinski et al. 2016), semen quality in bull (*Bos taurus*, Westfalewicz et al. 2019) and follicular fluid quality

in horse (Dutra et al. 2019). Lastly, in response to (changing) environments, e.g. global climate change there is a study in gilt-head sea bream (*Sparus aurata*), which showed that proteomic information, coupled with performance measurements, reflected the fish's health status and their ability to acclimate to new environmental conditions (i.e. ocean warming, Madeira et al. 2017). Also, artificially increased CO_2 in sea-water, mimicking future ocean acidification, resulted in differential protein expression in the barnacle *Balanus amphitrite* (Wong et al. 2011). Acquiring great tit proteomic data 'on top of' the genome-wide data (**Chapter 4**) from females in different reproductive stages, throughout the season in tissues from the HPGL axis would be invaluable for a better understanding of the mechanisms underlying timing of breeding.

Metabolome

The metabolome is the complete set of metabolites (both endogenous and exogenous) within an organism, which is, like other endo-phenotypes, influenced by the environment. Metabolites are numerous (in humans an estimated 500,000 at least) and range from amino and fatty acids to environmental contaminants and toxins. Using metabolome analysis for verifying physiological responses to environmental change and to understand mechanisms underlying complex traits relies on the assumption that metabolites play an important role in biological systems (Monteiro et al 2013) and that changing environments cause changes in the metabolomic pathways involved (de Leonardis et al. 2015; Pascual et al. 2017). The metabolite profile might be directly related to complex traits, because performance traits such as growth rate are related to energy metabolism by which metabolites are generated (te Pas et al. 2017). This could potentially be interesting in the light of seasonal timing of breeding (under changing climates), as egg-production is energetically costly (Harshman & Zera 2007; Nilsson & Raberg 2001; Vezina & Williams 2002; Walsberg 1983) and involves lipoprotein metabolism (Walzem 1996). In Atlantic salmon (*Salmo salar*), increased temperatures result in decreased food intake (via changed endocrine signalling) and impaired feed conversion leading to reduced growth and a strong impact on metabolic state (Kullgren et al. 2013). Layer chicks also showed decreased body weight gain under increased temperatures, which is detrimental for poultry health (Tomonaga et al. 2018). A study in *Pinus radiata* plants found crucial metabolites which can reschedule the metabolic strategy to adapt to high temperature (Escandón et al. 2018). Assessing avian metabolomics under different temperatures and even food regimes could therefore provide novel insights into biochemical pathways underlying timing of breeding (in response to climate change).

Microbiome

High-quality individuals (e.g. in terms of body mass in response to food availability) on average express higher fitness-related trait values (e.g. earlier breeding dates and larger clutch sizes) (Bêty et al. 2003; Devries et al. 2008). But again *how*? How is food intake translated at the mechanistic level (e.g. digestion, processing, and nutrient uptake) and how

are these processes integrated in the mechanistic underpinnings underlying (earlier) egg-laying? In organisms, the microbiome or the genetic material within the entire collection of microorganisms in for example the gut, contributes to critical functions, and is likely to have major implications for species evolution, behaviour and physiology (Fitzpatrick et al. 2018; Hanning & Diaz-Sanchez 2015; McKenney et al. 2018; Waite & Taylor 2014). For example, the role of microbiomes is increasingly recognized in regulating hormone and steroid production and, as a result, alter hormone-associated host gene expression profiles (Cryan & Dinan 2012; Evans, Morris & Marchesi 2013; Mayer 2011). Differences and variation in faecal bacterial communities between breeding and non-breeding black rhinos (*Diceros bicornis michaeli*), for example, are found to be associated with hormone production and breeding success and represent a potential biomarker for reproductive health (Antwis et al. 2019). In birds, there is evidence that physiological differences associated with sex and mating system affect gut microbiota (Grond et al. 2018) and that antibiotics depress nestling growth in house sparrows (*Passer domesticus*) (Kohl et al. 2018). Also, perceived external stressors seem to result in changes in the gut microbiome (Bailey et al. 2011; Meddings & Swain 2000; Palme et al. 2005; Sandrini et al. 2015). Though the number of studies on the diversity in avian gut microbiota has increased, studies on the functional aspects remain very limited (Waite & Taylor 2014). However, the existing bi-directional communication between microbiota and their host would be interesting in terms of individual quality or fitness and could potentially influence avian timing of breeding (in changing environments).

Integration of 'omics' data

Our current understanding of molecular mechanisms underlying complex traits remains incomplete. To study these, however, multiple, if not all endo-phenotypes deserve attention (Budzinski et al. 2016; te Pas et al. 2017; Pascual et al. 2017; Stylianou et al. 2008). By for example in integrative studies (Bornelöv et al. 2018; Qin et al. 2019; Shi et al. 2017; Wu et al. 2018), for which tools are currently being developed (Zhang et al. 2019). Not only, because every endo-phenotype harbours different biological information, but also because they interact. Gaining knowledge on physiological pathways and interaction networks would develop our understanding of complex traits.

Where to further look in the future?

Doing research usually means ending up with new, and sometimes even more, questions than one started with. In the light of the mechanisms underlying seasonal timing of breeding there is still much to discover and there are still many questions to answer and directions to look into. Many internal and external factors are influencing timing of breeding, whether it is over time, at different endo-phenotypic levels or in different tissues. As such, this paragraph could become endless and therefore I will restrict possible future directions to the results found in this thesis and refer to Williams (2012) and references therein for a more in depth discussion.

Temperature profiles, heat shock proteins and thermo-receptors

I show that temperature directly affects timing of egg-laying (**Chapter 3**). However, as discussed above, it is still unknown which information from the thermal profiles is used and how. Several studies have investigated which temperature window correlates with or predicts (variation in) timing of breeding, which could differ between species (e.g. Visser et al. 2006; Visser et al. 2009; Williams et al. 2015) or populations of a single species (Drake & Martin 2018). Further, there is evidence that the seasonal increase in temperature is important, opposed to the mean temperature and daily variation in temperature (Schaper et al. 2012, **Chapter 3**), but more experiments involving temperature profiles are needed, using the within-female approach I used in **Chapter 2**.

In the search for the *how,* heat shock proteins and thermo-receptors could be a direction to look in. In **Chapter 4**, for example, we found heat shock proteins *HSPA8, HSP90AA1* and *HSPA4* in hypothalamus, ovary and liver, respectively to be involved in gene modules that significantly correlated with temperature treatment. Further, *HSPA8* and *HSP90AA1* (together with two other genes important in temperature detection: *AKT1* and *MAPK1*) were found in the oestrogen signalling pathway. Heat shock proteins are produced by cells in response to exposure to stressful conditions, e.g. heat shock, but also cold and UV light. Heat shock proteins are 'turned on' by light and temperature. Photoperiod and temperature are two pivotal regulatory factors of breeding time and hence, heat shock proteins could be of interest. For example, the effects of light shock and heat shock, regulated through expression of heat shock protein 101 (*HSP101*), were synergistic in the development of the nodal roots of maize plants, whereas the effect of a thermos-period and photoperiod were additive (Lopez-Frias et al. 2011). A study in male juvenile domestic chickens found down-regulated heat shock protein 90B1 (*HSP90B1*) expression in the hypothalamus after 4 days of photo-stimulation or thyroxine (T4) treatment, which was associated with increased GnRH-I mRNA and plasma LH (Graham et al. 2009). This suggests that *HSP90B1* expression in the hypothalamus is down-regulated in response to the photo-induced increase in T3, generated by the TSHβ-DIO2 photoperiodic pathway in the pars tuberalis and may play a role in neuroendocrine functioning.

As members of the transient receptor potential (TRP) superfamily, thermoreceptors are expressed in sensory nerve endings and skin (Dhaka et al 2006), of which there are four (TRPV1-4) activated by heat (> 25°C) and two, TRPM8 and TRPA1, by cold, i.e. around room temperature (Patapoutian et al. 2003). TRP channels play important functional roles in the signal transduction machinery of hormone-secreting cells and have recently been implicated in reproductive physiology (Götz et al. 2017). While expression studies have demonstrated TRP channel expression at all levels of HPG axis, functional details about TRP channel action at the level of the individual cells controlling reproduction are just beginning to emerge. Identifying how ambient temperature is perceived and integrated at the brain level, would be the first step before the subsequent pathways of chemical and electrical signals leading to seasonal timing can be unravelled.

Many species integrate photoperiod and ambient temperature in order to time their seasonally reoccurring events. For example, the mystery of flowering in *Arabidopsis thaliana* was unravelled by starting with the influence of the circadian system, which resulted in the discovery of proteins linking circadian photoperiodism with temperature influences (Henderson & Dean 2004; Pineiro & Jarillo 2013). From here, Chen & Penfield (2018) found that two genes and an antisense RNA (i.e. a single stranded RNA that is complementary to a protein coding messenger RNA where it blocks its transcription into protein) integrate ambient temperature to control seed dormancy in *Arabidopsis*. As such, looking at other (model) species could potentially provide new insights and directions for future studies.

Downstream mechanisms… and clocks

What would the molecular mechanisms underlying individual 'cue-sensitivity' (Visser et al. 2011) be? So, where could the 'switch' that initiates egg-laying reside within the liver and/or ovary? Could it be in the 'competence' of the ovary to respond to LH (via LH-receptors) and FSH (via FSH-receptors), which would be consistent with a current model for follicle selection (Johnson 2015a)? Could there be a role for GnIH and GnRH receptors at the level of the ovary (Bentley et al. 2006; Maddineni et al. 2008; McGuire, Kangas & Bentley 2011) or IGF-1 receptors (see below, Onagbesan et al. 2009)? Or the communication between ovary and liver, for example in the E_2-dependent shift in lipid metabolism in the liver or the up-regulation of VTG/VLDL-receptors in the ovary? More data on rates of ovarian development and on up- regulation of vitellogenesis in relation to variation in daily photoperiod and temperature would be invaluable (Williams 2012) and (molecular) studies should involve at least all tissues from the HPGL axis.

Further, these potential mechanisms would need to be regulated, and imply the ovary to have a more autonomous role and receiving signals that bypass the classic neuro-endocrine pathway (**Box 2** in **Chapter 1**), maybe via the GnRH receptors in the ovary? Or other autocrine and paracrine regulatory mechanisms (see below)? A study in Japanese quail reported the expression of clock genes in the ovary and circadian patterns of genes involved in progesterone synthesis in pre-ovulatory follicles, suggesting that the ovary regulates her own functioning through its circadian clock (Nakao et al. 2007). In mammals, there is evidence that clocks in peripheral tissues downstream of the central nervous system receive information through multiple pathways, such as via autonomic nervous system activity, endocrine signaling and body temperature (e.g. Brown et al. 2002; Buhr et al. 2010; Guo et al. 2005; Ishida et al. 2005; Le Minh et al. 2001; Terazono et al. 2003), also *independently* of the suprachiasmatic nucleus in the hypothalamus (e.g. Oike et al. 2014; Sasaki et al. 2016; Tahara et al. 2015; Tahara & Shibata 2014).

Insulin-growth factor 1 (IGF-1)

In **Chapter 5**, I found individual differences in timing of breeding to be reflected significantly in mRNA expression of *IGF-1* in liver, with early breeding females showing increased expression for *IGF-1* in liver compared to late breeding females. This is suggestive of a role for IGF-1 in avian reproduction. IGF-1, predominantly synthesized by the liver, plays a clear regulatory role in growth and development, life span and various components of reproduction (Baker et al. 1993; Daftary & Gore 2005; Dantzer & Swanson 2012; Hellström et al. 2016; Holzenberger et al. 2003; Liu et al. 1993; Velazquez et al. 2008). Though, recent studies in avian species have focused on growth and body condition (Lodjak et al. 2014, 2017, 2018; Lodjak & Mägi 2018), evidence for a link between IGF-1 and reproduction in birds exist, but is restricted to poultry and mainly in relation to variation in productivity, such as the number of eggs and egg quality (Hocking et al. 1994; Nagaraja et al. 2000; Wu et al. 2016). The ovary expresses IGF-1 receptors (Onagbesan et al. 2009), but here IGF-1 might be mainly involved in steroidogenesis or differentiation and apoptosis in pre-hierarchical follicles, rather than rapid yolk development or VTG/VLDL receptor function (Onagbesan et al. 2009; Woods et al. 2007). Studies in chicken and rabbit suggests that IGF-1 is also produced by the ovary (downstream!), 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). Further, a few previous studies have proposed a regulatory role for FSH, inhibin, growth factors (e.g. IGF-1) in follicle selection (Palmer & Bahr 1992; Woods & Johnson 2005), though, to date, mechanisms that determine follicle selection remain unknown (Johnson & Woods 2009). As growth and reproduction are closely related, crosstalk between the endocrine systems controlling these fundamental processes is not to be ruled out (Hsin & Kenyon 1999; Hull & Harvey 2014). Interestingly, a recent study found that IGF-1 acts as a cellular circadian 'zeitgeber', by directly affecting the expression of the circadian gene *Bmal-1* and subsequent circadian clock gene expression in hypothalamic cells (Breit et al. 2018). This suggests, that IGF-1 potentially modulates clock-dependent processes in cells. Could there be a role for IGF-1 at the level of the ovary and its clock (see above)?

Concluding remarks

In this thesis, I explored the molecular basis of the physiological mechanism underlying seasonal timing of breeding in an avian model species; the great tit. I aimed to create early and late breeding females using genomic selection on egg-laying date and evaluate the genotypic and phenotypic response to this selection (**Chapter 2**). Here, I conclude that this was successful and that genomic selection is possible in wild populations, providing a powerful and promising tool to study physiological mechanisms underlying complex traits and understanding evolutionary dynamics in natural populations. In **Chapter 3** we aimed to determine whether temperature has a direct effect on timing of breeding, which I can conclude it does. Further, though with some reservation, increasing temperature due to climate change will likely not lead to a change in phenotypic sensitivity to the environment (i.e. plasticity in timing of breeding). I also aimed to determine whether individual differences in timing of breeding in females are reflected in differences in their molecular physiology and if so where. The conclusion here can be short, "yes and downstream in the HPGL axis" (**Chapters 4** and **5**). In addition, short term DNA methylation potentially plays an important role in this (**Chapter 6**). These results are promising for future studies and directions, and the integrative approach used in this thesis should be an example when studying complex traits. However, the vast amount of literature, the multitude of potentially influencing environmental factors, the complexity of the molecular underpinnings in the several endo-phenotypes, their pathways, networks and interactions, the tissues involved (known and unknown), and the many 'unknowns' make timing of breeding a difficult mystery to unravel.

References

Agarwal, N., Mishra, I., Komal, R., Rani, S. & Kumar, V. 2017, 'Circannual testis and moult cycles persist under photoperiods that disrupt circadian activity and clock gene cycles in spotted munia', *The Journal of Experimental Biology*, vol. 220, no. 22, pp. 4162 LP – 4168.

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.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A. & Mason, C.E. 2012, 'methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles.', *Genome biology*, vol. 13, no. 10, p. R87.

Alvarado, S., Mak, T., Liu, S., Storey, K.B. & Szyf, M. 2015, 'Dynamic changes in global and gene-specific DNA methylation during hibernation in adult thirteen-lined ground squirrels, Ictidomys tridecemlineatus.', *The Journal of experimental biology*, vol. 218, no. Pt 11, pp. 1787–95.

Amo, L., Caro, S.P. & Visser, M.E. 2011, 'Sleeping birds do not respond to predator odour', *PLoS ONE*, vol. 6, no. 11.

Andraweera, P.H., Dekker, G.A., Thompson, S.D., North, R.A., McCowan, L.M.E., Roberts, C.T. & Consortium, on behalf of the S. 2012, 'A functional variant in ANGPT1 and the risk of pregnancies with hypertensive disorders and small-for-gestational-age infants', *MHR: Basic science of reproductive medicine*, vol. 18, no. 6, pp. 325–32.

Antwis, R.E., Edwards, K.L., Unwin, B., Walker, S.L. & Shultz, S. 2019, 'Rare gut microbiota associated with breeding success, hormone metabolites and ovarian cycle phase in the critically endangered eastern black rhino', *Microbiome*, vol. 7, no. 1, p. 27.

van Asch, M., Salis, L., Holleman, L.J.M., van Lith, B. & Visser, M.E. 2013, 'Evolutionary response of the egg hatching date of a herbivorous insect under climate change', *Nature Climate Change*, vol. 3, p. 244.

van Asch, M., van Tienderen, P.H., Holleman, L.J.M. & Visser, M.E. 2007, 'Predicting adaptation of phenology in response to climate change, an insect herbivore example', *Global Change Biology*, vol. 13, no. 8, pp. 1596–604.

Asher, G. & Schibler, U. 2011, 'Crosstalk between components of circadian and metabolic cycles in mammals', *Cell Metabolism*, vol. 13, no. 2, pp. 125–37.

Astheimer, L.E.E.B. & Grau, C.R. 1990, 'A comparison of yolk growth rates in seabird eggs', *Ibis*, vol. 132, no. 3, pp. 380–94.

Audet, J.-N., Kayello, L., Ducatez, S., Perillo, S., Cauchard, L., Howard, J.T., O'Connell, L.A., Jarvis, E.D. & Lefebvre, L. 2018, 'Divergence in problem-solving skills is associated with differential expression of glutamate receptors in wild finches', *Science Advances*, vol. 4, no. 3.

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*, vol. 53, no. 3, pp. 1154–60.

Bailey, M.T., Dowd, S.E., Galley, J.D., Hufnagle, A.R., Allen, R.G. & Lyte, M. 2011, 'Exposure to a social stressor alters the structure of the intestinal microbiota: implications for stressor-induced immunomodulation.', *Brain, behavior, and immunity*, vol. 25, no. 3, pp. 397–407.

Baker, J., Liu, J., Robertson, E.J. & Efstratiadis, A. 1993, 'Role of insulin-like growth factors in embryonic and postnatal growth', *Cell*, vol. 75, no. 1, pp. 73–82.

Ball, G.F. 2007, 'The ovary knows more than you think! New views on clock genes and the positive feedback control of luteinizing hormone', *Endocrinology*, vol. 148, no. 7, pp. 3029–30.

Ball, G.F. & Balthazart, J. 2002, 'Neuroendocrine Mechanisms Regulating Reproductive Cycles and Reproductive Behavior in Birds', in D.W. Pfaff, A.P. Arnold, S.E. Fahrbach, A.M. Etgen & R.T.B.T.-H. Rubin Brain and Behavior (eds),*Hormones, Brain, and Behavior*, Academic Press, San Diego, pp. 649–798.

Bates, D., Maechler, M., Bolker, B. & Walker, S. 2015, 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48.

Beebee, T.J.C. 1995, 'Amphibian breeding and climate', *Nature*, vol. 374, no. 6519, pp. 219–20.

Belinky, F., Nativ, N., Stelzer, G., Zimmerman, S., Stein, T.I., Safran, M. & Lancet, D. 2015, 'PathCards: Multi-source consolidation of human biological pathways', *Database*, vol. 2015, no. 2, pp. 1–13.

Ben-Jonathan, N. & Hnasko, R. 2001, 'Dopamine as a prolactin (PRL) inhibitor', *Endocrine Reviews*, vol. 22, no. 6, pp. 724–63.

Bender, J. 2004, 'DNA methylation and epigenetics.', *Annual review of plant biology*, vol. 55, pp. 41–68.

Bengston, S.E., Dahan, R.A., Donaldson, Z., Phelps, S.M., van Oers, K., Sih, A. & Bell, A.M. 2018, 'Genomic tools for behavioural ecologists to understand repeatable individual differences in behaviour', *Nature Ecology & Evolution*, vol. 2, no. 6, pp. 944–55.

Benjamini, Y. & Hochberg, Y. 1995, 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300.

Bentley, G.E., Goldsmith, A.R., Dawson, A., Briggs, C. & Pemberton, M. 1998, 'Decreased Light Intensity Alters the Perception of Day Length by Male European Starlings (Sturnus vulgaris)', *Journal of Biological Rhythms*, vol. 13, no. 2, pp. 148–58.

Bentley, G.E., Kriegsfeld, L.J., Osugi, T., Ukena, K., O'Brien, S., Perfito, N., Moore, I.T., Tsutsui, K. & Wingfield, J.C. 2006, 'Interactions of gonadotropin-releasing hormone (GnRH) and gonadotropin-inhibitory hormone (GnIH) in birds and mammals.', *Journal of experimental zoology. Part A, Comparative experimental biology*, vol. 305, no. 9, pp. 807–14.

Bentz, A.B., Sirman, A.E., Wada, H., Navara, K.J. & Hood, W.R. 2016, 'Relationship between maternal environment and DNA methylation patterns of estrogen receptor alpha in wild Eastern Bluebird (Sialia sialis) nestlings: a pilot study', *Ecology and evolution*, vol. 6, no. 14, pp. 4741–52.

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.

van Bers, N.E.M., Oers, K., Kerstens, H.H.D., Dibbits, B.W., Crooijmans, R.P.M.A., Visser, M.E. & Groenen, M.A.M. 2010, 'Genome-wide SNP detection in the great tit Parus major using high throughput sequencing', *Molecular Ecology*, vol. 19, no. s1, pp. 89–99.

van Bers, N.E.M., Santure, A.W., Van Oers, K., De Cauwer, I., Dibbits, B.W., Mateman, C., Crooijmans, R.P.M.A., Sheldon, B.C., Visser, M.E., Groenen, M.A.M. & Slate, J. 2012, 'The design and cross-population application of a genome-wide SNP chip for the great tit Parus major.', *Molecular ecology resources*, vol. 12, no. 4, pp. 753–70.

Bety, J., Gauthier, G. & Giroux, J.-F. 2003, 'Body condition, migration, and timing of reproduction in snow geese: a test of the condition-dependent model of optimal clutch size.', *The American naturalist*, vol. 162, no. 1, pp. 110–21.

Bilezikjian, L.M., Vaughan, J.M. & Vale, W.W. 1993, 'Characterization and the regulation of inhibin/activin subunit proteins of cultured rat anterior pituitary cells.', *Endocrinology*, vol. 133, no. 6, pp. 2545–53.

Bind, M., Baccarelli, A., Zanobetti, A., Tarantini, L., Suh, H., Vokonas, P. & Schwartz, J. 2012, 'Air pollution and markers of coagulation, inflammation, and endothelial function: associations and epigene-environment interactions in an elderly cohort.', *Epidemiology (Cambridge, Mass.)*, vol. 23, no. 2, pp. 332–40.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z. & Galon, J. 2009, 'ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks', *Bioinformatics*, vol. 25, no. 8, pp. 1091–3.

Bird, A. 2002, 'DNA methylation patterns and epigenetic memory.', *Genes & development*, vol. 16, no. 1, pp. 6–21.

Bird, A. 2007, 'Perceptions of epigenetics', *Nature*, vol. 447, p. 396.

Bornelöv, S., Seroussi, E., Yosefi, S., Benjamini, S., Miyara, S., Ruzal, M., Grabherr, M., Rafati, N., Molin, A., Pendavis, K., Burgess, S.C., Andersson, L. & Friedman-Einat, M. 2018, 'Comparative omics and feeding manipulations in chicken indicate a shift of the endocrine role of visceral fat towards reproduction', *BMC Genomics*, vol. 19, no. 1, p. 295.

Boschiero, C., Moreira, G.C.M., Gheyas, A.A., Godoy, T.F., Gasparin, G., Mariani, P.D.S.C., Paduan, M., Cesar, A.S.M., Ledur, M.C. & Coutinho, L.L. 2018, 'Genome-wide characterization of genetic variants and putative regions under selection in meat and egg-type chicken lines', *BMC genomics*, vol. 19, no. 1, p. 83.

Both, C., Artemyev, A. V., Blaauw, B., Cowie, R.J., Dekhuijzen, A.J., Eeva, T., Enemar, A., Gustafson, L., Ivankina, E. V., Järvinen, A., Metcalfe, N.B., Nyhom, N.E.I., Potti, J., Ravussin, P.-A., Sanz, J.J., Silverin, B., Slater, F.M., Sokolov, L. V., Török, J., Winkel, W., Wright, J., Zang, H. & Visser, M.E. 2004, 'Large–scale geographical variation confirms that climate change causes birds to lay earlier', *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 271, no. 1549, pp. 1657–62.

Both, C., Van Asch, M., Bijlsma, R.G., Van Den Burg, A.B. & Visser, M.E. 2009, 'Climate change and unequal phenological changes across four trophic levels: Constraints or adaptations?', *Journal of Animal Ecology*, vol. 78, no. 1, pp. 73–83.

Both, C., Bouwhuis, S., Lessells, C.M. & Visser, M.E. 2006, 'Climate change and population declines in a long-distance migratory bird', *Nature*, vol. 441, no. 7089, pp. 81–3.

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.

Bourret, A., Belisle, M., Pelletier, F. & Garant, D. 2015, 'Multidimensional environmental influences on timing of breeding in a tree swallow population facing climate change.', *Evolutionary applications*, vol. 8, no. 10, pp. 933–44.

Bouwhuis, S., Sheldon, B.C. & Verhulst, S. 2011, 'Basal metabolic rate and the rate of senescence in the great tit', *Functional Ecology*, vol. 25, no. 4, pp. 829–38.

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.

Breit, A., Miek, L., Schredelseker, J., Geibel, M., Merrow, M. & Gudermann, T. 2018, 'Insulin-like growth factor-1 acts as a zeitgeber on hypothalamic circadian clock gene expression via glycogen synthase kinase-3beta signaling.', *The Journal of biological chemistry*, vol. 293, no. 44, pp. 17278–90.

Brinkhof, M.W.G., Cave, A.J., Hage, F.J. & Verhulst, S. 1993, 'Timing of reproduction and fledging success in the coot Fulica atra: evidence for a causal relationship', *Journal of Animal Ecology*, vol. 62, no. 3, pp. 577–87.

Brommer, J.E., Merilä, J., Sheldon, B.C. & Gustafsson, L. 2005, 'Natural selection and genetic variation for reproductive reaction norms in a wild bird population.', *Evolution; international journal of organic evolution*, vol. 59, no. 6, pp. 1362–71.

Brommer, J.E., Pietiäinen, H. & Kolunen, H. 2003, 'Natural selection on individual clutch size-laying date trends in the Ural owl', *Evolutionary ecology research*, vol. 5, no. 2, pp. 229–37.

Brown, J.L., Li, S.H. & Bhagabati, N. 1999, 'Long-term trend toward earlier breeding in an American bird: A response to global warming?', *Proceedings of the National Academy of Sciences*, vol. 96, no. 10, pp. 5565–9.

Brown, S.A., Zumbrunn, G., Fleury-Olela, F., Preitner, N. & Schibler, U. 2002, 'Rhythms of mammalian body temperature can sustain peripheral circadian clocks.', *Current biology : CB*, vol. 12, no. 18, pp. 1574–83.

Budki, P., Malik, S., Rani, S. & Kumar, V. 2014, 'Circadian rhythms are not involved in the regulation of circannual reproductive cycles in a sub-tropical bird, the spotted munia', *Journal of Experimental Biology*, vol. 217, no. 14, pp. 2569–79.

Budzinski, I.G.F., Moon, D.H., Lindén, P., Moritz, T. & Labate, C.A. 2016, 'Seasonal variation of carbon metabolism in the cambial zone of Eucalyptus grandis', *Frontiers in plant science*, vol. 7, p. 932.

Buhr, E.D. & Takahashi, J.S. 2013, 'Molecular Components of the Mammalian Circadian Clock', *Trends in Cell Biology*, vol. 24, pp. 3–27.

Buhr, E.D., Yoo, S. & Takahashi, J.S. 2010, 'Temperature as a universal resetting cue for mammalian circadian oscillators.', *Science (New York, N.Y.)*, vol. 330, no. 6002, pp. 379–85.

Bünning, E. 1969, 'Common Features of Photoperiodism in Plants and Animals', *Photochemistry and Photobiology*, vol. 9, no. 3, pp. 219–28.

Bushuev, A. V., Husby, A., Sternberg, H. & Grinkov, V.G. 2012, 'Quantitative genetics of basal metabolic rate and body mass in free-living pied flycatchers', *Journal of Zoology*, vol. 288, no. 4, pp. 245–51.

Butler, D., Cullis, B.R., Gilmour, A.R. & Gogel, D.J. 2009, 'ASReml-R Reference Manual, Release 3.0', *Department of Primary Industries and Fisheries*.

Calus, M.P.L. 2010, 'Genomic breeding value prediction: methods and procedures', *Animal*, vol. 4, no. 02, pp. 157–64.

Campbell, D.E.K. & Langlois, V.S. 2018, 'Thyroid hormones and androgens differentially regulate gene expression in testes and ovaries of sexually mature Silurana tropicalis', *General and Comparative Endocrinology*, vol. 267, pp. 172–82.

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., 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., Lambrechts, M.. M., Balthazart, J. & Perret, P. 2007, 'Non-photoperiodic factors and timing of breeding in blue tits: impact of environmental and social influences in semi-natural conditions.', *Behavioural processes*, vol. 75, no. 1, pp. 1–7.

Caro, S.P., Schaper, S. V., Hut, R.A., Ball, G.F. & Visser, M.E. 2013, 'The case of the missing mechanism: how does temperature influence seasonal timing in endotherms?', *PLOS Biology*, vol. 11, no. 4, p. e1001517.

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.

Cassone, V.M. 2015, 'Avian circadian organization', *Mechanisms of Circadian Systems in Animals and Their Clinical Relevance*, vol. 35, no. 1, pp. 69–94.

Cassone, V.M. & Yoshimura, T. 2014, *Circannual Cycles and Photoperiodism*, *Sturkie's Avian Physiology: Sixth Edition*, Sixth Edit., Elsevier.

Cayan, D.R., Kammerdiener, S.A., Dettinger, M.D., Caprio, J.M. & Peterson, D.H. 2001, 'Changes in the onset of spring in the Western United States', *Bulletin of the American Meteorological Society*, vol. 82, no. 3, pp. 399–416.

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.

Charmantier, A., Mccleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B. & Sheldon, B.C. 2008, 'Adaptive phenotypic plasticity in response to climate change in a wild bird population', *Science*, vol. 320, no. May, pp. 800–4.

Chen, M. & Penfield, S. 2018, 'Feedback regulation of COOLAIR expression controls seed dormancy and flowering time', *Science*, vol. 360, no. 6392, pp. 1014 LP – 1017.

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.

Clark, S.A. & van der Werf, J. 2013, *Genomic Best Linear Unbiased Prediction (gBLUP) for the Estimation of Genomic Breeding Values BT - Genome-Wide Association Studies and Genomic Prediction*, in C. Gondro, J. van der Werf & B. Hayes (eds), Humana Press, Totowa, NJ, pp. 321–30.

Cohen, J.M., Lajeunesse, M.J. & Rohr, J.R. 2018, 'A global synthesis of animal phenological responses to climate change', *Nature Climate Change*, vol. 8, no. 3, pp. 224–8.

Conner, J.K. 2003, 'Artificial selection: a powerful tool for ecologists', *Ecology*, vol. 84, no. 7, pp. 1650–60.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. 2009, 'Mapping complex disease traits with global gene expression', *Nature reviews. Genetics*, vol. 10, no. 3, pp. 184–94.

Cortijo, S., Wardenaar, R., Colome-Tatche, M., Gilly, A., Etcheverry, M., Labadie, K., Caillieux, E., Hospital, F., Aury, J., Wincker, P., Roudier, F., Jansen, R.C., Colot, V. & Johannes, F. 2014, 'Mapping the epigenetic basis of complex traits.', *Science (New York, N.Y.)*, vol. 343, no. 6175, pp. 1145–8.

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.

Crick, H.Q.P., Dudley, C., Glue, D.E. & Thomson, D.L. 1997, 'UK birds are laying eggs earlier', *Nature*, vol. 388, p. 526.

Cryan, J.F. & Dinan, T.G. 2012, 'Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour.', *Nature reviews. Neuroscience*, vol. 13, no. 10, pp. 701–12.

Daan, S. & Aschoff, J. 1975, 'Circadian rhythms of locomotor activity in captive birds and mammals: Their variations with season and latitude.', *Oecologia*, vol. 18, no. 4, pp. 269–316.

Daftary, S.S. & Gore, A.C. 2005, 'IGF-1 in the brain as a regulator of reproductive neuroendocrine function.', *Experimental biology and medicine (Maywood, N.J.)*, vol. 230, no. 5, pp. 292–306.

Dalkin, A.C., Haisenleder, D.J., Gilrain, J.T., Aylor, K., Yasin, M. & Marshall, J.C. 1999, 'Gonadotropin-releasing hormone regulation of gonadotropin subunit gene expression in female rats: Actions on follicle-stimulating hormone β messenger ribonucleic acid (mRNA) involve differential expression of pituitary activin (β-B) and follistatin mRNAs', *Endocrinology*, vol. 140, no. 2, pp. 903–8.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R. & Group, 1000 Genomes Project Analysis 2011, 'The variant call format and VCFtools', *Bioinformatics*, vol. 27, no. 15, pp. 2156–8.

Dantzer, B. & Swanson, E.M. 2012, 'Mediation of vertebrate life histories via insulin-like growth factor-1.', *Biological reviews of the Cambridge Philosophical Society*, vol. 87, no. 2, pp. 414–29.

Darwin, C. 1859, *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*, J. Murray (ed.), London.

Davies, M.N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R.A., Milosavljevic, A., Troakes, C., Al-Sarraj, S., Dobson, R., Schalkwyk, L.C. & Mill, J. 2012, 'Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood.', *Genome biology*, vol. 13, no. 6, p. R43.

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.

Dawson, A., King, V.M., Bentley, G.E. & Ball, G.F. 2001, 'Photoperiodic control of seasonality in birds.', *Journal of biological rhythms*, vol. 16, no. 4, pp. 365–80.

DePaolo, L.. 1997, 'Inhibins, Activins, and Follistatins: The Saga Continues', *Experimental Biology and Medicine*, vol. 214, no. 4, pp. 328–39.

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.

Devries, J.H., Brook, R.W., Howerter, D.W. & Anderson, M.G. 2008, 'Effects of Spring Body Condition and Age on Reproduction in Mallards (Anas Platyrhynchos)', *The Auk*, vol. 125, no. 3, pp. 618–28.

Dhaka, A., Viswanath, V. & Patapoutian, A. 2006, 'Trp ion channels and temperature sensation.', *Annual review of neuroscience*, vol. 29, pp. 135–61.

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.

Dmitrijeva, M., Ossowski, S., Serrano, L. & Schaefer, M.H. 2018, 'Tissue-specific DNA methylation loss during ageing and carcinogenesis is linked to chromosome structure, replication timing and cell division rates.', *Nucleic acids research*, vol. 46, no. 14, pp. 7022–39.

Dobzhansky, T. 1973, 'Nothing in Biology Makes Sense except in the Light of Evolution', *The American Biology Teacher*, vol. 35, no. 3, pp. 125–9.

Doolittle, M.H., LeBoeuf, R.C., Warden, C.H., Bee, L.M. & Lusis, A.J. 1990, 'A polymorphism affecting apolipoprotein A-II translational efficiency determines high density lipoprotein size and composition.', *The Journal of biological chemistry*, vol. 265, no. 27, pp. 16380–8.

Drake, A. & Martin, K. 2018, 'Local temperatures predict breeding phenology but do not result in breeding synchrony among a community of resident cavity-nesting birds', *Scientific Reports*, vol. 8, no. 1, p. 2756.

Drent, P.J., van Oers, K. & van Noordwijk, A.J. 2003, 'Realized heritability of personalities in the great tit (Parus major).', *Proceedings. Biological sciences / The Royal Society*, vol. 270, no. 1510, pp. 45–51.

Dufour, S., Weltzien, F.A., Sebert, M.E., Le Belle, N., Vidal, B., Vernier, P. & Pasqualini, C. 2005, 'Dopaminergic inhibition of reproduction in teleost fishes: Ecophysiological and evolutionary implications', *Annals of the New York Academy of Sciences*, vol. 1040, pp. 9–21.

Dunn, P.B.T. 2004, 'Breeding Dates and Reproductive Performance', *Birds and Climate Change*, vol. 35, Academic Press, pp. 69–87.

Dunn, P.O. & Winkler, D.W. 1999, 'Climate change has affected the breeding date of tree swallows throughout North America', *Proceedings. Biological sciences*, vol. 266, no. 1437, pp. 2487–90.

Dutra, G.A., Ishak, G.M., Pechanova, O., Pechan, T., Peterson, D.G., Jacob, J.C.F., Willard, S.T., Ryan, P.L., Gastal, E.L. & Feugang, J.M. 2019, 'Seasonal variation in equine follicular fluid proteome', *Reproductive Biology and Endocrinology*, vol. 17, no. 1, p. 29.

Elis, S., Dupont, J., Couty, I., Persani, L., Govoroun, M., Blesbois, E., Batellier, F. & Monget, P. 2007, 'Expression and biological effects of bone morphogenetic protein-15 in the hen ovary', *Journal of Endocrinology*, vol. 194, no. 3, pp. 485–97.

Escandón, M., Meijón, M., Valledor, L., Pascual, J., Pinto, G. & Cañal, M.J. 2018, 'Metabolome Integrated Analysis of High-Temperature Response in Pinus radiata', *Frontiers in Plant Science*, vol. 9, p. 485.

Evans, J.M., Morris, L.S. & Marchesi, J.R. 2013, 'The gut microbiome: the role of a virtual organ in the endocrinology of the host.', *The Journal of endocrinology*, vol. 218, no. 3, pp. R37-47.

Ewald, E.R., Wand, G.S., Seifuddin, F., Yang, X., Tamashiro, K.L., Potash, J.B., Zandi, P. & Lee, R.S. 2014, 'Alterations in DNA methylation of Fkbp5 as a determinant of blood-brain correlation of glucocorticoid exposure.', *Psychoneuroendocrinology*, vol. 44, pp. 112–22.

Excoffier, L., Hofer, T. & Foll, M. 2009, 'Detecting loci under selection in a hierarchically structured population', *Heredity*, vol. 103, no. 4, pp. 285–98.

Excoffier, L. & Lischer, H.E.L. 2010, 'Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows', *Molecular Ecology Resources*, vol. 10, no. 3, pp. 564–7.

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.

Farner, D.S. 1985, 'Annual Rhythms', *Annual Review of Physiology*, vol. 47, no. 1, pp. 65–82.

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.

Fitzpatrick, C.R., Copeland, J., Wang, P.W., Guttman, D.S., Kotanen, P.M. & Johnson, M.T.J. 2018, 'Assembly and ecological function of the root microbiome across angiosperm plant species', *Proceedings of the National Academy of Sciences*, vol. 115, no. 6, p. E1157 LP-E1165.

Follett, B.K. 1984, 'Birds', in E. Lamming (ed.),*Marschall's Physiology of Reproduction: reproductive cycles of vertebrates*, Churchill Livingstone, New York, NY, pp. 283–350.

Forister, M.L. & Shapiro, A.M. 2003, 'Climatic trends and advancing spring flight of butterflies in lowland California', *Global Change Biology*, vol. 9, no. 7, pp. 1130–5.

Frésard, L., Morisson, M., Brun, J., Collin, A., Pain, B., Minvielle, F. & Pitel, F. 2013, 'Epigenetics and phenotypic variability: some interesting insights from birds', *Genetics Selection Evolution*, vol. 45, no. 1, p. 16.

Garland, T. 2003, 'Selection experiments: an under-utilized tool in biomechanics and organismal biology', *Vertebrate Biomechanics and Evolution*, pp. 23–57.

Gaudet, P. & Dessimoz, C. 2017, *Gene Ontology: Pitfalls, Biases, and Remedies BT - The Gene Ontology Handbook*, in Christophe Dessimoz & N. Škunca (eds), Springer New York, New York, NY, pp. 189–205.

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., Hemerik, L. & Visser, M.E. 2005, 'A new statistical tool to predict phenology under climate change scenarios', *Global Change Biology*, vol. 11, no. 4, pp. 600–6.

Gienapp, P., Laine, V.N., Mateman, A.C., van Oers, K. & Visser, M.E. 2017, 'Environment-dependent genotype-phenotype associations in avian breeding time', *Frontiers in Genetics*, vol. 8, no. AUG, pp. 1–9.

Gienapp, P., Postma, E. & Visser, M.E. 2006, 'Why Breeding Time Has Not Responded To Selection for Earlier Breeding in a Songbird Population', *Evolution*, vol. 60, no. 11, p. 2381.

Gienapp, P., Reed, T.E. & Visser, M.E. 2014, 'Why climate change will invariably alter selection pressures on phenology', *Proceedings of the Royal Society B: Biological Sciences*, vol. 281, no. August, p. 20141611.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A. & Merilä, J. 2008, 'Climate change and evolution: disentangling environmental and genetic responses.', *Molecular ecology*, vol. 17, no. 1, pp. 167–78.

Gienapp, P., Väisänen, R.A. & Brommer, J.E. 2010, 'Latitudinal variation in breeding time reaction norms in a passerine bird', *Journal of Animal Ecology*, vol. 79, no. 4, pp. 836–42.

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.

Götz, V., Qiao, S., Beck, A. & Boehm, U. 2017, 'Transient receptor potential (TRP) channel function in the reproductive axis.', *Cell calcium*, vol. 67, pp. 138–47.

Graham, G., Sharp, P.J., Li, Q., Wilson, P.W., Talbot, R.T., Downing, A. & Boswell, T. 2009, 'HSP90B1, a thyroid hormone-responsive heat shock protein gene involved in photoperiodic signaling.', *Brain research bulletin*, vol. 79, no. 3–4, pp. 201–7.

Grant, S.G.N. & Blackstock, W.P. 2001, 'Proteomics in neuroscience: from protein to network', *The Journal of Neuroscience*, vol. 21, no. 21, pp. 8315 LP – 8318.

Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. 2003, 'Comparing protein abundance and mRNA expression levels on a genomic scale.', *Genome biology*, vol. 4, no. 9, p. 117.

Gregory, S.J. & Kaiser, U.B. 2004, 'Regulation of gonadotropins by inhibin and activin', *Seminars in Reproductive Medicine*, vol. 22, no. 3, pp. 253–67.

Grond, K., Sandercock, B.K., Jumpponen, A. & Zeglin, L.H. 2018, 'The avian gut microbiota: community, physiology and function in wild birds', *Journal of Avian Biology*, vol. 49, no. 11, p. e01788.

Guo, H., Brewer, J.M., Champhekar, A., Harris, R.B.S. & Bittman, E.L. 2005, 'Differential control of peripheral circadian rhythms by suprachiasmatic-dependent neural signals.', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 8, pp. 3111–6.

Guo, Y., Xiao, P., Lei, S., Deng, F., Xiao, G.G., Liu, Y., Chen, X., Li, L., Wu, S., Chen, Y., Jiang, H., Tan, L., Xie, J., Zhu, X., Liang, S. & Deng, H. 2008, 'How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes.', *Acta biochimica et biophysica Sinica*, vol. 40, no. 5, pp. 426–36.

Gwinner, E. 1975, 'Effects of season and external testosterone on the freerunning circadian activity rhythm of European starlings (Sturnus vulgaris)', *Journal of Comparative Physiology*, vol. 103, pp. 315–28.

Gwinner, E. 1986, *Circannual rhythms: endogenous annual clocks in the organization of seasonal processes*, Springer Berlin Heidelberg, Berlin.

Halekoh, U. & Højsgaard, S. 2014, 'A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models - The R Package pbkrtest', *Journal of Statistical Software*, vol. 59, no. 9, pp. 1–32.

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.

Hanning, I. & Diaz-Sanchez, S. 2015, 'The functionality of the gastrointestinal microbiome in non-human animals', *Microbiome*, vol. 3, p. 51.

Harrington, R., Woiwod, I. & Sparks, T. 1999, 'Climate change and trophic interactions', *Trends in Ecology & Evolution*, vol. 14, no. 4, pp. 146–50.

Harshman, L.G. & Zera, A.. . 2007, 'The cost of reproduction: the devil in the details.', *Trends in ecology & evolution*, vol. 22, no. 2, pp. 80–6.

Hayes, B.J., Bowman, P.J., Chamberlain, A.C., Verbyla, K. & Goddard, M.E. 2009, 'Accuracy of genomic breeding values in multi-breed dairy cattle populations', *Genetics Selection Evolution*, vol. 41, no. 1, p. 51.

Heidaritabar, M., Vereijken, A., Muir, W.M., Meuwissen, T., Cheng, H., Megens, H.-J., Groenen, M.A.M. & Bastiaansen, J.W.M. 2014, 'Systematic differences in the response of genetic variation to pedigree and genome-based selection methods', *Heredity*, vol. 113, no. 6, 2014/07/30., pp. 503–13.

Hellström, A., Ley, D., Hansen-Pupp, I., Hallberg, B., Ramenghi, L.A., Lofqvist, C., Smith, L.E.H. & Hard, A. 2016, 'Role of Insulinlike Growth Factor 1 in Fetal Development and in the Early Postnatal Life of Premature Infants.', *American journal of perinatology*, vol. 33, no. 11, pp. 1067–71.

Helm, B. & Visser, M.E. 2010, 'Heritable circadian period length in a wild bird population', *Proceedings of the Royal Society B: Biological Sciences*, vol. 277, no. 1698, pp. 3335–42.

Henderson, I.R. & Dean, C. 2004, 'Control of Arabidopsis flowering: the chill before the bloom.', *Development (Cambridge, England)*, vol. 131, no. 16, pp. 3829–38.

Hinde, R.A. 1952, *The behaviour of the great tit (Parus major) and some other related species*, Behaviour: Supplement, E.J. Brill.

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.

Hoffmann, A.A. & Sgrò, C.M. 2011, 'Climate change and evolutionary adaptation', *Nature*, vol. 470, p. 479.

Holsinger, K.E. & Weir, B.S. 2009, 'Genetics in geographically structured populations: defining, estimating and interpreting FST', *Nature Reviews Genetics*, vol. 10, p. 639.

Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P.C., Cervera, P. & Le Bouc, Y. 2003, 'IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice.', *Nature*, vol. 421, no. 6919, pp. 182–7.

Hsin, H. & Kenyon, C. 1999, 'Signals from the reproductive system regulate the lifespan of C. elegans.', *Nature*, vol. 399, no. 6734, pp. 362–6.

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.

von Humboldt, A. & Bonpland, A. 1814, *Personal narrative of travels to the equinoctial regions of America during the years 1799-1804*, London, Longman, Hurst, Rees, Orme and Brown.

Hüppop, O. & Hüppop, K. 2003, 'North Atlantic Oscillation and timing of spring migration in birds', *Proceedings. Biological sciences*, vol. 270, no. 1512, pp. 233–40.

Husby, A., Nussey, D.H., Visser, M.E., Wilson, A.J., Sheldon, B.C. & Kruuk, L.E.B. 2010, 'Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (Parus major) populations', *Evolution*, vol. 64, no. 8, pp. 2221–37.

Inouye, D.W., Barr, B., Armitage, K.B. & Inouye, B.D. 2000, 'Climate change is affecting altitudinal migrants and hibernating species', *Proceedings of the National Academy of Sciences*, vol. 97, no. 4, pp. 1630 LP – 1633.

IPCC 2014, 'Climate Change 2014 Synthesis Report Summary Chapter for Policymakers', *IPCC*, p. 31.

Ishida, A., Mutoh, T., Ueyama, T., Bando, H., Masubuchi, S., Nakahara, D., Tsujimoto, G. & Okamura, H. 2005, 'Light activates the adrenal gland: timing of gene expression and glucocorticoid release.', *Cell metabolism*, vol. 2, no. 5, pp. 297–307.

Jaenisch, R. & Bird, A. 2003, 'Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.', *Nature genetics*, vol. 33 Suppl, pp. 245–54.

Jannink, J.-L., Lorenz, A.J. & Iwata, H. 2010, 'Genomic selection in plant breeding: from theory to practice', *Briefings in Functional Genomics*, vol. 9, no. 2, pp. 166–77.

John, S., Sabo, P.J., Thurman, R.E., Sung, M., Biddie, S.C., Johnson, T.A., Hager, G.L. & Stamatoyannopoulos, J.A. 2011, 'Chromatin accessibility pre-determines glucocorticoid receptor binding patterns', *Nature Genetics*, vol. 43, p. 264.

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

Johnson, A.L. 2015b, *Reproduction in the Female*, *Sturkie's Avian Physiology*, Sixth Edit., Elsevier.

Johnson, A.L. & Woods, D.C. 2009, 'Dynamics of avian ovarian follicle development: Cellular mechanisms of granulosa cell differentiation', *General and Comparative Endocrinology*, vol. 163, no. 1–2, pp. 12–7.

Johnsson, M., Jonsson, K.B., Andersson, L., Jensen, P. & Wright, D. 2016, 'Quantitative Trait Locus and Genetical Genomics Analysis Identifies Putatively Causal Genes for Fecundity and Brooding in the Chicken', *G3 Genes|Genomes|Genetics*, vol. 6, no. 2, pp. 311–9.

Johnston, R.A., Paxton, K.L., Moore, F.R., Wayne, R.K. & Smith, T.B. 2016, 'Seasonal gene expression in a migratory songbird', *Molecular Ecology*, vol. 25, no. 22, pp. 5680–91.

Juengel, J.L., Bodensteiner, K.J., Heath, D.A., Hudson, N.L., Moeller, C.L., Smith, P., Galloway, S.M., Davis, G.H., Sawyer, H.R. & McNatty, K.P. 2004, 'Physiology of GDF9 and BMP15 signalling molecules', *Animal Reproduction Science*, vol. 82–83, pp. 447–60.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. 2017, 'KEGG: New perspectives on genomes, pathways, diseases and drugs', *Nucleic Acids Research*, vol. 45, no. D1, pp. D353–61.

Kanehisa, M. & Goto, S. 2000, 'KEGG: Kyoto Encyclopedia of Genes and Genomes', *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30.

Kang, L., Cui, X., Zhang, Y., Yang, C. & Jiang, Y. 2013, 'Identification of miRNAs associated with sexual maturity in chicken ovary by Illumina small RNA deep sequencing', *BMC Genomics*, vol. 14, no. 1.

Kang, S.W., Madkour, M. & Kuenzel, W.J. 2017, 'Tissue-specific expression of DNA methyltransferases involved in early-life nutritional stress of chicken, Gallus gallus', *Frontiers in genetics*, vol. 8, p. 204.

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.

Kassam, I., Wu, Y., Yang, J., Visscher, P.M. & McRae, A.F. 2019, 'Tissue-specific sex-differences in human gene expression', *Human Molecular Genetics*.

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.

Kelly, M.J. & Qiu, J. 2010, 'Estrogen signaling in hypothalamic circuits controling reproduction', *Brain Research*, vol. 1364, no. 9, pp. 44–52.

Kettlewell, H.B.D. 1958, 'A survey of the frequencies of Biston betularia (L.) (Lep.) and its melanic forms in Great Britain', *Heredity*, vol. 12, p. 51.

Kim, D., Langmead, B. & Salzberg, S.L. 2015, 'HISAT: a fast spliced aligner with low memory requirements', *Nature Methods*, vol. 12, no. 4, pp. 357–60.

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.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. 2001, 'The strength of phenotypic selection in natural populations.', *The American naturalist*, vol. 157, no. 3, pp. 245–61.

Klengel, T. & Binder, E.B. 2015, 'Epigenetics of Stress-Related Psychiatric Disorders and Gene x Environment Interactions.', *Neuron*, vol. 86, no. 6, pp. 1343–57.

Koh, G.Y. 2013, 'Orchestral actions of angiopoietin-1 in vascular regeneration', *Trends in Molecular Medicine*, vol. 19, no. 1, pp. 31–9.

Kohl, K.D., Brun, A., Bordenstein, S.R., Caviedes-Vidal, E. & Karasov, W.H. 2018, 'Gut microbes limit growth in house sparrow nestlings (Passer domesticus) but not through limitations in digestive capacity.', *Integrative zoology*, vol. 13, no. 2, pp. 139–51.

Kotrschal, A., Kolm, N. & Penn, D.J. 2016, 'Selection for brain size impairs innate, but not adaptive immune responses', *Proceedings of the Royal Society B: Biological Sciences*, vol. 283, no. 1826.

Kotrschal, A., Lievens, E.J.P., Dahlbom, J., Bundsen, A., Semenova, S., Sundvik, M., Maklakov, A.A., Winberg, S., Panula, P. & Kolm, N. 2013, 'Artificial selection on relative brain size reveals a positive genetic correlation between brain size and proactive personality in the guppy', *Evolution*, vol. 68, no. 4, pp. 1139–49.

Kullgren, A., Jutfelt, F., Fontanillas, R., Sundell, K., Samuelsson, L., Wiklander, K., Kling, P., Koppe, W., Larsson, D.G.J., Bjornsson, B.T. & Jonsson, E. 2013, 'The impact of temperature on the metabolome and endocrine metabolic signals in Atlantic salmon (Salmo salar).', *Comparative biochemistry and physiology. Part A, Molecular & integrative physiology*, vol. 164, no. 1, pp. 44–53.

Kuo, Y.-M., Shiue, Y.-L., Chen, C.-F., Tang, P.-C. & Lee, Y.-P. 2005, 'Proteomic analysis of hypothalamic proteins of high and low egg production strains of chickens.', *Theriogenology*, vol. 64, no. 7, pp. 1490–502.

Lack, D. 1968, *Ecological adaptations for breeding in birds*, Methuen, London.

Laine, V.N., Atema, E., Vlaming, P., Verhagen, I., Mateman, C., Ramakers, J.J.C., van Oers, K., Spoelstra, K. & Visser, M.E. 2019, 'The genomics of circadian timing in a wild bird, the great tit (Parus major)', *Frontiers in Ecology and Evolution*, vol. 7, p. 152.

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.

Laird, P.W. 2003, 'The power and the promise of DNA methylation markers.', *Nature reviews. Cancer*, vol. 3, no. 4, pp. 253–66.

Lambrechts, M.M., Perret, P., Maistre, M. & Blondel, J. 1999, 'Do experiments with captive non–domesticated animals make sense without population field studies? A case study with blue tits' breeding time', *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 266, no. 1426, pp. 1311–5.

Lambrechts, M.M. & Visser, M.E. 1999, 'Proximate aspects of the timing of reproduction', in N.J. Adams & R.H. Slotow (eds),*Proceedings of the 22nd International Ornithol. Congress, Durban*, BirdLife South Africa, Johannesburg, pp. 231–3.

Lande, R. & Arnold, S.J. 1983, 'The measurement of selection on correlated characters', *Evolution*, vol. 37, no. 6, pp. 1210–26.

Langfelder, P. & Horvath, S. 2008, 'WGCNA: an R package for weighted correlation network analysis', *BMC Bioinformatics*, vol. 9, no. 1, p. 559.

Law, J.A. & Jacobsen, S.E. 2010, 'Establishing, maintaining and modifying DNA methylation patterns in plants and animals.', *Nature reviews. Genetics*, vol. 11, no. 3, pp. 204–20.

Leenen, F.A.D., Muller, C.P. & Turner, J.D. 2016, 'DNA methylation: conducting the orchestra from exposure to phenotype?', *Clinical epigenetics*, vol. 8, p. 92.

Lehmann, M., Spoelstra, K., Visser, M.E. & Helm, B. 2012, 'Effects of temperature on circadian clock and chronotype: an experimental study on a passerine bird.', *Chronobiology international*, vol. 29, no. 8, pp. 1062–71.

de Leonardis, A.M., Fragasso, M., Beleggia, R., Ficco, D.B.M., de Vita, P. & Mastrangelo, A.M. 2015, 'Effects of Heat Stress on Metabolite Accumulation and Composition, and Nutritional Properties of Durum Wheat Grain', *International journal of molecular sciences*, vol. 16, no. 12, pp. 30382–404.

Levins, R. 1968, *Evolution in changing environments: some theoretical explorations*, Princeton University Press.

Lewontin, R.C. 1974, *The genetic basis of evolutionary change*, Columbia University Press.

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

Liebl, A.L., Schrey, A.W., Richards, C.L. & Martin, L.B. 2013, 'Patterns of DNA Methylation Throughout a Range Expansion of an Introduced Songbird', *Integrative and Comparative Biology*, vol. 53, no. 2, pp. 351–8.

Lin, Z., Hayes, B.J. & Daetwyler, H.D. 2014, 'Genomic selection in crops, trees and forages: a review', *Crop and Pasture Science*, vol. 65, no. 11, pp. 1177–91.

Lindner, M., Viitaniemi, H.M., Verhagen, I., Laine, V.N., Visser, M.E., van Oers, K. & Husby, A. n.d., 'Epigenetic regulation of the onset of reproduction in a wild songbird', *Manuscript in preparation*.

Litscher, E.S. & Wassarman, P.M. 2015, 'Evolution, structure, and synthesis of vertebrate egg-coat proteins', *Trends in Developmental Biology*, no. 8, pp. 65–76.

Liu, J.P., Baker, J., Perkins, A.S., Robertson, E.J. & Efstratiadis, A. 1993, 'Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r).', *Cell*, vol. 75, no. 1, pp. 59–72.

Liu, L., Wylie, R.C., Andrews, L.G. & Tollefsbol, T.O. 2003, 'Aging, cancer and nutrition: the DNA methylation connection.', *Mechanisms of ageing and development*, vol. 124, no. 10–12, pp. 989–98.

Lodjak, J. & Mägi, M. 2018, 'Insulin-like growth factor 1 relieves the constraints on the growth of young wild passerines', *Ibis*, vol. 160, no. 3, pp. 688–92.

Lodjak, J., Mägi, M., Sild, E. & Mänd, R. 2017, 'Causal link between insulin-like growth factor 1 and growth in nestlings of a wild passerine bird', *Functional Ecology*, vol. 31, no. 1, pp. 184–91.

Lodjak, J., Mägi, M. & Tilgar, V. 2014, 'Insulin-like growth factor 1 and growth rate in nestlings of a wild passerine bird', *Functional Ecology*, vol. 28, no. 1, pp. 159–66.

Lodjak, J., Mänd, R. & Mägi, M. 2018, 'Insulin-like growth factor 1 and life-history evolution of passerine birds', *Functional Ecology*, vol. 32, no. 2, pp. 313–23.

Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., Kolde, R., Koltšina, M., Nilsson, T.K., Vilo, J., Salumets, A. & Tõnisson, N. 2014, 'DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns', *Genome Biology*, vol. 15, no. 4, p. 3248.

Lopez-Frias, G., Martinez, L.M., Ponce, G., Cassab, G.I. & Nieto-Sotelo, J. 2011, 'Role of HSP101 in the stimulation of nodal root development from the coleoptilar node by light and temperature in maize (Zea mays L.) seedlings.', *Journal of experimental botany*, vol. 62, no. 13, pp. 4661–73.

Love, M.I., Huber, W. & Anders, S. 2014, 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, vol. 15, no. 12, pp. 1–21.

Love, O.P., Gilchrist, G.H., Descamps, S., Semeniuk, C.A.D. & Bêty, J. 2010, 'Pre-laying climatic cues can time reproduction to optimally match offspring hatching and ice conditions in an Arctic marine bird', *Oecologia*, vol. 164, no. 1, pp. 277–86.

Lowe, R., Shirley, N., Bleackley, M., Dolan, S. & Shafee, T. 2017, 'Transcriptomics technologies', *PLOS Computational Biology*, vol. 13, no. 5, p. e1005457.

Luan, X., Cao, Z., Xing, Z., Liu, M., Gao, M., Meng, B. & Fan, R. 2017, 'Comparative proteomic analysis of pituitary glands from Huoyan geese between pre-laying and laying periods using an iTRAQ-based approach.', *PloS one*, vol. 12, no. 9, p. e0185253.

Lynch, M. & Walsh, B. 1998, *Genetics and analysis of quantitative traits*, 1st ed., Sinauer Associates, Sunderland, MA.

Lyon, B.E., Chaine, A.S. & Winkler, D.W. 2008, 'A Matter of Timing', *Science*, vol. 321, no. 5892, pp. 1051 LP – 1052.

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

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

Madeira, D., Araújo, J.E., Vitorino, R., Costa, P.M., Capelo, J.L., Vinagre, C. & Diniz, M.S. 2017, 'Molecular plasticity under ocean warming: proteomics and fitness data provides clues for a better understanding of the thermal tolerance in fish', *Frontiers in physiology*, vol. 8, p. 825.

Maegawa, S., Hinkal, G., Kim, H.S., Shen, L., Zhang, L., Zhang, J., Zhang, N., Liang, S., Donehower, L.A. & Issa, J.J. 2010, 'Widespread and tissue specific age-related DNA methylation changes in mice.', *Genome research*, vol. 20, no. 3, pp. 332–40.

Maffucci, J.A. & Gore, A.C. 2009, 'Hypothalamic neural systems controlling the female reproductive life cycle: Gonadotropin-releasing hormone, glutamate, and GABA', *International review of cell and molecular biology*, vol. 274, p. 69.

Mäkinen, H., Viitaniemi, H.M., Visser, M.E., Verhagen, I., van Oers, K. & Husby, A. 2019, 'Temporally replicated DNA methylation patterns in great tit using reduced representation bisulfite sequencing', *Scientific Data*, vol. 6, no. 1, p. 136.

Mantel, N. 1967, 'The Detection of Disease Clustering and a Generalized Regression Approach', *Cancer Research*, vol. 27, no. 2 Part 1, pp. 209 LP – 220.

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

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

te Marvelde, L., Webber, S.L., Meijer, H.A.J. & Visser, M.E. 2012, 'Energy expenditure during egg laying is equal for early and late breeding free-living female great tits', *Oecologia*, vol. 168, no. 3, pp. 631–8.

Matthysen, E., Adriaensen, F. & Dhondt, A.A. 2011, 'Multiple responses to increasing spring temperatures in the breeding cycle of blue and great tits (Cyanistes caeruleus, Parus major)', *Global Change Biology*, vol. 17, no. 1, pp. 1–16.

Mauvoisin, D., Dayon, L., Gachon, F. & Kussmann, M. 2015, 'Proteomics and circadian rhythms: It's all about signaling!', *Proteomics*, vol. 15, no. 2–3, pp. 310–7.

Mayer, E.A. 2011, 'Gut feelings: the emerging biology of gut-brain communication.', *Nature reviews. Neuroscience*, vol. 12, no. 8, pp. 453–66.

Maynard Smith, J. 1989, *Evolutionary genetics*, Oxford University Press.

McCleery, R.H. & Perrins, C.M. 1998, '¨temperature and egg-laying trends', *Nature*, vol. 391, no. 6662, pp. 30–1.

Mccleery, R.H., Pettifor, R.A., Armbruster, P., Meyer, K., Sheldon, B.C. & Perrins, C.M. 2004, 'Components of Variance Underlying Fitness in a Natural Population of the Great Tit Parus major', *The American Naturalist*, vol. 164, no. 3, pp. 62–72.

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

McKay, J.A., Xie, L., Harris, S., Wong, Y.K., Ford, D. & Mathers, J.C. 2011, 'Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice.', *Molecular nutrition & food research*, vol. 55, no. 7, pp. 1026–35.

McKemy, D.D., Neuhausser, W.M. & Julius, D. 2002, 'Identification of a cold receptor reveals a general role for TRP channels in thermosensation.', *Nature*, vol. 416, no. 6876, pp. 52–8.

McKenney, E.A., Koelle, K., Dunn, R.R. & Yoder, A.D. 2018, 'The ecosystem services of animal microbiomes.', *Molecular ecology*, vol. 27, no. 8, pp. 2164–72.

McKinnon, L., Picotin, M., Bolduc, E., Juillet, C. & Bêty, J. 2012, 'Timing of breeding, peak food availability, and effects of mismatch on chick growth in birds nesting in the High Arctic', *Canadian Journal of Zoology*, vol. 90, no. 8, pp. 961–71.

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

Meddings, J.B. & Swain, M.G. 2000, 'Environmental stress-induced gastrointestinal permeability is mediated by endogenous glucocorticoids in the rat.', *Gastroenterology*, vol. 119, no. 4, pp. 1019–28.

Meethal, S. V., Liu, T., Chan, H.W., Ginsburg, E., Wilson, A.C., Gray, D.N., Bowen, R.L., Vonderhaar, B.K. & Atwood, C.S. 2009, 'Identification of a regulatory loop for the synthesis of neurosteroids: a steroidogenic acute regulatory protein-dependent mechanism involving hypothalamic-pituitary-gonadal axis receptors', *Journal of neurochemistry*, vol. 110, no. 3, 2009/05/29., pp. 1014–27.

Meijer, T., Nienaber, U., Langer, U. & Trillmich, F. 1999, 'Temperature and Timing of Egg-Laying of European Starlings', *The Condor: Ornithological Applications*, vol. 101, no. 1, pp. 124–32.

Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S. & Jaenisch, R. 2005, 'Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis', *Nucleic acids research*, vol. 33, no. 18, pp. 5868–77.

Menzel, A. & Dose, V. 2005, 'Analysis of long-term time series of the beginning of flowering by Bayesian function estimation', *Meteorologische Zeitschrift*, vol. 14, no. 3, pp. 429–34.

Merilä, J. & Hendry, A.P. 2014, 'Climate change, adaptation, and phenotypic plasticity: The problem and the evidence', *Evolutionary Applications*, vol. 7, no. 1, pp. 1–14.

Merilä, J. & Hoffmann, A.A. 2016, *Evolutionary Impacts of Climate Change*, Oxford University Press.

Merilä, J., Sheldon, B.C. & Kruuk, L.E.B. 2001, 'Explaining statis: Microevolutionary studies in natural populations', *Genetica*, vol. 112–113, no. 1, pp. 199–222.

Meuwissen, T., Hayes, B. & Goddard, M. 2016, 'Genomic selection: A paradigm shift in animal breeding', *Animal Frontiers*, vol. 6, no. 1, pp. 6–14.

Le Minh, N., Damiola, F., Tronche, F., Schutz, G. & Schibler, U. 2001, 'Glucocorticoid hormones inhibit food-induced phase-shifting of peripheral circadian oscillators.', *The EMBO journal*, vol. 20, no. 24, pp. 7128–36.

Miragaia, R.J., Gomes, T., Chomka, A., Jardine, L., Riedel, A., Hegazy, A.N., Whibley, N., Tucci, A., Chen, X., Lindeman, I., Emerton, G., Krausgruber, T., Shields, J., Haniffa, M., Powrie, F. & Teichmann, S.A. 2019, 'Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation', *Immunity*, vol. 50, no. 2, pp. 493-504.e7.

Møller, A.P., Rubolini, D. & Lehikoinen, E. 2008, 'Populations of migratory bird species that did not show a phenological response to climate change are declining', *Proceedings of the National Academy of Sciences*, vol. 105, no. 42, pp. 16195 LP – 16200.

Monaghan, P. & Nager, R.G. 1997, 'Why don't birds lay more eggs?', *Trends in Ecology & Evolution*, vol. 12, no. 7, pp. 270–4.

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

Myung, J., Hong, S., DeWoskin, D., De Schutter, E., Forger, D.B. & Takumi, T. 2015, 'GABA-mediated repulsive coupling between circadian clock neurons in the SCN encodes seasonal time', *Proceedings of the National Academy of Sciences*, vol. 112, no. 29, pp. E3920–9.

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

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

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

Nakao, N., Ono, H. & Yoshimura, T. 2008, 'Thyroid hormones and seasonal reproductive neuroendocrine interactions', *Reproduction*, vol. 136, no. 1, pp. 1–8.

Nakao, N., Yasuo, S., Nishimura, A., Yamamura, T., Watanabe, T., Anraku, T., Okano, T., Fukada, Y., Sharp, P.J., Ebihara, S. & Yoshimura, T. 2007, 'Circadian clock gene regulation of steroidogenic acute regulatory protein gene expression in preovulatory ovarian follicles', *Endocrinology*, vol. 148, no. 7, pp. 3031–8.

Neal-Perry, G. & Santoro, N.F. 2006, *CHAPTER 50 - Aging in the Hypothalamic–Pituitary–Ovarian Axis*, in J.D.B.T.-K. and N.P. of R. (Third E. Neill (ed.), Academic Press, St Louis, pp. 2729–55.

Needham, K.B., Bergeon Burns, C.M., Graham, J.L., Bauer, C.M., Kittilson, J.D., Ketterson, E.D., Hahn, T.P. & Greives, T.J. 2019, 'Changes in processes downstream of the hypothalamus are associated with seasonal follicle development in a songbird, the dark-eyed junco (Junco hyemalis)', *General and Comparative Endocrinology*, vol. 270, pp. 103–12.

Nielsen, J.T. & Moller, A.P. 2006, 'Effects of food abundance, density and climate change on reproduction in the sparrowhawk Accipiter nisus.', *Oecologia*, vol. 149, no. 3, pp. 505–18.

Nilsson, J. & Raberg, L. 2001, 'The resting metabolic cost of egg laying and nestling feeding in great tits.', *Oecologia*, vol. 128, no. 2, pp. 187–92.

Nilsson, J.Å., Åkesson, M. & Nilsson, J.F. 2009, 'Heritability of resting metabolic rate in a wild population of blue tits', *Journal of Evolutionary Biology*, vol. 22, no. 9, pp. 1867–74.

Nilsson, J.F. & Nilsson, J.Å. 2016, 'Fluctuating selection on basal metabolic rate', *Ecology and Evolution*, vol. 6, no. 4, pp. 1197–202.

van Noordwijk, A.J., McCleery, R.H. & Perrins, C.M. 1995, 'Selection for the Timing of Great Tit Breeding in Relation to Caterpillar Growth and Temperature', *Journal of Animal Ecology*, vol. 64, no. 4, pp. 451–8.

van Noordwijk, A.J. & Muller, C.B. 1994, 'On adaptive plasticity in reproductive traits, illustrated with laydate in the Great tit and colony inception in a bumble bee', in P. Jarman & A. Rossiter (eds),*Animal Societies: individuals, interaction and organization*, Kyoto University Press, pp. 180–94.

Nussey, D.H., Postma, E., Gienapp, P. & Visser, M.E. 2005, 'Selection on Heritable Phenotypic Plasticity in a Wild Bird Population', *Science*, vol. 310, no. 5746, pp. 304–6.

Nussey, D.H., Wilson, A.J. & Brommer, J.E. 2007, 'The evolutionary ecology of individual phenotypic plasticity in wild populations', *Journal of Evolutionary Biology*, vol. 20, no. 3, pp. 831–44.

van Oers, K., Buchanan, K.L., Thomas, T.E. & Drent, P.J. 2011, 'Correlated response to selection of testosterone levels and immunocompetence in lines selected for avian personality', *Animal Behaviour*, vol. 81, no. 5, pp. 1055–61.

van Oers, K., Drent, P.J., de Goede, P. & van Noordwijk, A.J. 2004, 'Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities', *Proceedings. Biological sciences*, vol. 271, no. 1534, pp. 65–73.

Oike, H., Oishi, K. & Kobori, M. 2014, 'Nutrients, Clock Genes, and Chrononutrition', *Current nutrition reports*, vol. 3, no. 3, pp. 204–12.

Okabe, M. 2013, 'The cell biology of mammalian fertilization', *Development*, vol. 140, no. 22, pp. 4471–9.

Onagbesan, O., Bruggeman, V. & Decuypere, E. 2009, 'Intra-ovarian growth factors regulating ovarian function in avian species: a review.', *Animal reproduction science*, vol. 111, no. 2–4, pp. 121–40.

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.

Palme, R., Rettenbacher, S., Touma, C., El-Bahr, S.M. & Mostl, E. 2005, 'Stress hormones in mammals and birds: comparative aspects regarding metabolism, excretion, and noninvasive measurement in fecal samples.', *Annals of the New York Academy of Sciences*, vol. 1040, pp. 162–71.

Palmer, S.S. & Bahr, J.M. 1992, 'Follicle stimulating hormone increases serum oestradiol-17 beta concentrations, number of growing follicles and yolk deposition in aging hens (Gallus gallus domesticus) with decreased egg production.', *British poultry science*, vol. 33, no. 2, pp. 403–14.

Parmesan, C. 2006, 'Ecological and evolutionary responses to recent climate change', *Annual Review of Ecology, Evolution, and Systematics*, Annual Reviews, pp. 637–69.

Parmesan, C. 2007, 'Influences of species, latitudes and methodologies on estimates of phenological response to global warming', *Global Change Biology*, vol. 13, no. 9, pp. 1860–72.

Parmesan, C. & Yohe, G. 2003, 'A globally coherent fingerprint of climate change impacts across natural systems.', *Nature*, vol. 421, no. 6918, pp. 37–42.

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.

te Pas, F.M., Madsen, O., Calus, P.M. & Smits, A.M. 2017, 'The Importance of Endophenotypes to Evaluate the Relationship between Genotype and External Phenotype', *International Journal of Molecular Sciences*.

Pascual, J., Canal, M.J., Escandon, M., Meijon, M., Weckwerth, W. & Valledor, L. 2017, 'Integrated Physiological, Proteomic, and Metabolomic Analysis of Ultra Violet (UV) Stress Responses and Adaptation Mechanisms in Pinus radiata.', *Molecular & cellular proteomics : MCP*, vol. 16, no. 3, pp. 485–501.

Patapoutian, A., Peier, A.M., Story, G.M. & Viswanath, V. 2003, 'ThermoTRP channels and beyond: mechanisms of temperature sensation.', *Nature reviews. Neuroscience*, vol. 4, no. 7, pp. 529–39.

Payne, S.H. 2015, 'The utility of protein and mRNA correlation', *Trends in Biochemical Sciences*, vol. 40, no. 1, pp. 1–3.

Peek, C.B., Ramsey, K.M., Marcheva, B. & Bass, J. 2012, 'Nutrient sensing and the circadian clock', *Trends in Endocrinology & Metabolism*, vol. 23, no. 7, pp. 312–8.

Pegoraro, M., Bafna, A., Davies, N.J., Shuker, D.M. & Tauber, E. 2016, 'DNA methylation changes induced by long and short photoperiods in Nasonia.', *Genome research*, vol. 26, no. 2, pp. 203–10.

Penning, T.M. & Drury, J.E. 2007, 'Human aldo–keto reductases: Function, gene regulation, and single nucleotide polymorphisms', *Archives of Biochemistry and Biophysics*, vol. 464, no. 2, pp. 241–50.

Peñuelas, J. & Filella, I. 2001, 'Responses to a Warming World', *Science*, vol. 294, no. 5543, pp. 793 LP – 795.

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.

Perfito, N., Jeong, S.Y., Silverin, B., Calisi, R.M., Bentley, G.E. & Hau, M. 2012, 'Anticipating spring: wild populations of great tits (Parus major) differ in expression of key genes for photoperiodic time measurement', *PLOS ONE*, vol. 7, no. 4, p. e34997.

Perrins, C.M. 1965, 'Population Fluctuations and Clutch-Size in the Great Tit, Parus major L.', *Journal of Animal Ecology*, vol. 34, no. 3, pp. 601–47.

Perrins, C.M. 1970, 'the Timing of Birds' Breeding Seasons', *Ibis*, vol. 112, no. 2, pp. 242–55.

Perrins, C.M. 1979, *British tits*, Collins.

Perrins, C.M. 1991, 'Tits and their caterpillar food supply', *Ibis*, vol. 133, pp. 49–54.

Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D.S., Schmidt, A., Deutsch, E.W., Rost, H., Sun, Z., Rinner, O., Reiter, L., Shen, Q., Michaelson, J.J., Frei, A., Alberti, S., Kusebauch, U., Wollscheid, B., Moritz, R.L., Beyer, A. & Aebersold, R. 2013, 'A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis.', *Nature*, vol. 494, no. 7436, pp. 266–70.

Pigliucci, M. 2001, *Phenotypic plasticity: beyond nature and nurture*, Johns Hopkins University Press.

Pineiro, M. & Jarillo, J.A. 2013, 'Ubiquitination in the control of photoperiodic flowering.', *Plant science : an international journal of experimental plant biology*, vol. 198, pp. 98–109.

Pittendrigh, C.S. & Daan, S. 1976, 'A functional analysis of circadian pacemakers in nocturnal rodents - I. The stability and lability of spontaneous frequency', *Journal of Comparative Physiology A*, vol. 106, no. 3, pp. 223–52.

Platt, T., Fuentes-Yaco, C. & Frank, K.T. 2003, 'Spring algal bloom and larval fish survival', *Nature*, vol. 423, no. 6938, pp. 398–9.

Pörtner, H.O. & Farrell, A.P. 2008, 'Physiology and Climate Change', *Science*, vol. 322, no. 5902, pp. 690 LP – 692.

Primmer, C.R., Papakostas, S., Leder, E.H., Davis, M.J. & Ragan, M.A. 2013, 'Annotated genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and evolution research.', *Molecular ecology*, vol. 22, no. 12, pp. 3216–41.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J. & Sham, P.C. 2007, 'PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses', *The American Journal of Human Genetics*, vol. 81, no. 3, pp. 559–75.

Qin, H., Niu, T. & Zhao, J. 2019, 'Identifying Multi-Omics Causers and Causal Pathways for Complex Traits', *Frontiers in Genetics*, vol. 10, p. 110.

Ramakers, J.J.C., Gienapp, P. & Visser, M.E. 2019, 'Phenological mismatch drives selection on elevation, but not on slope, of breeding time plasticity in a wild songbird', *Evolution*, vol. 73, no. 2, pp. 175–87.

Reddon, A.R., Chouinard-Thuly, L., Leris, I. & Reader, S.M. 2018, 'Wild and laboratory exposure to cues of predation risk increases relative brain mass in male guppies', *Functional Ecology*, vol. 32, no. 7, pp. 1847–56.

Reed, T.E., Grøtan, V., Jenouvrier, S., Sæther, B. & Visser, M.E. 2013, 'Population growth in a wild bird is buffered against phenological mismatch.', *Science*, vol. 340, no. 6131, pp. 488–91.

Reppert, S.M. & Wever, D.R. 2002, 'Coordination of circadian timing system', *Nature*, vol. 418, no. August, pp. 935–41.

Reznick, D.N. & Ricklefs, R.E. 2009, 'Darwin's bridge between microevolution and macroevolution', *Nature*, vol. 457, p. 837.

Richards, E.J. 2006, 'Inherited epigenetic variation--revisiting soft inheritance.', *Nature reviews. Genetics*, vol. 7, no. 5, pp. 395–401.

Rivals, I., Personnaz, L., Taing, L. & Potier, M. 2007, 'Enrichment or depletion of a GO category within a class of genes: which test?', *Bioinformatics*, vol. 23, no. 4, pp. 401–7.

Riyahi, S., Sanchez-Delgado, M., Calafell, F., Monk, D. & Senar, J.C. 2015, 'Combined epigenetic and intraspecific variation of the DRD4 and SERT genes influence novelty seeking behavior in great tit Parus major.', *Epigenetics*, vol. 10, no. 6, pp. 516–25.

Robinson, M.R., Santure, A.W., Decauwer, I., Sheldon, B.C. & Slate, J. 2013, 'Partitioning of genetic variation across the genome using multimarker methods in a wild bird population.', *Molecular ecology*, vol. 22, no. 15, pp. 3963–80.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A. 2003, 'Fingerprints of global warming on wild animals and plants', *Nature*, vol. 421, no. 6918, pp. 57–60.

Roughgarden, J. 1979, *Theory of population genetics and evolutionary ecology: an introduction*, Macmillan.

Roy, D.B. & Sparks, T.H. 2000, 'Phenology of British butterflies and climate change', *Global Change Biology*, vol. 6, no. 4, pp. 407–16.

Rudolph, L.M., Bentley, G.E., Calandra, R.S., Paredes, A.H., Tesone, M., Wu, T.J. & Micevych, P.E. 2016, 'Peripheral and central mechanisms involved in the hormonal control of male and female reproduction', *Journal of Neuroendocrinology*, vol. 28, no. 7.

Ruf, T. 1999, 'The Lomb-Scargle periodogram in biological rhythm research: analysis of incomplete and unequally spaced time-series', *Biological Rhythm Research*, vol. 30, no. 2, pp. 178–201.

Saino, N., Ambrosini, R., Albetti, B., Caprioli, M., De Giorgio, B., Gatti, E., Liechti, F., Parolini, M., Romano, A., Romano, M., Scandolara, C., Gianfranceschi, L., Bollati, V. & Rubolini, D. 2017, 'Migration phenology and breeding success are predicted by methylation of a photoperiodic gene in the barn swallow', *Scientific reports*, vol. 7, p. 45412.

Salvante, K.G., Walzem, R.L. & Williams, T.D. 2007, 'What comes first, the zebra finch or the egg: temperature-dependent reproductive, physiological and behavioural plasticity in egg-laying zebra finches.', *The Journal of experimental biology*, vol. 210, no. Pt 8, pp. 1325–34.

Sánchez, F. & Smitz, J. 2012, 'Molecular control of oogenesis', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, vol. 1822, no. 12, pp. 1896–912.

Sandrini, S., Aldriwesh, M., Alruways, M. & Freestone, P. 2015, 'Microbial endocrinology: host-bacteria communication within the gut microbiome.', *The Journal of endocrinology*, vol. 225, no. 2, pp. R21-34.

Santure, A.W., De Cauwer, I., Robinson, M.R., Poissant, J., Sheldon, B.C. & Slate, J. 2013, 'Genomic dissection of variation in clutch size and egg mass in a wild great tit (Parus major) population', *Molecular Ecology*, vol. 22, no. 15, pp. 3949–62.

Santure, A.W., Gratten, J., Mossman, J. a, Sheldon, B.C. & Slate, J. 2011, 'Characterisation of the transcriptome of a wild great tit Parus major population by next generation sequencing.', *BMC genomics*, vol. 12, no. 1, p. 283.

Sasaki, H., Hattori, Y., Ikeda, Y., Kamagata, M., Iwami, S., Yasuda, S., Tahara, Y. & Shibata, S. 2016, 'Forced rather than voluntary exercise entrains peripheral clocks via a corticosterone/noradrenaline increase in PER2::LUC mice.', *Scientific reports*, vol. 6, p. 27607.

Schaper, S. V. 2012, 'Seasonal timing in a changing climate: the physiological basis of phenotypic plasticity and its evolutionary potential', University of Groningen.

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.

Schaper, S. V., Rueda, C., Sharp, P.J., Dawson, A. & Visser, M.E. 2011, 'Spring phenology does not affect timing of reproduction in the great tit (Parus major)', *Journal of Experimental Biology*, vol. 214, no. 21, pp. 3664–71.

Schultz, T.F. & Kay, S.A. 2003, 'Circadian clocks in daily and seasonal control of development', *Science*, vol. 301, no. 5631, pp. 326–8.

Sellix, M.T. & Menaker, M. 2010, 'Circadian clocks in the ovary.', *Trends in endocrinology and metabolism*, vol. 21, no. 10, pp. 628–36.

Sepers, B., van den Heuvel, K., Lindner, M., Viitaniemi, H.M., Husby, A. & van Oers, K. 2019, 'Avian ecological epigenetics: pitfalls and promises', *Accepted for publication in Journal of Ornithology*.

Serizawa, M., Kinoshita, M., Rodler, D., Tsukada, A., Ono, H., Yoshimura, T., Kansaku, N. & Sasanami, T. 2011, 'Oocytic expression of zona pellucida protein ZP4 in Japanese quail (Coturnix japonica)', *Animal Science Journal*, vol. 82, no. 2, pp. 227–35.

Sharma, A., Singh, D., Das, S. & Kumar, V. 2018, 'Hypothalamic and liver transcriptome from two crucial life-history stages in a migratory songbird', *Experimental Physiology*, vol. 103, no. 4, pp. 559–69.

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

Sheldon, B.C., Kruuk, L.E.B. & Merilä, J. 2003, 'Natural selection and inheritance of breeding time and clutch size in the collared flycatcher', *Evolution*, vol. 57, pp. 406–20.

Shi, S., Shen, Y., Zhang, S., Zhao, Z., Hou, Z., Zhou, H., Zou, J. & Guo, Y. 2017, 'Combinatory Evaluation of Transcriptome and Metabolome Profiles of Low Temperature-induced Resistant Ascites Syndrome in Broiler Chickens', *Scientific Reports*, vol. 7, no. 1, p. 2389.

Shultz, M.T., Piatt, J.F., Harding, A.M.A., Kettle, A.B. & Van Pelt, T.I. 2009, *Timing of breeding and reproductive performance in murres and kittiwakes reflect mismatched seasonal prey dynamics*, vol. 393, pp. 247–58.

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.

Slate, J., Santure, A.W., Feulner, P.G.D., Brown, E.A., Ball, A.D., Johnston, S.E. & Gratten, J. 2010, 'Genome mapping in intensively studied wild vertebrate populations', *Trends in Genetics*, vol. 26, no. 6, pp. 275–84.

Slieker, R.C., Relton, C.L., Gaunt, T.R., Slagboom, P.E. & Heijmans, B.T. 2018, 'Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception.', *Epigenetics & chromatin*, vol. 11, no. 1, p. 25.

Sliwowska, J.H., Fergani, C., Gawałek, M., Skowronska, B., Fichna, P. & Lehman, M.N. 2014, 'Insulin: Its role in the central control of reproduction', *Physiology & Behavior*, vol. 133, pp. 197–206.

Smith, A.K., Kilaru, V., Klengel, T., Mercer, K.B., Bradley, B., Conneely, K.N., Ressler, K.J. & Binder, E.B. 2015, 'DNA extracted from saliva for methylation studies of psychiatric traits: evidence tissue specificity and relatedness to brain.', *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, vol. 168B, no. 1, pp. 36–44.

de Sousa Abreu, R., Penalva, L.O., Marcotte, E.M. & Vogel, C. 2009, 'Global signatures of protein and mRNA expression levels.', *Molecular bioSystems*, vol. 5, no. 12, pp. 1512–26.

Spoelstra, K., Verhagen, I., Meijer, D. & Visser, M.E. 2018, 'Artificial light at night shifts daily activity patterns but not the internal clock in the great tit (Parus major)', *Proceedings of the Royal Society B: Biological Sciences*, vol. 285, no. 1875, p. 20172751.

Stenseth, N.C., Mysterud, A., Ottersen, G., Hurrell, J.W., Chan, K. & Lima, M. 2002, 'Ecological Effects of Climate Fluctuations', *Science*, vol. 297, no. 5585, pp. 1292 LP – 1296.

Stevenson, I.R. & Bryant, D.M. 2000, 'Climate change and constraints on breeding.', *Nature*, vol. 406, no. 6794, pp. 366–7.

Stevenson, T.J. & Prendergast, B.J. 2013, 'Reversible DNA methylation regulates seasonal photoperiodic time measurement.', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 41, pp. 16651–6.

Stojkovic, K., Wing, S.S. & Cermakian, N. 2014, 'A central role for ubiquitination within a circadian clock protein modification code', *Frontiers in Molecular Neuroscience*, vol. 7, no. August, pp. 1–7.

Stokkan, K.A., Yamazaki, S., Tei, H., Sakaki, Y. & Menaker, M. 2001, 'Entrainment of the circadian clock in the liver by feeding', *Science*, vol. 291, no. 5503, pp. 490–3.

Stylianou, I.M., Affourtit, J.P., Shockley, K.R., Wilpan, R.Y., Abdi, F.A., Bhardwaj, S., Rollins, J., Churchill, G.A. & Paigen, B. 2008, 'Applying gene expression, proteomics and single-nucleotide polymorphism analysis for complex trait gene identification.', *Genetics*, vol. 178, no. 3, pp. 1795–805.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., Kuhn, M., Bork, P., Jensen, L.J. & Von Mering, C. 2015, 'STRING v10: Protein-protein interaction networks, integrated over the tree of life', *Nucleic Acids Research*, vol. 43, no. D1, pp. D447–52.

Tahara, Y. & Shibata, S. 2014, 'Chrono-biology, chrono-pharmacology, and chrono-nutrition.', *Journal of pharmacological sciences*, vol. 124, no. 3, pp. 320–35.

Tahara, Y., Shiraishi, T., Kikuchi, Y., Haraguchi, A., Kuriki, D., Sasaki, H., Motohashi, H., Sakai, T. & Shibata, S. 2015, 'Entrainment of the mouse circadian clock by sub-acute physical and psychological stress', *Scientific Reports*, vol. 5, p. 11417.

Tajima, F. 1991, 'Determination of window size for analyzing DNA sequences', *Journal of Molecular Evolution*, vol. 33, no. 5, pp. 470–3.

Tammen, S.A., Friso, S. & Choi, S. 2013, 'Epigenetics: the link between nature and nurture.', *Molecular aspects of medicine*, vol. 34, no. 4, pp. 753–64.

Taylor, D.H., Chu, E.T., Spektor, R. & Soloway, P.D. 2015, 'Long non-coding RNA regulation of reproduction and development', *Molecular Reproduction and Development*, vol. 82, no. 12, pp. 932–56.

Terazono, H., Mutoh, T., Yamaguchi, S., Kobayashi, M., Akiyama, M., Udo, R., Ohdo, S., Okamura, H. & Shibata, S. 2003, 'Adrenergic regulation of clock gene expression in mouse liver', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 11, 2003/05/16., pp. 6795–800.

Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Helaouet, P., Johns, D.G., Jones, I.D., Leech, D.I., Mackay, E.B., Massimino, D., Atkinson, S., Bacon, P.J., Brereton, T.M., Carvalho, L., Clutton-Brock, T.H., Duck, C., Edwards, M., Elliott, J.M., Hall, S.J.G., Harrington, R., Pearce-Higgins, J.W., Høye, T.T., Kruuk, L.E.B., Pemberton, J.M., Sparks, T.H., Thompson, P.M., White, I., Winfield, I.J. & Wanless, S. 2016, 'Phenological sensitivity to climate across taxa and trophic levels', *Nature*, vol. 535, no. 7611, pp. 241–5.

Thackeray, S.J., Sparks, T.H., Frederiksen, M., Burthe, S., Bacon, P.J., Bell, J.R., Botham, M.S., Brereton, T.M., Bright, P.W., Carvalho, L., Clutton-Brock, T.H., Dawson, A., Edwards, M., Elliott, J.M., Harrington, R., Johns, D., Jones, I.D., Leech, D.I., Roy, D.B., Scott, W.A., Smith, M., Smithers, R.J., Winfield, I.J. & Wanless, S. 2010, 'Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments.', *Global Change Biology*, vol. 16, pp. 3304–13.

Theuns, J., Del-Favero, J., Dermaut, B., van Duijn, C.M., Backhovens, H., Van den Broeck, M. V., Serneels, S., Corsmit, E., Van Broeckhoven, C. V. & Cruts, M. 2000, 'Genetic variability in the regulatory region of presenilin 1 associated with risk for Alzheimer's disease and variable expression.', *Human molecular genetics*, vol. 9, no. 3, pp. 325–31.

Thomas, D.W., Blondel, J., Perret, P., Lambrechts, M.M. & Speakman, J.R. 2001, 'Energetic and Fitness Costs of Mismatching Resource Supply and Demand in Seasonally Breeding Birds', *Science*, vol. 291, no. 5513, pp. 2598 LP – 2600.

Tieleman, B.I., Versteegh, M.A., Fries, A., Helm, B., Dingemanse, N.J., Gibbs, H.L. & Williams, J.B. 2009, 'Genetic modulation of energy metabolism in birds through mitochondrial function', *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1662, pp. 1685–93.

Tischkau, S.A., Howell, R.E., Hickok, J.R., Krager, S.L. & Bahr, J.M. 2011, 'The luteinizing hormone surge regulates circadian clock gene expression in the chicken ovary', *Chronobiology International*, vol. 28, no. 1, pp. 10–20.

Tomonaga, S., Okuyama, H., Tachibana, T. & Makino, R. 2018, 'Effects of high ambient temperature on plasma metabolomic profiles in chicks.', *Animal science journal = Nihon chikusan Gakkaiho*, vol. 89, no. 2, pp. 448–55.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. & Pachter, L. 2012, 'Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks', *Nature Protocols*, vol. 7, no. 3, pp. 562–78.

Trivedi, A.K., Kumar, J., Rani, S. & Kumar, V. 2014, 'Annual life history-dependent gene expression in the hypothalamus and liver of a migratory songbird: Insights into the molecular regulation of seasonal metabolism', *Journal of Biological Rhythms*, vol. 29, no. 5, pp. 332–45.

Trupiano, D., Rocco, M., Renzone, G., Scaloni, A., Montagnoli, A., Terzaghi, M., Di Iorio, A., Chiatante, D. & Scippa, G.S. 2013, 'Poplar woody root proteome during the transition dormancy-active growth', *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, vol. 147, no. 4, pp. 1095–100.

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.

Ubuka, T., Bentley, G.E. & Tsutsui, K. 2013, 'Neuroendocrine regulation of gonadotropin secretion in seasonally breeding birds', *Frontiers in Neuroscience*, vol. 7, no. 7 MAR, pp. 1–17.

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.

VanRaden, P.M., Van Tassell, C.P., Wiggans, G.R., Sonstegard, T.S., Schnabel, R.D., Taylor, J.F. & Schenkel, F.S. 2009, 'Invited review: reliability of genomic predictions for North American Holstein bulls.', *Journal of dairy science*, vol. 92, no. 1, pp. 16–24.

Velazquez, M.A., Spicer, L.J. & Wathes, D.C. 2008, 'The role of endocrine insulin-like growth factor-I (IGF-I) in female bovine reproduction.', *Domestic animal endocrinology*, vol. 35, no. 4, pp. 325–42.

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.

Verhagen, I., Laine, V.N., Mateman, A.C., Pijl, A., de Wit, R., van Lith, B., Kamphuis, W., Viitaniemi, H.M., Williams, T.D., Caro, S.P., Meddle, S.L., Gienapp, P., van Oers, K. & Visser, M.E. 2019, 'Fine-tuning of seasonal timing of breeding is regulated downstream in the underlying neuro-endocrine system in a small songbird', *The Journal of Experimental Biology*, p. jeb.202481.

Verhulst, E.C., Mateman, A.C., Zwier, M. V., Caro, S.P., Verhoeven, K.J.F. & van Oers, K. 2016, 'Evidence from pyrosequencing indicates that natural variation in animal personality is associated with DRD4 DNA methylation.', *Molecular ecology*, vol. 25, no. 8, pp. 1801–11.

Verhulst, S. & Nilsson, J.A. 2008, 'The timing of birds' breeding seasons: a review of experiments that manipulated timing of breeding', *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 363, no. 1490, pp. 399–410.

Verhulst, S. & Tinbergen, J.M. 1991, 'Experimental Evidence for a Causal Relationship between Timing and Success of Reproduction in the Great Tit Parus m. major', *Journal of Animal Ecology*, vol. 60, no. 1, pp. 269–82.

Vezina, F. & Williams, T.D. 2002, 'Metabolic costs of egg production in the European starling (Sturnus vulgaris).', *Physiological and biochemical zoology : PBZ*, vol. 75, no. 4, pp. 377–85.

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.

Viitaniemi, H.M., Verhagen, I., Visser, M.E., Honkela, A., van Oers, K. & Husby, A. 2019, 'Seasonal variation in genome-wide DNA methylation patterns and the onset of seasonal timing of reproduction in great tits', *Genome Biology and Evolution*, vol. 11, no. 3, pp. 970–83.

Visser, M.E. 2008, 'Keeping up with a warming world ; assessing the rate of adaptation to climate change', *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 275, no. March, pp. 649–59.

Visser, M.E. & Both, C. 2005, 'Shifts in phenology due to global climate change: the need for a yardstick', *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1581, pp. 2561–9.

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., Caro, S.P., van Oers, K., Schaper, S. V. & Helm, B. 2010, 'Phenology, seasonal timing and circannual rhythms: towards a unified framework', *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1555, pp. 3113–27.

Visser, M.E. & Gienapp, P.G. 2019, 'Evolutionary and demographic consequences of phenological mismatches', *Nature Ecology & Evolution*.

Visser, M.E. & Holleman, L.J.M. 2001, 'Warmer springs disrupt the synchrony of oak and winter moth phenology', *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 268, no. 1464, pp. 289–94.

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.

Vogel, C. & Marcotte, E.M. 2012, 'Insights into the regulation of protein abundance from proteomic and transcriptomic analyses.', *Nature reviews. Genetics*, vol. 13, no. 4, pp. 227–32.

Voigt, W., Perner, J., Davis, A.J., Eggers, T., Schumacher, J., Bährmann, R., Fabian, B., Heinrich, W., Köhler, G., Lichter, D., Marstaller, R. & Sander, F.W. 2003, 'Trophic levels are differentially sensitive to climate', *Ecology*, vol. 84, pp. 2444–53.

Waite, D.W. & Taylor, M.W. 2014, 'Characterizing the avian gut microbiota: membership, driving influences, and potential function.', *Frontiers in microbiology*, vol. 5, p. 223.

Wallace, A.R. 1858, 'On the tendency of varieties to depart indefinitely from the original type.', *Zoological Journal of the Linnean Society*, vol. 3, pp. 53–62.

Walsberg, G.E. 1983, 'Avian ecological energetics', in D.S. Farner, J.R. King & K.C.B.T. Parkes (eds),*Avian Biology*, Academic Press, San Diego, pp. 161–220.

Walther, G.-R. 2010, 'Community and ecosystem responses to recent climate change', *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1549, pp. 2019–24.

Walther, G., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J., Hoegh-Guldberg, O. & Bairlein, F. 2002, 'Ecological responses to recent climate change', *Nature*, vol. 416, no. 6879, pp. 389–95.

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

Wang, J., Duan, Y., Meng, Q., Gong, R., Guo, C., Zhao, Y. & Zhang, Y. 2018, 'Integrated analysis of DNA methylation profiling and gene expression profiling identifies novel markers in lung cancer in Xuanwei, China.', *PloS one*, vol. 13, no. 10, p. e0203155.

Wang, X., Pipes, L., Trut, L.N., Herbeck, Y., Vladimirova, A. V., Gulevich, R.G., Kharlamova, A. V., Johnson, J.L., Acland, G.M., Kukekova, A. V. & Clark, A.G. 2018, 'Genomic responses to selection for tame/aggressive behaviors in the silver fox (Vulpes vulpes)', *Proceedings of the National Academy of Sciences*, vol. 115, no. 41, pp. 10398 LP – 10403.

Weir, J.B. de V 1949, 'New methods for calculating metabolic rate with special reference to protein metabolism', *The Journal of Physiology*, vol. 109, no. 1–2, pp. 1–9.

Westfalewicz, B., Dietrich, M., Slowinska, M., Judycka, S. & Ciereszko, A. 2019, 'Seasonal changes in the proteome of cryopreserved bull semen supernatant.', *Theriogenology*, vol. 126, pp. 295–302.

Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A.M., Lieberenz, M., Savitski, M.M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J.-H., Bantscheff, M., Gerstmair, A., Faerber, F. & Kuster, B. 2014, 'Mass-spectrometry-based draft of the human proteome', *Nature*, vol. 509, p. 582.

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*, Princeton University Press, Princeton and Oxford.

Williams, T.D., Bourgeon, S., Cornell, A., Ferguson, L., Fowler, M., Fronstin, R.B. & Love, O.P. 2015, 'Mid-winter temperatures, not spring temperatures, predict breeding phenology in the European starling (Sturnus vulgaris)', *Royal Society Open Science*, vol. 2, no. 1.

Wilschut, R.A., Oplaat, C., Snoek, L.B., Kirschner, J. & Verhoeven, K.J.F. 2016, 'Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage', *Molecular Ecology*, vol. 25, no. 8, pp. 1759–68.

Winder, M. & Schindler, D.E. 2004, 'Climate change uncouples trophic interactions in an aquatic ecosystem', *Ecology*, vol. 85, no. 8, pp. 2100–6.

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.

Wingfield, J.C., Hahn, T.P., Levin, R. & Honey, P. 1992, 'Environmental predictability and control of gonadal cycles in birds', *Journal of Experimental Zoology*, vol. 261, no. 2, pp. 214–31.

Wingfield, J.C. & Kenagy, G.J. 1991, 'Natural regulation of reproductive cycles', in M. Schreibman & R.E. Jones (eds),*Vertebrate endocrinology: fundamentals and biomedical implications*, Academic Press, New York, NY, pp. 181–241.

Wolc, A., Zhao, H.H., Arango, J., Settar, P., Fulton, J.E., O'Sullivan, N.P., Preisinger, R., Stricker, C., Habier, D., Fernando, R.L., Garrick, D.J., Lamont, S.J. & Dekkers, J.C.M. 2015, 'Response and inbreeding from a genomic selection experiment in layer chickens', *Genetics Selection Evolution*, vol. 47, no. 1, p. 59.

Woltereck, R. 1909, 'Weitere experimentelle untersuchungen über Artänderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden', *Verhandlungen der Deutschen Zoologischen Gesellschaft*, pp. 110–72.

Wong, K.K.W., Lane, A.C., Leung, P.T.Y. & Thiyagarajan, V. 2011, 'Response of larval barnacle proteome to CO2-driven seawater acidification.', *Comparative biochemistry and physiology. Part D, Genomics & proteomics*, vol. 6, no. 3, pp. 310–21.

Woods, D.C., Haugen, M.J. & Johnson, A.L. 2007, 'Actions of Epidermal Growth Factor Receptor/Mitogen-Activated Protein Kinase and Protein Kinase C Signaling in Granulosa Cells from Gallus gallus Are Dependent upon Stage of Differentiation1', *Biology of Reproduction*, vol. 77, no. 1, pp. 61–70.

Woods, D.C. & Johnson, A.L. 2005, 'Regulation of follicle-stimulating hormone-receptor messenger RNA in hen granulosa cells relative to follicle selection.', *Biology of reproduction*, vol. 72, no. 3, pp. 643–50.

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.

Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L.R., Marioni, R.E., Martin, N.G., Montgomery, G.W., Deary, I.J., Wray, N.R., Visscher, P.M., McRae, A.F. & Yang, J. 2018, 'Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits', *Nature Communications*, vol. 9, no. 1, p. 918.

Xie, F., Deng, F., Wu, L., Mo, X., Zhu, H., Wu, J., Guo, Y., Zeng, K., Wang, M., Zhu, X., Xia, W., Wang, L., He, P., Bing, P., Lu, X., Zhang, Y. & Lei, S. 2018, 'Multiple correlation analyses revealed complex relationship between DNA methylation and mRNA expression in human peripheral blood mononuclear cells.', *Functional & integrative genomics*, vol. 18, no. 1, pp. 1–10.

Xie, S., Yang, X., Wang, D., Zhu, F., Yang, N., Hou, Z. & Ning, Z. 2018, 'Thyroid transcriptome analysis reveals different adaptive responses to cold environmental conditions between two chicken breeds', *PLoS ONE*, vol. 13, no. 1, p. e0191096.

Yoshimura, T. 2010, 'Neuroendocrine mechanism of seasonal reproduction in birds and mammals.', *Animal science journal = Nihon chikusan Gakkaiho*, vol. 81, no. 4, pp. 403–10.

Yoshimura, T. 2013, 'Thyroid hormone and seasonal regulation of reproduction', *Frontiers in Neuroendocrinology*, vol. 34, no. 3, pp. 157–66.

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.

Zhang, B., Zhou, Y., Lin, N., Lowdon, R.F., Hong, C., Nagarajan, R.P., Cheng, J.B., Li, D., Stevens, M., Lee, H.J., Xing, X., Zhou, J., Sundaram, V., Elliott, G., Gu, J., Shi, T., Gascard, P., Sigaroudinia, M., Tlsty, T.D., Kadlecek, T., Weiss, A., O'Geen, H., Farnham, P.J., Maire, C.L., Ligon, K.L., Madden, P.A.F., Tam, A., Moore, R., Hirst, M., Marra, M.A., Zhang, B., Costello, J.F. & Wang, T. 2013, 'Functional DNA methylation differences between tissues, cell types, and across individuals discovered using the M&M algorithm', *Genome Research*, vol. 23, no. 9, pp. 1522–40.

Zhang, F., Chen, W., Zhu, Z., Zhang, Q., Deary, I.J., Wray, N.R., Visscher, P.M., McRae, A.F. & Yang, J. 2019, 'OSCA: a tool for omic-data-based complex trait analysis', *bioRxiv*.

Zhang, W., Guo, Y., Li, J., Huang, L., Kazitsa, E.G. & Wu, H. 2016, 'Transcriptome analysis reveals the genetic basis underlying the seasonal development of keratinized nuptial spines in Leptobrachium boringii', *BMC Genomics*, vol. 17, no. 1, p. 978.

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.

Zhu, T., Zheng, S.C., Paul, D.S., Horvath, S. & Teschendorff, A.E. 2018, 'Cell and tissue type independent age-associated DNA methylation changes are not rare but common', *Aging*, vol. 10, no. 11, pp. 3541–57.

Summary

With climate change being one of the major threats to current biodiversity, it is essential for species to adapt sufficiently in order to survive. Some species adapt their phenology faster, in reaction to increasing temperatures, compared to others, resulting in mismatched timing. For many seasonal breeding avian species in temperate zones, such as the great tit, the reproductive period is short and coincides with warmer temperatures and increased food supplies required for successful rearing of offspring. Therefore, seasonal breeders time their reproductive cycle to the changing seasons in order to maximize reproductive success and offspring survival. With springs getting warmer earlier in the year, it is of importance for great tit females to start laying earlier to be able to raise their offspring in an optimal period (i.e. sufficient food abundance). However, females show large variation in timing of breeding, which lies in the underlying physiology: different cues are used and translated by a cascade of neuro-endocrine processes along the hypothalamic-pituitary-gonadal-liver (HPGL) axis into a laying date. Natural selection could act on this variation between females, but it is still unclear on which of the compartments (brain, ovary, liver) of the HPGL axis cues act and thus where the variation in timing between females arises. It is of importance to understand how the components of the physiological mechanism contribute to genetic variation in timing before one is able to understand how natural selection can act on timing of reproduction.

In this thesis, the main aim was to explore the molecular basis of the physiological mechanisms underlying avian seasonal timing of breeding. A promising way to do this is by comparing (extremely) early and (extremely) late laying females. In **Chapter 2**, I describe a large-scale selection experiment, where we created selection lines for early and late egg-laying using genomic selection. In this chapter we show that genomic selection on a complex trait such as timing of breeding is possible, because we find that the early and late selection line birds differ genomically and that this difference increases over the generations. In addition, we find that F_3 generation birds differ also phenotypically, with a significant average difference in egg-laying dates of ~10 days between selection lines.

By housing pairs of the selection lines in climate-controlled aviaries and in outdoor aviaries for two consecutive years and in contrasting environments (either artificial or semi-natural), I was able to determine that temperature has a direct effect on timing of breeding instead of via food phenology and that females laid on average earlier in the warm environment (**Chapter 3**). Further, because we obtained two laying dates per female, we evaluated whether our selection on laying date also changed the birds' phenotypic plasticity and found early selection line females to initiated egg laying consistently ~9 days earlier compared to late selection line females in outdoor aviaries, but no difference in the degree of plasticity. This suggests that while natural selection may lead to a change in phenotype in the average environment it is unlikely to result in a correlated response on the degree of plasticity in timing of breeding.

I also aimed to determine whether individual differences in timing of breeding in females are reflected in differences in their molecular biology and if so where. In **Chapter 4** we generated comprehensive RNA expression data from a set of three tissues important in the neuro-endocrine cascade (HPGL axis) underlying avian seasonal timing of breeding, from three different time points and from two temperature treatments and two selection lines for breeding time. Time was the strongest driver in this study, but we found an interesting interaction between time and temperature in hypothalamus, with several genes involved in circadian rhythms differentially expressed. Even though the hypothalamus has been considered the final integration point of environmental cues and guide top down hormonal regulation and in this way direct ovarian function to time breeding, we find evidence for downstream regulation of timing of breeding in **Chapter 5**. Differences in key reproductive candidate gene expression between phenotypically early and late laying females were found exclusively in the ovary and liver. This also suggests that adaptation in the HPGL axis to changing environments might be downstream.

The effects of the environment need to be translated into gene transcription (**Chapter 4** and **5**), for which DNA methylation is a likely key regulator. Therefore, in **Chapter 6**, we investigated in great tits whether methylation changes were tissue-specific or tissue-general and whether such methylation changes were associated with expression changes within and between tissues. Overall, we found a positive correlation between changes in DNA methylation in red blood cells and liver, both genome-wide as well as for the sites within the promoter region or transcription start site (TSS) separately. Within the TSS of genes, hyper-methylation over time in red blood cells was highly correlated with a decrease in the expression of the associated gene in the ovary. Tissue-general changes in DNA methylation could potentially be informative for changes in gene expression in inaccessible tissues.

I explored the molecular basis of the physiological mechanism underlying seasonal timing of breeding in an avian model species; the great tit. I looked at the phenotype, investigated candidate gene and genome-wide gene expression. In addition, we looked at DNA methylation (in relation to gene expression). The main conclusions are that (1) genomic selection is possible in wild populations, (2) temperature directly influences timing of breeding and (3) that timing of breeding is regulated downstream in the HPLG axis. However, we are only scratching the surface of this complex trait and further studies (also considering other 'endo-phenotypes' and their interactions, see **Chapter 7**) are necessary in order to make predictions about whether birds in general, and great tits specifically, will adapt to rapidly changing environments.

Samenvatting

Nu klimaatverandering een grote dreiging vormt voor de huidige biodiversiteit, is het voor soorten essentieel zich aan te passen om te kunnen overleven. De ene soort pas zijn fenologie sneller aan, als respons op stijgende temperaturen, in vergelijking met de andere, wat resulteert in een 'mismatch in timing'. Voor veel seizoens-broedende vogels in gematigd klimaat, zoals de koolmees, is de voortplantingsperiode kort, welke samenvalt met warmere temperaturen en verhoogde voedselbeschikbaarheid die nodig is voor het succesvol grootbrengen van hun jongen. Daarom stemmen seizoens-broeders hun voortplantings-cyclus af op de veranderende seizoenen om zo hun voortplantingssucces en het overleven van hun jongen te vergroten. Met lentes die steeds vroeger in het jaar warmer worden, is het zaak dat koolmeesvrouwen ook vroeger gaan leggen, om in de optimale periode (i.e. voldoende voedselaanbod) hun jongen groot te brengen. Echter, vrouwen laten een grote variatie in de timing van de voorplanting zien. Deze variatie komt voort uit de onderliggende fysiologie, omdat, na integratie en vertaling van omgevingsfactoren, het leggen van ei het resultaat is van een cascade van neuro-endocriene processen langs de hypothalamus-hypofyse-gonadale-lever as (HPGL as). Natuurlijke selectie zou op deze variatie kunnen inhaken, maar het is nog steeds onduidelijk op welke van de componenten (hersenen, eierstok, lever) van de HPGL as omgevingsfactoren invloed uitoefenen en dus waar de fenotypische variatie wordt veroorzaakt. Het is daarom van belang om te begrijpen hoe de componenten van het fysiologisch mechanisme bijdragen aan de genetische variatie in de timing van de voortplanting voordat er onderzocht kan worden hoe natuurlijke selectie deze timing beïnvloedt.

In deze thesis is het hoofddoel om de moleculaire basis van de fysiologische mechanismen onderliggend aan de timing van seizoens-voortplanting bij de koolmees te verkennen. Een veelbelovende manier om dat te doen is door (extreem) vroeg en (extreem) laat leggende vrouwen te vergelijken. In **Hoofdstuk 2** beschrijf ik een grootschalig selectie experiment, waarmee we selectielijnen voor vroege en late ei-leg creëerden door middel van genetische selectie. We laten zien dat genetische selectie op de timing van de voortplanting mogelijk is. We vinden namelijk dat vogels van de vroeg en late selectielijn genetisch verschillen en dat dit verschil groter wordt per generatie. Daarnaast vinden we dat de F_3 generatie ook van elkaar verschillen in fenotype, met een gemiddeld significant verschil in legdatum van ~10 dagen tussen de selectielijnen.

Door selectielijn-paren twee opeenvolgende jaren in klimaat-gecontroleerde en buiten volières te huisvesten in respectievelijk kunstmatige en semi-natuurlijke contrasterende omgevingen, kon er vastgesteld worden dat temperatuur een direct effect heeft op timing van de voortplanting in plaats van via voedselfenologie en dat koolmeesvrouwen gemiddeld gezien vroeger leggen in een warme omgeving (**Hoofdstuk 3**). Verder, omdat we twee legdatums per vrouw verkregen, is er gekeken of onze genetische selectie van invloed is op de fenotypische plasticiteit van de koolmezen. In de buiten volières initieerden vroege selectielijnvrouwen het ei leggen ~9 dagen vroeger in vergelijking met late selectielijn-vrouwen. Echter, er was geen verschil in de mate van plasticiteit tussen de selectielijnen.

Dit suggereert dat, terwijl natuurlijke selectie tot veranderingen in het fenotype kan leiden in de gemiddelde omgeving, het onwaarschijnlijk is dat het resulteert in een gecorreleerde response in de mate van plasticiteit in timing van de voortplanting.

Een ander doel was te bepalen of (individuele) verschillen in timing van de voortplanting bij koolmeesvrouwen weerspiegeld werden in hun moleculaire biologie en zo ja, waar in de HPGL as. Aan de hand van RNA-expressie data, gegenereerd uit drie weefsels die belangrijk zijn in de HPGL as, verzameld op drie verschillende tijdspunten van twee selectielijnen en twee temperatuur omgevingen, laten we in **Hoofdstuk 4** zien dat tijdspunt het sterkste effect heeft. Interessant genoeg, vinden we een interactie tussen tijdspunt en temperatuur in de hypothalamus, waarbij een aantal genen naar voren kwamen die betrokken zijn bij circadiaanse ritmes. Ook al wordt de hypothalamus beschouwd als het laatste integratiepunt van omgevingsfactoren, de uitvoerder van 'top down' hormoon regulatie en op deze manier de functie van de eierstok stuurt en daarmee het timen van de voortplanting, vinden we in **Hoofdstuk 5** aanwijzingen voor een meer prominente rol 'downstream' (eierstok en lever) in het sturen van de timing van de voortplanting. Verschillen in de expressie van kandidaat-genen met een belangrijke rol in de voortplanting tussen fenotypisch vroeg en laat leggende vrouwen werden alleen gevonden in de eierstok en lever. Dit suggereert ook dat eventuele aanpassing in de HPGL as in respons of veranderende omgevingen 'downstream' zou kunnen plaatsvinden.

De effecten van de omgeving moeten vertaald worden naar gen transcriptie (**Hoofdstuk 4** en **5**), waarin DNA methylatie waarschijnlijk een sleutelrol speelt. Daarom, onderzoeken we in **Hoofdstuk 6** of veranderingen in methylatie weefselspecifiek of weefsel breed zijn en of dergelijke veranderingen gepaard gaan met veranderingen in genexpressie binnen en tussen weefsels. Over het geheel vonden we een positieve correlatie tussen veranderingen in DNA methylatie in rode bloedcellen en lever, zowel genoom-wijd als voor de CpG sites binnen de promoter en de transcription start site (TSS) afzonderlijk. Binnen de TSS van de genen in de eierstok, was de hyper-methylatie over de tijd in rode bloedcellen sterk gecorreleerd met een afname in expressie. Weefsel brede veranderingen in DNA methylatie zouden potentieel informatief kunnen zijn voor veranderingen in genexpressie in ontoegankelijke weefsels.

Ik onderzocht de moleculaire basis van het fysiologisch mechanisme onderliggend aan seizoenstiming van de voorplanting van een modelsoort in de ecologie: de koolmees. Ik heb gekeken naar het fenotype, onderzocht genoomwijde en individuele genexpressie en veranderingen in DNA methylatie (in relatie tot veranderingen in genexpressie). De belangrijkste conclusies zijn dat (1) genetische selectie mogelijk is in wilde populaties, (2) temperatuur de timing van de voortplanting direct beïnvloedt en (3) dat deze timing 'downstream' in de HPGL as wordt gecontroleerd. Echter, we staan nog maar aan het begin met het ontrafelen van de moleculaire basis en verder onderzoek (ook met betrekking tot andere 'endo-phenotypes' en hun interacties, **Hoofdstuk 7**) is noodzakelijk om voorspellingen te kunnen maken over of vogels in het algemeen en koolmezen in het bijzonder, zich kunnen aanpassen aan snel veranderende omgevingen.

Curriculum Vitae

Irene Charlotte Verhagen, was born on the 16th of October 1985 in Sneek, the Netherlands. After secondary school, she moved to Groningen in 2003, where she tried two Bachelor studies (Archaeology and Biology) at the University of Groningen. She decided to move to Leeuwarden in 2005 to enrol in the HBO Bachelor's programme Animal Management at Van Hall Larenstein, where she obtained her degree in 2009 with a specialisation in wildlife management – ex situ.

Following her bachelor, she decided to start the MSc in Animal Sciences at the Wageningen University. Here, both ecology and physiology were of high interest. Under supervision of Fred de Boer, she went to Welgevonden Private Game Reserve in South Africa for her minor to study nitrogen and phosphorus cycling between soil, vegetation and large herbivores (zebra and wildebeast). For her major, she went to Pittsburgh, USA to investigate postnatal and pubertal testicular development through quantification of the Leydig cell population under different hormone treatments in the rhesus monkey under supervision of Katja Teerds, Suresh Ramaswamy and Tony Plant. In 2012, she obtained her master's degree.

Her professional career started in 2012 as a medical analyst at the Isala fertility centre, in Zwolle, where she made several people very happy. However, 'making babies' was not for life. Because an interest in physiology does not have to rule out an interest in ecology and vice versa, she started a PhD project in 2015 studying 'the mechanisms underlying seasonal timing of reproduction in the great tit' at the Department of Animal Ecology of the Netherlands Institute of Ecology (NIOO-KNAW) under the supervision of Marcel Visser and Veronika Laine. The results are presented in this thesis. In November 2019 she will start as a Specialist Scientific Information within the Research & Education Support department at the Wageningen University & Research Library in Wageningen where she will support researchers with the management of their research data and publications.

List of publications

Peer reviewed

Verhagen, I., Laine, V.N., Mateman, A.C., Pijl, A., de Wit, R., van Lith, B., Kamphuis, W., Viitaniemi, H.M., Williams, T.D., Caro, S.P., Meddle, S.L., Gienapp, P., van Oers, K. & Visser, M.E. (2019) Fine-tuning of seasonal timing of breeding is regulated downstream in the underlying neuro-endocrine system in a small song bird. *Journal of Experimental Biology*, 222: jeb202481.

Laine, V.N., **Verhagen, I.**, Mateman, A.C., Pijl, A., Williams, T.D., Gienapp, P., van Oers, K. & Visser, M.E. (2019) Gene expression patterns underlying timing of reproduction in a song bird. *BMC Genomics*, 20(1), 693.

Mäkinen, H., Viitaniemi, H.M., Visser, M.E., **Verhagen, I.**, van Oers, K. & Husby, A. (2019) Temporally replicated reduced representation bisulfite sequencing data on DNA methylation pattern in great tit. *Scientific Data,* 6(1), 136.

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,* 33(9), 1708-1721.

Laine, V.N., Atema, E., Vlaming, P., **Verhagen, I.**, Mateman, A.C., Ramakers, J.J.C., van Oers, K., Spoelstra, K. & Visser, M.E. (2019) The genomics of rhythms in wild birds: linking genetic variation to circannual and circadian timing in the great tit (*Parus major*). *Frontiers in Ecology and Evolution*, 7, 152.

Bosse, M., Spurgin L.G., Laine V.N., Cole E.F., Firth, J.A., Gienapp, P., Gosler, A.G., McMahon, K., Poissant, J., **Verhagen, I.**, Groenen, M.A.M., van Oers, K., Sheldon, B.C., Visser, M.E. & Slate, J. (2019). Response to Perrier and Charmantier: on the importance of time scales when studying adaptive evolution. *Evolution Letters*, evl3.112.

Viitaniemi, H.M., **Verhagen, I.**, Visser, M.E., Honkela, A., van Oers, K. & Husby, A. (2019) Seasonal variation in genome-wide DNA methylation patterns and the onset of seasonal timing of reproduction in great tits. *Genome Biology and Evolution*, 11(3), 970-983.

Hsu, B., **Verhagen, I.**, ,Gienapp, P., Darras, V.M., Visser, M.E. & Ruuskanen, S. (2019) Between- and within-individual variation of maternal thyroid hormone deposition in wild great tits (*Parus major*). *The American Naturalist*, 194(4), E96-E108

Spoelstra, K., **Verhagen, I.**, Meijer, D. & Visser, M.E. (2018) Artificial light at night shifts daily activity patterns but not the internal clock in the great tit (*Parus major*). *Proceedings of the Royal Society B*, 285: 20172751.

Bosse, M., Spurgin, L.G., Laine, V.N., Cole, E.F., Firth, J.A., Gienapp, P., Gosler, A.G., McMahon, K., Poissant, J., **Verhagen, I.**, Groenen, M.A.M., van Oers, K., Sheldon, B.C., Visser, M.E., Slate, J. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, 358(6361), 365-368.

Zwerts, J., Prins, H., Bomhoff, D., **Verhagen, I**., Swarth, J., & de Boer, F.W.F. (2015). Competition between a Lawn-Forming Cynodon dactylon and a Tufted Grass Species Hyparrhenia hirta on a South-African Dystrophic Savanna. *PLoS One*, 10(10), 1-17.

Verhagen, I., Ramaswamy, S., Teerds, K., Keijer, J., & Plant, T. (2014). Time course and role of luteinizing hormone and follicle-stimulating hormone in the expansion of the Leydig cell population at the time of puberty in the rhesus monkey (*Macaca mulatta*). *Andrology*, 2(6), 1-7.

Submitted/under review

Verhagen, I., Lindner, M., Viitaniemi, H.M., Laine, V.N., Visser, M.E., van Oers, K. & Husby, A. Temporal changes in DNA methylation and RNA expression in the great tit – assessing correlations within and between tissues. *BMC Genomics*

To be submitted/in preparation

Verhagen, I., Tomotoni, B.M., Gienapp, P. & Visser, M.E. Temperature has a causal and plastic effect on timing of breeding in a small songbird.

Lindner, M., Viitaniemi, H.M., **Verhagen, I.**, Visser, M.E., van Oers, K. & Husby, A. Epigenetic regulation of the onset of reproduction in a wild songbird.

Popular

Verhagen, I. (2017) De koolmees (*Parus major*) – model voor ecologisch onderzoek. *Biotechniek,* 56, 27-31.

Acknowledgements

This thesis would not have come about without all the feathered cheeky beady-eyed acrobats (i.e. great tits) we laid claim to in both the wild and the aviaries. You never asked to be a model species for scientific research, let alone being actually used in it. I am content that your sacrifice, which I have not taken lightly, has not been in vain and that some of you are not just stored in -80°C for eternity, but are turned into many (collaborating) papers describing first time ever experiments and showing new insights. Thank you for your resilience, sacrifice, uniqueness (in personality) and beauty. It was an absolute pleasure to work with you. You have made me appreciate the class of birds in general and the species of great tits in particular.

Also, I could not have done this without the support of colleagues, friends and family, for which I am very grateful. Therefore, a personal message to you:

WIAS training and education statement

Completed in fulfilment of the requirements for the Education Certificate of the *Wageningen Institute of Animal Sciences*

Basic training activities (3.0 ECTS)

WIAS Introduction Day (2016)
Course *Philosophy of Science and/or Ethics* (2017)
Course *Essential Skills* (2017)

Disciplinary Competences (15.6 ECTS)
Research proposal PhD (2015)
Course *Life Histories*, WUR (2015)
Course *Life History Theory*, RUG (2015)
Course *Animal Experimentation* (Article 9), KNAW (2015)
Course *Introduction to R for statistical analysis*, PE&RC, SENSE, WUR (2015)
Course *Survival Analysis*, PE&RC, WUR (2016)
Course *Linear models*, PE&RC, SENSE, WUR (2017)

Professional Competences (5.8 ECTS)
Workshop *How to grow a PhD - PhD event*, KNAW (2015)
PE&RC first years PhD weekend (2015)
PhD workshop Netherlands Society for Behavioral Biology, Soesterberg, NL (2015)
Workshop *Pitch jou onderzoek*, KNAW (2016)
Workshop *Principles of good academic research*, KNAW (2017)
Workshop *Scientific Integrity*, KNAW (2017)
Wageningen Ecology and Evolution Seminars (WEES), organizing member (2015-2018)
Journal club at NIOO-KNAW (2015-2018)

Presentation Skills (4 ECTS)
Birds and climate change - Utrechtse Biologen Vereniging (invited), oral, Utrecht, NL (2016)
International Symposium for Avian Endocrinology (ISAE), oral, Niagara Falls, USA (2016)
Biotechnische dagen (invited), oral, Venray, NL (2017)
Society for Integrative and Comparative Biology (SICB), oral, New Orleans, USA (2017)
European Society for Evolutionary Biology (ESEB) congress, poster, Groningen, NL (2017)

Teaching competences (2.5 ECTS)
Supervising fieldwork student (2015)
Supervising MSc student (2016)

EDUCATION AND TRAINING TOTAL: 30.9 ECTS

Colophon

The research presented in this thesis was conducted at the Department of Animal Ecology of the Netherlands Institute of Ecology (NIOO-KNAW), Wageningen.

The research described in this thesis was financially supported by the European Research Council.

This thesis is NIOO Thesis 172

Photos by courtesy of Elizabeth M. Grevenhof

Cover art by Wandalin van den Abeele

Printed by ProefschriftMaken.nl, Wageningen, The Netherlands