Letters

Selection on reproductive plasticity in a wild population of blue tits, *Cyanistes caeruleus*

Heung Ying Janet Chik[1], Catalina Estrada[1], Yiqing Wang[1], Priyesha Vijendra[1], Alex Lord[1], Julia Schroeder[1]

[1]Department of Life Sciences, Imperial College London, Silwood Park, Ascot, SL5 7PY.

Corresponding author: Heung Ying Janet Chik <heung.chik18@imperial.ac.uk>

Running title: Blue tit reproductive plasticity

Keywords: Phenotypic plasticity, reaction norm, avian reproduction, quantitative genetics, selection pressure, climate change

Abstract word count: 249

Total word count: 4,642

Abstract

In the plant-insect-insectivorous-bird system, changing climates can result in mis-timing in bird reproduction, potentially impacting chick survival. To adapt to earlier prey emergence, birds can make use of phenotypic plasticity, which can be characterized by reaction norms. Despite gaining focus in research, studies on avian reproductive reaction norms as traits are scarce, particularly on laying-date-budburst-date and clutch-size-laying-date reaction norms. Here we examined the possibility of evolution of these two reaction norms from a quantitative genetics viewpoint, and tested whether 1) there is among-individual variance in reaction norms properties (intercept and slope); 2) there is selection on these reaction norm properties; and 3) variances and selection pressures differ between the two reaction norms. Data of oak (genus Quercus) budburst and blue tit (*Cyanistes caeruleus*) reproduction were collected from a wild population for 18 years. We used bivariate random regression models with a Bayesian approach to test for among-individual variance in reaction norm properties and their covariance with fitness. Individuals significantly differed in intercepts and slopes of both laying-date-budburst-date and clutch-size-laying-date reaction norms, and directional selection was present for an earlier laying date and a larger clutch size, but not on either plasticity. Results suggested that variation in reaction norm properties can be attributed to genetic and environmental effects, and that stabilizing selection on plasticity could be tested as a next step. This study compliments previous research on the evolution of reaction norms and helps shed light on their genetic structure, the context of their selection, and their mediation in birds.

Impact summary

In the face of climate change, it is essential to understand how organisms adapt to changing environments and novel challenges in both the short and the long term. Since phenotypic plasticity provides a powerful means for temperate birds to adapt to dramatic environmental changes, understanding its origin, how it is mediated, and its possibility to evolve provides valuable insights in predicting the fates of temperate bird populations in the future. Our study focused on two little-examined forms of plasticity, laying-date-budburst-date and clutch-size-laying-date. We found individuals differ in both forms of plasticity, and direct selection only acts on the traits: laying date and clutch size, but not on either forms of plasticity. Our results complemented existing research to demonstrate that there is potential for evolution of plasticity, and brought attention to areas where further investigation is required, namely, the cues and physiological systems used by birds to time reproduction, and the genetic structure of plasticity. We provided new data from a UK population of blue tits, which enabled further comparisons of plasticity expression among populations across various geographical locations to better the understanding of avian adaptation mechanisms, and the biological and environmental means through which they are mediated. We also demonstrated a powerful but underused statistical method to estimate selection pressures, and thus assisted in promoting its use in the fields of evolutionary biology and quantitative genetics.

Introduction

Climate change poses many impacts on ecosystems, one of which being phenological mismatch, or the mistiming of life-history events in different trophic levels of a food chain. To maintain a food chain, phenological events across trophic levels are synchronized, such that the peak in abundance of the food source coincides with that of the food demand of higher trophic levels. However, as temperatures continue to rise (IPCC, 2018), spring phenological events in over 1,700 species have advanced at an average rate of 2.3 days per decade (Parmesan & Yohe, 2003), including those of the plant-insect-insectivore system, namely tree budburst (Menzel et al., 2001; Badeck et al., 2004), insect emergence (Roy & Sparks, 2000), and avian breeding (Both et al., 2005).

Phenological advancements lead to mismatch when species experience shifts of different magnitudes or directions, resulting in predator and prey, which are reliant on each other’s phenology, tracking climate change to different degrees. This is because different species possess different response mechanisms to changing environments. For example, in oaks (genus *Quercus*), leaf development and leaf palatability to herbivorous insects are largely dependent on temperature (Buse et al., 1999). In poikilothermic species whose body temperature varies greatly with the environment, such as the oak-leaf-eating winter moth (*Operophtera brumata*), egg-hatching and thus emergence is directly dependent on accumulated heat (Embree, 1970; Dewar & Watt, 1992), and hence temperature. Therefore, as temperature rises, one could expect similar phenological shifts in both oak and moth, hence a relatively small mismatch between the two. In contrast, insectivorous birds such as great tits (*Parus major*) and blue tits (*Cyanistes caeruleus*) have a more complex response mechanism. In such species, selection occurs after the hatching of eggs, when chicks are rapidly growing and thus possess the greatest need for caterpillars as food (Charmantier et al., 2008). As there exists a time lag between the start of breeding and egg-hatching, females must rely on environmental cues at the time of egg-laying to predict the conditions (and thus food availability) at which chicks have the greatest nutritional need, and time egg-laying accordingly. These cues might become unreliable as a result of climate change. Specifically, Visser et al. (2004) argued that there are two possible explanations to the variations in avian reproduction phenology – that climate change affects the environment at the time of decision-making differently than it affects the environment at the time of selection; or that birds may rely on different and/or multiple cues to different degrees, some of which, such as photoperiod, would remain relatively stable, leading to the species possessing a smaller phenological shift. If the shift in the insectivore does not match that of the insect and/or the host plant, mismatch arises (Fig. 1).

In insectivorous birds, phenological mismatch can result in increased selection pressure for earlier laying dates (Van Noordwijk et al., 1995; Visser et al., 1998; Both & Visser, 2001). While microevolutionary response to selection in laying date at the population-level has been observed (Møller et al., 2006; Gienapp et al., 2008), it is speculated to be less important for birds to adapt to shifting prey phenology, since laying date is only moderately heritable, leading to microevolution being a slow process and unable to keep up with the more rapid prey phenological changes (Gienapp et al., 2008; Charmantier & Gienapp, 2014). Instead, it is argued that phenotypic plasticity plays a greater role in providing a faster way to adapt to rapidly changing environments (Gienapp et al., 2008).

Phenotypic plasticity is defined as the expression of more than one phenotypic value from a single genotype across changing environments, and can be characterized by the reaction norm, a regression line of phenotypic values of the focal trait against an environmental gradient, of which the intercept represents the trait value at the average environment, and the slope represents plasticity (Fig 2, Stearns, 1989). Plasticity can be found in avian life-history traits such as laying date (Nussey et al., 2005; Charmantier et al., 2008; Porlier et al., 2012; Thorley & Lord, 2015), but more importantly, it has been shown that plasticity in laying date varies among individuals and is heritable, thus it could be subjected to selection and hence evolution (Nussey et al., 2005; Porlier et al., 2012).

Previous studies on plasticity have largely focused on laying-date-temperature reaction norms, which, while crucial, may be insufficient in providing a complete picture in understanding how birds adapt to phenological mismatch. This is because, while temperature remains the main driver of spring phenology in the plant-insect-insectivore system, it is ultimately tree budburst phenology that determines the emergence of food peaks for emerging insects, thus forming the basis of synchrony among trophic levels (Dewar & Watt, 1992). Therefore, it may be preferable to examine laying-date-budburst-date reaction norms instead of temperature to investigate avian adaptation through plasticity, especially when it remains largely uncertain what cues birds use to time egg-laying. Moreover, one should also consider plasticity in other life-history traits related to laying date, such as clutch size, as this trait is closely related to fitness (Rowe et al., 1994). The interaction between clutch size and laying date results in a trade-off to determine an optimal clutch size for every laying date (Lack, 1954). This optimization is governed by two considerations: that reproductive value of an egg declines seasonally with laying date; and that clutch size increases with laying date, since later breeding means that parents have more time to gather resources (Brommer et al., 2003). Overall, these considerations result in the trend that clutch size decreases with laying date (Brommer et al., 2003). Hence in this study, we characterized plasticity as the slope of clutch-size-laying-date reaction norms. A more plastic female may be advantageous and favored over a less plastic one since she can better adjust her optimal clutch size – should she lay early, she can lay more eggs with higher reproductive values; should she lay late, she can better minimize phenological mismatch, since a smaller clutch means she can more quickly proceed to incubation, as most insectivorous passerines are constrained to lay a maximum of one egg per day (Perrins, 1979).

Despite the importance of laying-date-budburst-date (LDBD) and clutch-size-laying-date (CSLD) plasticity, little research exists that examines the evolution of plasticity in such traits. Here, we investigated the possibility for evolution of both forms of plasticity using long-term data of a wild blue tit (*Cyanistes caeruleus*) population. To our knowledge, this is the first study on laying-date-budburst-date correlation using individually coupled tree and nest box data. We used an advanced but underused one-step statistical approach (Arnold et al., 2019) to assess selection on plasticity. We tested three hypotheses: 1) that there is among-individual variance in LDBD and CSLD reaction norms properties (intercept and slope), allowing the possibility for an evolutionary response; 2) that there is selection on these reaction norm properties; and 3) that variances and selection pressures differ between LDBD and CSLD reaction norms. Together with past research, this study shall provide insights on how plasticity may evolve as a trait.

Methods

Study species and site

This study used long-term data of a nest-box population of blue tits at Silwood Park, United Kingdom. The blue tit is a small passerine that commonly dwells in deciduous or mixed woodlands and breeds readily in holes or nest boxes (Svensson et al., 2009). Its breeding season commences in late-March and typically lasts until June, where females lay a single brood with up to 19 eggs according to previous records and feed their young with predominantly caterpillars. The study site has an area of approximately 100 ha and consists of deciduous woodlands ranging from 60 to several hundred years of age (Crawley, 2005). The site is dominated primarily by the English oak (*Quercus robur*), among other deciduous tree species. Within the site, 200 nest boxes were installed in 2002, with further changes in subsequent years, totaling 259 nest boxes as of 2019.

Data collection

Collection of blue tit breeding data began in 2002. Every year data collection began in late-March, and nest boxes were examined every other day for signs of nest-building and egg-laying. We recorded the laying date, defined as the date on which the first egg of each clutch is laid, in ‘April Days’, the number of days passed since the 1st of April (= Day 0) in a given year. Upon allowing 15 days for females to complete their clutches, we caught blue tits in their nest boxes and recorded the final clutch size. Birds were identified by uniquely numbered metal rings from the British Trust for Ornithology (BTO). We sexed birds by the presence of a brood patch, a patch of featherless, highly vascularized skin on the abdomen of females, and aged them by the coloration of the primary covert feathers on the wing. We separated birds into two age groups, ‘one year old or less (≤1)’, and ‘over one year old (>1)’. We then allowed 11 days before revisiting nests to check for egg hatching, upon which hatching date is recorded. We measured, weighed and fitted chicks with BTO metal rings and counted the number of ringed chicks. From 2002 to 2011, this was done when chicks were 7 days old, and from 2012 onwards, when chicks were 14 days old. We recorded the number of fledglings by revisiting nest boxes when chicks were 19 days old.

Oak leaf budburst phenology was monitored starting from 2007. We monitored in total 3,945 oak trees, each of which has a unique ID and a record of its nearest nest box. 10% of trees are monitored annually, and the remaining 90% biannually. Each year, we carried out tree monitoring from April to May. We visited each tree every two or three days and recorded the overall budburst score for that tree from not yet budding (stage 0) to fully tanninized (stage 6, Fig. 3) based on the majority of its leaves, until all trees have been scored at stage 6. Dates on which a tree reached a certain stage were recorded as April Days.

LDBD and CSLD relationships

We ran all models in R v.3.5.1 (R Core Team, 2018). Laying dates and budburst dates were standardized annually, so that within a year the mean date was scaled to zero, and the standard deviation to one. For each breeding record, we summarized the average dates on which the oak trees corresponding to its nest box reached budburst stages 1 to 6 respectively. We selected the average date where trees reach stage 3 as the budburst date, as this date presents the highest correlation with laying date (r = 0.0771) before the final stage 6, and thus most likely a cue for egg-laying. To examine the overall LDBD and CSLD relationships, we used the R package *lme4* (Bates et al., 2019) to run two models, a linear mixed model with laying date as the response variable and budburst date as the explanatory variable, and a generalized linear mixed model with clutch size as the response variable, laying date as the explanatory variable, and 'Poisson' as the data distribution family. In both models, we modelled age as an interacting fixed effect, and bird ID and year as random effects on the intercept to account for repeated measures of the same individuals and in the same year respectively.

Assessing selection pressures on reaction norms

This study follows the one-step approach demonstrated by Arnold et al. (2019), using a bivariate generalized linear mixed model to assess selection pressures on LDBD and CSLD reaction norms. We illustrate the approach here with the LDBD reaction norm. First, we constructed a mixed model for individual reaction norm properties:

$$\mathrm{LD} \sim \mathrm{BD} * \mathrm{age} + (\mathrm{LD})_{year} + (\mathrm{LD})_{ID} + (\mathrm{LD{:}\,BD})_{ID}$$

where LD is laying date, BD is the fixed covariate of budburst date, age is a fixed factor, $(\mathrm{LD})_{year}$ is the random effect of year on the intercept, $(\mathrm{LD})_{ID}$ is the random effect of individual birds' intercepts, and $(\mathrm{LD{:}\,BD})_{ID}$ is the random effect of individual birds' slopes. At the individual random effect, this model has the following variance-covariance structure:

$$P_2 = \begin{bmatrix} {\sigma^2}_{LD} & \sigma_{LD,LD:BD} \\ \sigma_{LD,LD:BD} & {\sigma^2}_{LD:BD} \end{bmatrix}_{ID}$$

where ${\sigma^2}_{LD}$ is the among-individual variance in intercept, ${\sigma^2}_{LD:BD}$ is the among-individual variance in slope, and $\sigma_{LD,LD:BD}$ is the covariance between the intercept and the slope.

Next, we extended this model to a bivariate one by considering the equation

$$\omega = \mu_{\omega} + (\omega)_{ID}$$

where ω is individual fitness, μ_{ω} is the mean fitness of the population, and $(\omega)_{ID}$ here is the deviation of individual birds' fitness values from the mean. Since there is one fitness value for each individual, $(\omega)_{ID}$ can also be treated as a random effect of individual birds' fitness in the bivariate model, resulting in the following variance-covariance structure in the final model, with three levels at the individual effect:

$$P = \begin{bmatrix} {\sigma^2}_{LD} & \sigma_{LD,LD:BD} & \sigma_{LD,\omega} \\ \sigma_{LD,LD:BD} & {\sigma^2}_{LD:BD} & \sigma_{LD:BD,\omega} \\ \sigma_{LD,\omega} & \sigma_{LD:BD,\omega} & {\sigma^2}_{\omega} \end{bmatrix}_{ID} + {\sigma^2}_{year} + {\sigma^2}_{residual}$$

where ${\sigma^2}_{\omega}$ is the among-individual variance in fitness, $\sigma_{LD,\omega}$ is the covariance between individual reaction norm intercept and fitness, $\sigma_{LD:BD,\omega}$ is the covariance between individual reaction norm slope and fitness, ${\sigma^2}_{year}$ is the among-year variance in intercept in LD, and ${\sigma^2}_{residual}$ is the residual variance in LD. $\sigma_{LD,\omega}$ and $\sigma_{LD:BD,\omega}$ are thus selection differentials for individual intercepts and slopes respectively, and represents the total selection on individual intercepts and slopes. If these selection differentials are concatenated into a vector

S, then direct selection can further be obtained by calculating the selection gradients of individual intercepts and slopes β using

$$\beta = {P_2}^{-1} S$$

where ${P_2}^{-1}$ denotes the inverse matrix of P_2 (Lande & Arnold, 1983).

Likewise, we assessed selection on CSLD reaction norm using the same approach. We used lifetime breeding success (LBS) in terms of the number of ringed chicks a female produced throughout her breeding career as the fitness measure (ω), and discarded breeding data of the latest year (2019) under the assumption that all females breeding in that year have not yet completed their breeding careers. We removed outlier observations with clutch sizes over 20 as they were likely the result of recording errors and/or of multiple birds breeding in the same nest box. In total, 1,284 females, 3,945 oak trees, 615 breeding observations for LDBD analyses and 1,957 observations for CSLD analyses were included in the final models. We used the R package *MCMCglmm* v2.26 (Hadfield, 2019), a package using Bayesian modelling approach to run these models, with 15 million iterations, 1.5 million burn-ins and 10,000 thinning intervals. Degrees of belief in the prior distributions were set to be equal to the dimension of the among-individual variance-covariance matrix i.e. 3 to ensure propriety (Hadfield, 2009). We determined model posterior modes as significantly different from zero when their 95% credible intervals (CI) do not overlap zero.

Results

Descriptive statistics

Of the 1,449 blue tit females recorded in this study, 945 (65.217%) were one-time breeders, and the maximum number of breeding observations was seven. A summary of breeding records was provided in Table 1. The average LBS of the 1,284 females used in the analyses was 8.425 chicks, with a range of 0 to 53, and a variance of 62.435. Summaries of blue tit breeding data and oak budburst data were provided in Tables 2 and 3, respectively.

LDBD relationship

The linear mixed model revealed no statistically significant relationship between laying date and budburst date (Table 4 & Fig. 4). Birds with age >1 did not differ significantly from those of age ≤1.

Selection on LDBD reaction norm

There was non-zero among individual variance in both reaction norm intercept and slope, and a non-zero selection differential on intercepts for an earlier laying date (Table 5). There was zero selection differential on reaction norm slope. After transforming selection differentials, the selection gradient for individual intercepts was -0.890 (95% CI = -1.397 – -0.411), and that of individual slopes was -0.283 (95% CI = -1.044 – 0.576), indicating direct directional selection on reaction norm intercept but not on slope.

CSLD relationship

The generalized linear mixed model revealed a statistically significant and negative correlation between clutch size and laying date (Table 6 & Fig. 5). For birds of age ≤1, the slope translated to approximately 8.64% decrease in clutch size per unit increase in laying date. Birds of age >1 did not significantly differ in either intercept or slope (Table 6).

Selection on CSLD reaction norm

There was among-individual variance in CSLD intercepts and slopes, though smaller than that displayed in the LDBD model (Table 7). There was a small but significant selection differential on individual intercepts for a larger clutch size. A significant selection differential on individual slopes, on the other hand, could not be detected.

Transforming total selection differentials resulted in a selection gradient of 1.93 (95% CI =

0.931 – 3.01) on individual intercepts and a gradient of 0.163 (95% CI = -0.927 – 1.05) on individual slopes. Results indicated a non-zero direct selection pressure on individual intercepts for a larger clutch size, but a lack of direct selection on plasticity.

Discussion

LDBD and CSLD relationships

Oak budburst is coupled with caterpillar emergence and abundance (Nilsson & Källander, 2006), hence if birds are to successfully time reproduction, one would expect a positive correlation between laying date and oak budburst date – the earlier the emergence of new leaves, the earlier the egg-laying. However, the results demonstrated otherwise, implying this population did not utilize oak budburst as an important cue for egg-laying. This was in line with some studies which showed that blue tits did not respond to oak phenology in both egg-laying and reproductive hormone secretion (Visser et al., 2002; Nilsson & Källander, 2006; Schaper et al., 2011), but contradicted others (e.g. Bourgault et al., 2010). In particular, Thorley and Lord (2015) showed a correlation between temperature and laying date in this population, implying temperature as a main cue used by birds. Since temperature is often correlated with biotic changes, it is likely that causal relationships between oak phenology and bird breeding found in other studies are merely apparent, and further effort is required to tease apart the contributions of different cues and the physiological mechanisms used to time avian reproduction.

On the other hand, clutch size displayed a decline with laying date, consistent with the predicted outcome of the trade-off between the two traits (Lack, 1954). An earlier laying date coincides better with the caterpillar abundance peak, and translates to more food available for chick rearing, resulting in heavier chicks with a higher chance of survival after fledging (Perrins, 1965). As spring passes, food availability diminishes and parents are unable to feed as many chicks as during the start of the season. Furthermore, feeding effort of parents do not increase proportionally with brood size (Gibb, 1955), and the larger the brood, the less food each chick receives. Considering these, it is therefore a better strategy to lay fewer eggs as the season progresses, so as to ensure success of all chicks in the brood. This ultimately creates the negative relationship between clutch size and laying date shown in this study.

Among-individual variation in LDBD and CSLD reaction norms

We demonstrated that individuals possess the ability to adjust laying date and clutch sizes under their own regimes, as there was significant among-individual variance in the slopes and intercepts of both LDBD and CSLD reaction norms. Since trait variation is essential for natural selection, there is capacity for LDBD and CSLD reaction norms to be subjected to selection. This among-individual variation could be attributed to two sources. First is a genetic system controlling the expression of reaction norms, such that genetically related individuals display less variation than non-related individuals would. Second is the effect of the environment. Birds have the ability to learn – those that have experienced a warmer spring begin egg-laying earlier in the subsequent year and vice versa, for example (Nussey et al., 2005). Within a population, individuals may experience a unique set of environmental changes throughout their lifetimes based on their location, its associated microclimate, and chance. Thus each individual could ‘learn’ differently and develop varying reaction norm properties, optimized to their local environments (Brommer et al., 2003).

To quantify the contributions of genes and the environment to the variation in reaction norm properties, heritability analysis is essential. Heritability in the broad sense refers to the proportion of genetic variance over the total phenotypic variance (Goldberger, 1979), and requires a genetic pedigree to be estimated. The proportion of non-genetic variation is attributed to environmental effects. Heritability has been proven in laying-date-temperature reaction norms properties (Nussey et al., 2005; Charmantier et al., 2008; Porlier et al., 2012), but not on LDBD and CSLD reaction norms, calling for further research effort. In addition, to examine environmental effects, it would be sensible to compare reaction norms of birds of different number of breeding records. If learning plays a critical role in shaping plasticity, one could expect older and more experienced birds to display a steeper slope in their reaction norms than younger birds. While this difference was not found in this study, it could be because the true age difference in the two age groups remained largely unknown. Thus, breeding records might prove a better predictor.

Furthermore, our results show that LDBD reaction norm properties possess much higher among-individual variance than those of the CSLD reaction norm. Brommer et al. (2012) theorized that there should exist an optimal reaction norm to maximize reproductive output in a particular set of environments i.e. in a particular population. When there is deviation of individual reaction norms from the optimum, fitness is reduced, and selection drives individuals towards the optimal reaction norm, ultimately decreasing among-individual variance. The results here may thus mean that in the Silwood Park population, CSLD reaction norm has already been pushed closer to the optimum by selection than has LDBD reaction norm (Charmantier et al., 2008). Without heritability estimates, it is difficult to conclude that CSLD reaction norm had the capacity to evolve. Nevertheless, one could expect evolution on the CSLD reaction norm to halt before the LDBD reaction norm does, due to little and decreasing individual variation in the former.

Selection on LDBD and CSLD reaction norms

The results indicated direct selection pressure towards an earlier laying date and a larger clutch size, which is consistent with each other and with previous literature (Van Noordwijk et al., 1995; Brommer et al., 2012; Thorley & Lord, 2015). An earlier laying date allows better synchrony with food abundance, and a larger associated clutch size, pushing individuals towards a higher overall reproductive output. These selection pressures, along with the presence of among-individual variation, strongly suggest the ability for both reaction norm intercepts to evolve. Contrary to common findings (Brommer et al., 2005; Nussey et al., 2005), there was no significant covariance between intercept and slope in both reaction norms, meaning that in this population selection on the intercept is unlikely to result in indirect selection on the slope.

The lack of significant covariance between slope (plasticity) and fitness in both LDBD and CSLD reaction norms inferred that the more plastic females did not perform better or worse than the less plastic females in reproductive output. Selection gradients also indicated no directional selection on plasticity for both reaction norms. In LDBD plasticity, this could be explained by the possibility that selection favors reaction norms that enable birds to achieve maximum synchrony with oak budburst. As the emergence of food peak is only momentary, there is a narrow window for birds to reproduce. This means that to ensure a well-timed laying date with caterpillar abundance across years, an optimal reaction norm slope i.e. an intermediate response to budburst, is needed (Reed et al., 2006). When budburst phenology varies among years, as in the study site (Table 3), an overly-plastic female would hurry laying too much ahead of the caterpillar peak in an early-budburst year, and delay laying too much in a late-budburst year, thus falling out of synchrony. On the contrary, a non-plastic female would lay too late in an early-budburst year, and too early in a late-budburst year, and likewise fall out of synchrony. As such, the highest fitness should be associated with the optimal plasticity, and selection should drive individuals towards the single reaction norm slope in favor of more extreme ones. In other words, stabilizing selection would occur. In CSLD plasticity, the case is similar – an overly-plastic female suffers a reduction in number of chicks produced greater than the gain from improved chick survival, and vice versa (Brommer et al., 2012). An essential next step, therefore, would be to examine stabilizing selection on reaction norm slope, which could be achieved by detecting directional selection on the square term of the slope (Reed et al., 2006; Brommer et al., 2012).

It is worth noting, however, that the inability to detect directional selection in this study could also be attributed to limitations in estimating fitness. Silwood Park blue tits are an open population (Table 1), and LBS estimates are thus prone to errors, as females might have raised broods elsewhere, resulting in the underestimation of reproductive output overall, and an upward bias in LBS towards birds with more recorded breeding observations. Nevertheless, LBS remains one of the most widely used fitness measures (e.g. Slate et al., 2000; Brommer et al., 2005; Nussey et al., 2005) in wild populations, and thus appropriate for this study.

Bivariate random regression models to estimate selection

In this study we demonstrated a statistical approach capable of estimating among-individual variance in reaction intercept and slope, covariance in intercept and slope, and selection differentials on intercept and slope simultaneously. This approach has advantages over a conventional two-step method, which requires: 1) characterizing the among-individual variance in reaction norm intercepts and slopes, and 2) calculating selection pressure by regressing a lifetime reproductive fitness measure on these intercepts and slopes (e.g. Nussey et al., 2005). The latter is essentially stats-on-stats, where statistical errors in Step 1 would be carried over to Step 2 (Arnold et al., 2019). In addition, to perform Step 2, one can either utilize estimates from a simple linear regression, or best linear unbiased predictors (BLUP) of random effects from mixed models in Step 1 (Brommer et al., 2012). The former allows only data of individuals with a fairly large number of repeated measurements, thus discarding potentially a large proportion of data, while the latter violates the assumption that BLUP values are derived when all variables affecting the response variable have been included (Brommer et al., 2012). Bivariate models do not have these limitations and are thus an advanced way to assess selection pressure. In addition, they allow also the estimation of selection on non-linear reaction norms, by fitting a quadratic or higher order function as the individual trait-environment/trait-trait relationship (Arnold et al., 2019). These models are, however, unable to detect non-linear selection pressures on plasticity, which requires the covariance between fitness and the square term of the slope. Since in the model the slope is developed from the within-individual covariance between the focal trait and the environment/predictor trait, one cannot directly manipulate it to obtain a square term. As such, conventional methods will need to be employed for non-linear selection analyses in the future.

Conclusions

Using breeding data of a wild blue tit population, along with tightly coupled oak phenology data, we examined whether laying-date-budburst-date and clutch-size-laying-date reaction norms have the potential to evolve. Laying date did not correlate with budburst date, while clutch size decreased with laying date. We found significant among-individual variance in the properties of both reaction norms, the intercept (individual laying date/clutch size) and the slope (plasticity), suggesting the possibility of a genetic effect. We found directional selection for an earlier laying date and a larger clutch size, but no directional selection of both LDBD and CSLD plasticities, suggesting stabilizing selection might be present instead. While research in phenotypic plasticity is gaining momentum, it will take further effort to unravel the mechanisms by which evolution of plasticity operates. By providing new findings on plasticity and demonstrating an advanced statistical method, this study contributed to unravelling the genetic structure of plasticity, the context of selection on plasticity, and the cues through which organisms mediate plastic responses.

Acknowledgements

We thank C. Wilson of the British Trust for Ornithology for providing rings, training and supervision in the field, and numerous field assistants for collecting blue tits breeding and oak phenology data throughout the years. We thank J. Hadfield for advice and support on statistical model construction. The Blue Tit Monitoring Project was initiated with a UK Natural Environment Research Council grant, and supported by Imperial College London. The authors declare no conflicts of interest.

Author contributions

HYJC conceived the study. HYJC, CE, YW and PV collected the data in 2019, and CE and JS coordinated data collection. HC analyzed the data with input from YW and PV. HYJC drafted the initial version of the manuscript and all authors contributed to later versions. AL and JS supervised the project. All authors approved the final version of the manuscript to be published.

References

Arnold, P. A., Nicotra, A. B. & Kruuk, L. E. B. (2019). Sparse evidence for selection on phenotypic plasticity in response to temperature. *Phil. Trans. R. Soc. B* 374.

Badeck, F. W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J. et al. (2004). Responses of spring phenology to climate change. *New Phytol.* 162: 295–309.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H. et al. (2019). Package “lme4.”

Both, C., Piersma, T. & Roodbergen, S. P. (2005). Climatic change explains much of the 20th century advance in laying date of Northern Lapwing *Vanellus vanellus* in the Netherlands. *Ardea* 93: 79–88.

Both, C. & Visser, M. E. (2001). Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature* 411: 296–298.

Bourgault, P., Thomas, D., Perret, P. & Blondel, J. (2010). Spring vegetation phenology is a robust predictor of breeding date across broad landscapes: a multi-site approach using the Corsican blue tit (*Cyanistes caeruleus*). *Oecologia* 162: 885–892.

Brommer, J. E., Kontiainen, P. & Pietiäinen, H. (2012). Selection on plasticity of seasonal life-history traits using random regression mixed model analysis. *Ecol. Evol.* 2: 695–704.

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* 59: 1362–1371.

Brommer, J. E., Pietiäinen, H. & Kolunen, H. (2003). Natural selection on individual clutch size-laying date trends in the Ural owl. *Evol. Ecol. Res.* 5: 229–237.

Buse, A., Dury, S. J., Woodburn, R. J. W., Perrins, C. M. & Good, J. E. G. (1999). Effects of elevated temperature on multi-species interactions: the case of Pedunculate Oak, Winter Moth and Tits. *Funct. Ecol.* 13: 74–82.

Charmantier, A. & Gienapp, P. (2014). Climate change and timing of avian breeding and migration: Evolutionary versus plastic changes. *Evol. Appl.* 7: 15–28.

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C. M., Kruuk, L. E. B. & Sheldon, B.

C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* 320: 800–803.

Crawley, M. J. (2005). The Flora of Berkshire. Brambleby Books, Harpenden, UK.

Dewar, R. C. & Watt, A. D. (1992). Predicted changes in the synchrony of larval emergence and budburst under climatic warming. *Oecologia* 89: 557–559.

Doblas, L. L. (2017). Field Season Protocol for the Blue Tit Project.

Embree, D. G. (1970). The diurnal and seasonal pattern of hatching of winter moth eggs, *Operophtera brumata* (Geometridae: Lepidoptera). *Can. Entomol.* 102: 759–768.

Gibb, J. (1955). Feeding rates of great tits. *Br. Birds* 48: 49–58.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A. & Merilä, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* 17: 167–178.

Goldberger, A. S. (1979). Heritability. *Economica* 46: 327–347.

Hadfield, J. (2009). Tutorial for MCMCglmm package in R.

Hadfield, J. (2019). Package ‘MCMCglmm’

IPCC. (2018). Global Warming of 1.5 °C: Summary for Policymakers.

Lack, D. (1954). The Natural Regulation of Animal Numbers. Oxford University Press, Oxford, UK.

Lande, R. & Arnold, S. J. (1983). The measurement of selection on correlated characters.

Evolution 37: 1210–1226.

Menzel, A., Estrella, N. & Fabian, P. (2001). Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. *Glob. Chang. Biol.* 7: 657–666.

Møller, A. P., Flenstad-Jensen, E. & Mardal, W. (2006). Rapidly advancing laying date in a seabird and the changing advantage of early reproduction. *J. Anim. Ecol.* 75: 657–665.

Nilsson, J. Å. & Källander, H. (2006). Leafing phenology and timing of egg laying in great tits *Parus major* and blue tits *P. caeruleus*. *J. Avian Biol.* 37: 357–363.

Nussey, D. H., Postma, E., Gienapp, P. & Visser, M. E. (2005). Selection on heritable phenotypic plasticity in a wild bird population. *Science* 310: 304–306.

Parmesan, C. & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421: 37–42.

Perrins, C. M. (1979). British Tits. HarperCollins Publishers, London, UK.

Perrins, C. M. (1965). Population fluctuations and clutch-size in the Great Tit, *Parus major*. *J. Anim. Ecol.* 34: 601–647.

Porlier, M., Charmantier, A., Bourgault, P., Perret, P., Blondel, J. & Garant, D. (2012). Variation in phenotypic plasticity and selection patterns in blue tit breeding time: Between- and within-population comparisons. *J. Anim. Ecol.* 81: 1041–1051.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Reed, T. E., Wanless, S., Harris, M. P., Frederiksen, M., Kruuk, L. E. B. & Cunningham, E. J. (2006). Responding to environmental change: plastic responses vary little in a synchronous breeder. *Proc. R. Soc. B* 273: 2713–2719.

Rowe, L., Ludwig, D. & Schluter, D. (1994). Time, condition, and the seasonal decline of avian clutch size. *Am. Nat.* 143: 698–722.

Roy, D. B. & Sparks, T. H. (2000). Phenology of British butterflies and climate change. *Glob. Chang. Biol.* 6: 407–416.

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*). *J. Exp. Biol.* 214: 3664–3671.

Slate, J., Kruuk, L. E. B., Marshall, T. C., Pemberton, J. M. & Clutton-Brock, T. H. (2000). Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus*). *Proc. R. Soc. B* 267: 1657–1662.

Stearns, S. C. (1989). The evolutionary significance of phenotypic plasticity. *Bioscience* 39: 436–445.

Svensson, L., Mullarney, K., Zetterström, D. & Grant, P. J. (2009). Collins Bird Guide. 2nd edition. HarperCollins Publishers, London, UK.

Thorley, J. B. & Lord, A. M. (2015). Laying Date is a plastic and repeatable trait in a population of blue tits *Cyanistes caeruleus*. *Ardea* 103: 69–78.

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. *J. Anim. Ecol.* 64: 451–458.

Visser, M. E., Lessells, C. M., Van Noordwijk, A. J. & Tinbergen, J. M. (1998). Warmer springs lead to mistimed reproduction in great tits. *Proc. R. Soc. B* 265: 1867–1870.

Visser, M. E., Both, C. & Lambrechts, M. M. (2004). Global climate change leads to mistimed avian reproduction. *Adv. Ecol. Res.* 35: 89–110.

Visser, M. E., Silverin, B., Lambrechts, M. M. & Tinbergen, J. M. (2002). No evidence for tree phenology as a cue for the timing of reproduction in Tits *Parus spp*. *Avian Sci.* 2: 77–86.

Tables and figures

Table 1. Number of females and associated number of breeding observations (2002-2019).

Number of breeding observations	*Number of females*
1	945
2	283
3	144
4	55
5	14
6	6

7		2
	Total	1449

Table 2. Summary of blue tit breeding data. LD = raw laying date, CS = clutch size.

Year	*No. of nests*	*Mean LD (Range)*	*Var (LD)*	*Mean CS (Range)*	*Var (CS)*	*No. of ringed chicks*	*Mean no. of ringed chicks per female*
2002	103	12.60 (3–25)	28.65	9.70 (4–14)	3.33	853	8.28
2003	131	22.27 (5–34)	19.31	8.11 (1–13)	2.57	734	5.60
2004	153	23.14 (12–41)	26.91	8.86 (3–14)	3.19	991	6.48
2005	193	19.63 (6–34)	38.36	9.00 (1–15)	4.40	668	3.46
2006	82	24.56 (19–35)	12.27	9.27 (6–12)	1.90	676	8.24
2007	179	14.91 (7–28)	16.89	9.69 (4–19)	3.40	890	4.97
2008	93	17.56 (5–30)	40.44	10.57 (6–16)	4.99	776	8.34
2009	122	13.74 (5–41)	34.09	10.98 (6–18)	4.16	811	6.65
2010	82	20.51 (7–42)	46.38	10.41 (5–14)	4.00	708	8.63
2011	134	13.47 (1–36)	23.23	10.01 (6–15)	2.31	1175	8.77
2012	171	15.81 (1–42)	66.98	9.24 (3–15)	5.74	483	2.82
2013	153	33.16 (24–53)	31.59	8.37 (3–17)	2.72	468	3.06
2015	134	21.61 (12–48)	57.23	8.66 (2–13)	3.32	540	4.03
2016	139	25.57 (10–49)	64.17	8.07 (4–13)	2.81	393	2.83
2017	135	12.68 (-3–42)	94.44	9.21 (4–14)	3.74	541	4.01
2018	108	21.17 (1–38)	22.10	9.58 (5–15)	2.99	714	6.61
2019	171	9.34 (-7–37)	70.66	9.55 (5–15)	3.20	884	5.17

Table 3. Summary of oak budburst data. BD = raw budburst date, defined as when a tree reaches stage 3.

Year	*No. of oaks measured*	*Mean no. of oaks measured per nest box*	*Mean BD (range)*	*Var (BD)*
2007	423	10.32	11.06 (1–25)	18.31
2008	661	7.18	25.22 (3–40)	40.62

2009	1032	8.82	17.64 (-2–40)	33.87
2010	1629	10.31	28.34 (10–63)	44.64
2011	1699	11.48	12.42 (-2–29)	17.71
2012	1813	11.62	25.96 (-7–56)	136.03
2013	1844	8.78	33.82 (2–52)	13.26
2015	346	2.98	21.16 (10–41)	23.60
2016	534	2.64	25.29 (12–41)	38.32
2017	477	2.59	17.02 (5–55)	58.82
2018	812	4.00	22.93 (16–42)	12.87

Table 4. LMM summary on LDBD relationship. BD = budburst date. *Yearly variance undetected due to annual standardization of LD. **2013 data excluded due to lack of age records.

Fixed effects

	Estimate	*Standard error*	*t-value*
Intercept	-0.074	0.050	-1.465
BD	0.024	0.049	0.479
age>1	-0.053	0.081	-0.653
BD:age>1	0.142	0.083	1.704

Random effects

Group	*Variance*	*No. of groups*
ID	0.204	482
year	0.000*	10**
residual	0.719	615

Table 5. Model summary of LDBD reaction norms, showing variance-covariance matrix for individual reaction norm intercept, slope and LBS, plus other random and fixed effects. Variances are on the diagonal while covariances are on the sides. Bolded text indicates

selection differentials on individual intercepts and slopes respectively. DIC = 8,379. LD = laying date, BD = budburst date, LBS = lifetime breeding success.

Variance-covariance matrix

	Post. Mean (95% CI)		
	LD	LD:BD	LBS
LD	0.225 (0.155 – 0.355)	0.028 (-0.038 – 0.079)	**-0.244 (-0.347 – -0.117)**
LD:BD	0.028 (-0.038 – 0.079)	0.152 (0.099 – 0.244)	**-0.088 (-0.189 – 0.077)**
LBS	**-0.244 (-0.347 – -0.117)**	**-0.088 (-0.189 – 0.077)**	1.452 (1.269 – 1.630)

Random effects

	Post. Mean	*95% CI*	*Effective sample size*
year	0.164	0.045 – 0.336	1168
residual	0.649	0.540 – 0.762	1350

Fixed effects

	Post. Mean	*95% CI*	*pMCMC*	*Eff sample size*
intercept	-0.023	-0.290 – 0.275	0.865	1350
LD:BD	0.055	-0.080 – 0.164	0.379	1509
LD:age>1	-0.081	-0.250 – 0.084	0.375	1350
LD:BD:age>1	0.108	-0.068 – 0.316	0.271	1350

Table 6. Summary of generalized linear mixed model on CSLD relationship. LD = laying date. Sample size = 1,957. AIC = 8,168.3. *Individual variance close to zero. **2013 data excluded due to lack of age records.

Fixed effects

	Estimate	*Standard error*	*z-value*	*Pr(>|z|)*
Intercept	2.229	0.023	94.870	<0.001
LD	-0.090	0.009	-9.532	<0.001
age>1	0.006	0.018	0.336	0.737
BD:age>1	-0.035	0.017	-1.940	0.052

Random effects

Group	*Variance*	*No. of groups*

ID	0.000*	1266
year	0.007	15**

Table 7. Model summary of clutch-size-laying-date plasticity, showing variance-covariance matrix for individual reaction norm intercepts, slopes and LBS, and other random effects. Variances are on the diagonal while covariances are on the sides. Bolded text indicates selection differentials on individual intercepts and slopes respectively. DIC = 16,033. CS = clutch size, LD = laying date, LBS = lifetime breeding success.

Variance-covariance matrix

	Post. Mean (95% CI)		
	CS	CS:LD	LBS
CS	0.019 (0.016 – 0.022)	0.001 (-0.001 – 0.004)	**0.032 (0.013 – 0.059)**
CS:LD	0.001 (-0.001 – 0.004)	0.022 (0.018 – 0.026)	**0.002 (-0.021 – 0.031)**
LBS	**0.032 (0.013 – 0.059)**	**0.002 (-0.021 – 0.031)**	1.434 (1.290 – 1.641)

Random effects

	Post. Mean	*95% CI*	*Effective sample size*
year	0.085	0.034 – 0.150	1136
residual	0.010	0.008 – 0.012	1350

Fixed effects

	Post. Mean	*95% CI*	*pMCMC*	*Eff sample size*
intercept	2.203	2.063 – 2.352	<0.001	1246
CS:LD	-0.093	-0.122 – -0.069	<0.001	1350
CS:age>1	0.016	-0.025 – 0.059	0.468	1512
CS:LD:age>1	-0.035	-0.082 – 0.016	0.182	1350

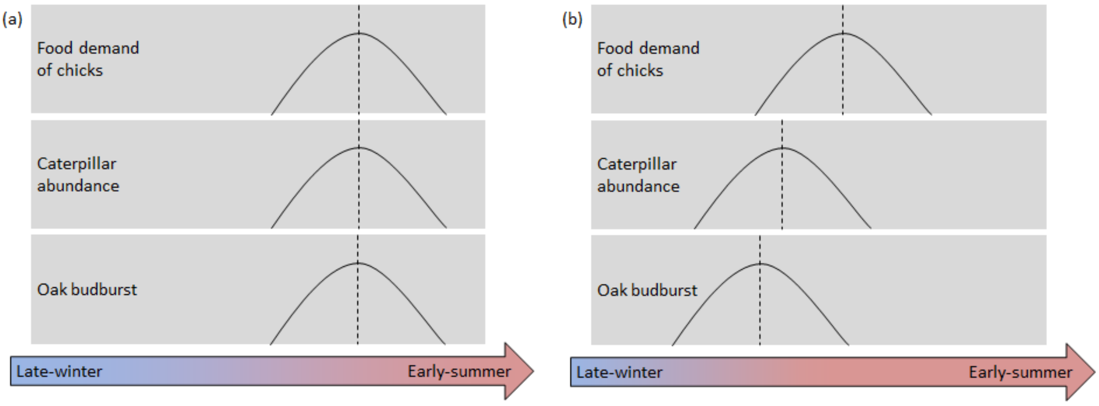

Fig. 1. Schematic diagram of phenological shifts. (a) depicts phenological match where oak budburst, caterpillar emergence and chick food demand align with one another. (b) depicts mismatch under advancing spring, where laying date of birds remains late, and chick food demand peaks later than the peak of food availability.

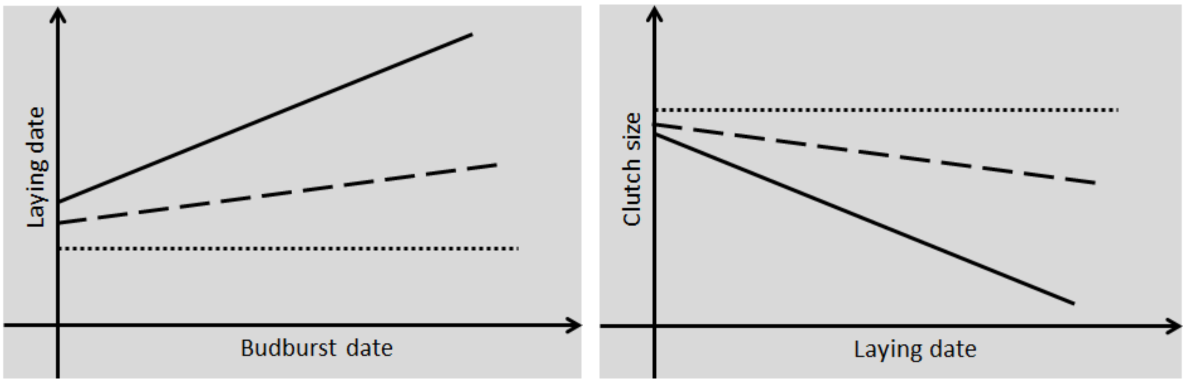

Fig. 2. Schematic diagrams of LDBD and CSLD reaction norms examined in this study. Lines represent linear regressions of individual reaction norm, which differ in intercept (mean trait value of the individual) and slope (plasticity). Solid line represents a more plastic reaction norm; dashed line represents a less plastic one, and dotted line represents a non-plastic one.

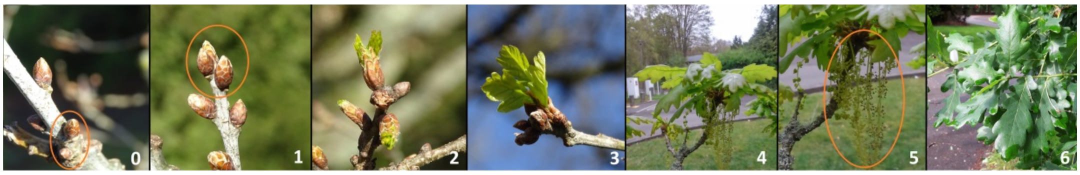

Fig. 3. Leaf stages at different scores. Adopted from Doblas, 2017

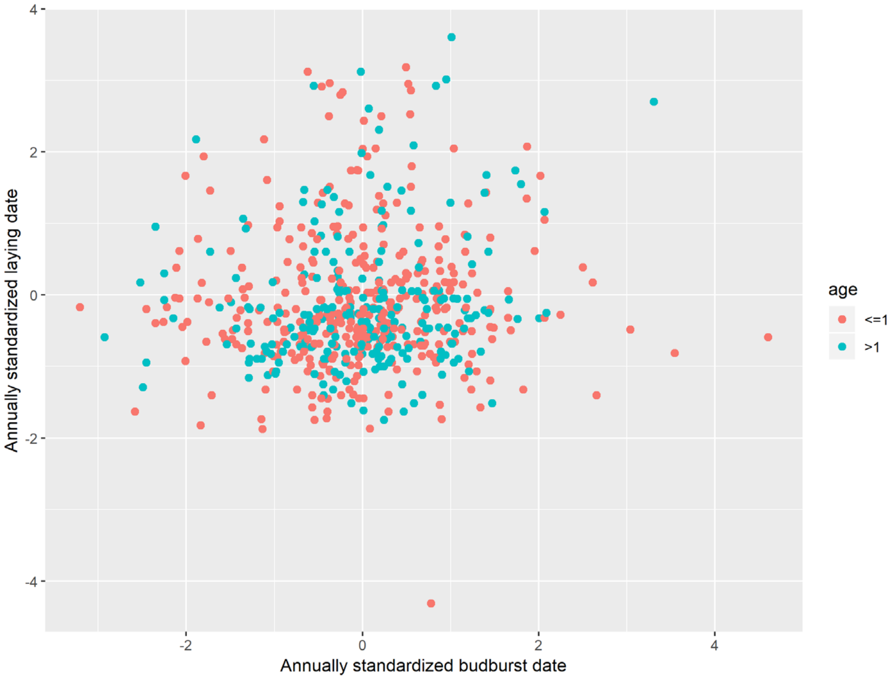

Fig. 4. Relationship between laying date and budburst date. Red dots represent birds one year old or less; blue dots represent birds more than one year old.

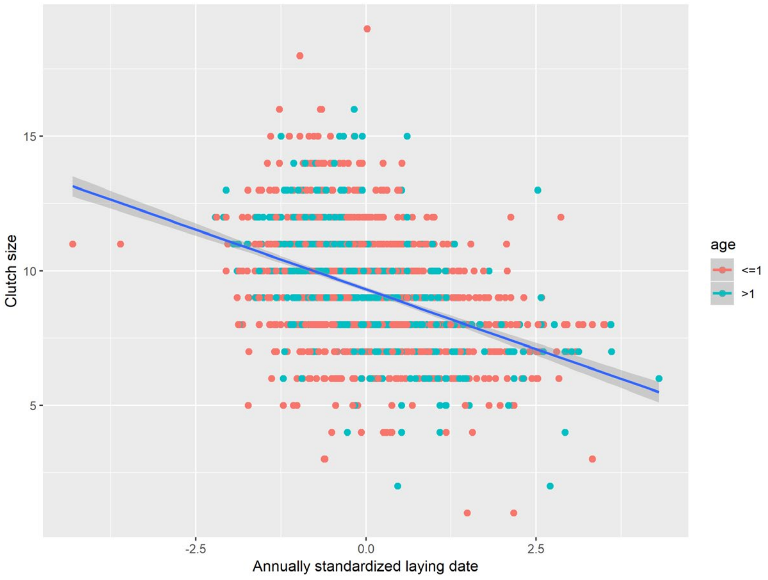

Fig. 5. Relationship between clutch size and laying date. Red dots represent birds one year old or less; blue dots represent birds more than one year old. Shaded area represents 95% confident interval.