

# Variation In Shape And Consistency Of Selection Between Populations Of The Threatened Hihi (*Notiomystis cincta*).

## Abstract

The shape and intensity of natural selection can vary between years, potentially resulting in a chronic reduction of fitness as individuals need to track a continually changing optimum of fitness (i.e., a “lag load”). In endangered species, often characterised by small population size, the lack of genetic diversity is expected to limit the response to this constant need to adjust to fluctuating selection, increasing the fitness burden and thus the risk of extinction. Here, we use long-term monitoring data to assess whether the type of selection for a key fitness trait (i.e., lay date) differs between two reintroduced populations of a threatened passerine bird, the hihi (*Notiomystis cincta*). We apply recent statistical developments to test for the presence or absence of fluctuation in selection in both the Tiritiri Mātangi Island and the Kārori sanctuary populations. Our results support the presence of stabilizing selection in Tiritiri Mātangi with a potential moving optimum for lay date. In Kārori our results favour a regime of directional selection. Although the shape of selection may differ, for both populations an earlier lay date generally increases fitness in both environments. Further, the moving optimum models of lay date on Tiritiri Mātangi, suggesting that selection varies between years, imply a substantial lag load in addition to the fitness burden caused by the population laying too late. Our results highlight the importance of characterising the form and temporal variation of selection for each population to predict the effects of environmental change and to inform management.

**Keywords:** Adaptation, Conservation, Fluctuating environment, Lay date, Phenology, Phenotypic plasticity, Stitchbird

## INTRODUCTION

Environmental variation is a fundamental aspect of evolution: fluctuations in environmental conditions over time can generate fluctuations in natural selection, which in turn can shape phenotypic variation (Darwin, 1859). The theoretical relationship between fluctuating selection and phenotypic variation has been well-studied, notably through moving optimum models. In such models, selection favours an optimal trait value that varies over time, in turn favouring new phenotypes that better match with the new optimum of fitness (Kopp & Matuszewski, 2014). When this variation is stochastic (rather than a single sudden or a gradual change of the optimum), the changing environment is expected to induce a continual change in selection direction and intensity, leading to fluctuations in the optimal phenotype over time (Estes & Arnold, 2007; Matuszewski *et al.*, 2014; Chevin *et al.*, 2017). Moving optimum models have emphasised the importance of stochastic environmental variation in the maintenance of phenotypic heterogeneity via genetic polymorphism (Bürger, 1999) or via mechanisms that help to mitigate environmental fluctuations such as phenotypic plasticity (Tufto, 2000; Chevin & Lande, 2015) and bet-hedging (Tufto, 2015). Yet, they also highlight that fluctuations in selective pressure can prevent natural populations from being perfectly adapted to their current environment (Eshel & Hamilton, 1984): perpetual fluctuations create a constant delay between the contemporary selective pressure and the response to this selection, ultimately leading to a fitness burden called a '*lag load*' (Figure 1)(Maynard Smith, 1976; Bürger & Lynch, 1995). Analytical predictions further demonstrate that, in a stochastic environment characterized by random fluctuations of the optimum phenotype, small populations are likely to experience a simultaneous decrease in their growth rate and increase in population size variance, both likely to increase extinction risk (Chevin *et al.*, 2017).

Despite the evolutionary importance of fluctuating selection, evidence of lag load's impact in wild populations is limited in the literature, and empirical demonstration for the consequences of it remains scarce. In one recent example, Bonnet & Postma (2018) demonstrated temporal

fluctuations in selection on body size of snow voles (*Chionomys nivalis*), although the mechanism for this selection was unclear. In another recent study, McAdam et al. (2019) demonstrated that annual changes in food availability imposed change in the optimum litter size for a North American red squirrel (*Tamiasciurus hudsonicus*) population. However, female squirrels were only partly able to plastically adjust their litter size to match the new optimum, resulting in an evolutionary lag load. More generally, two significant challenges in inferring both the fluctuations in selection and the resultant lag load exist. First, the long-term studies necessary for such analyses are relatively rare. Second, previous models have not disentangled changes in the trait distribution from changes in optimum fitness over time (but see Chevin *et al.*, 2015; Gamelon *et al.*, 2018). Both issues were recently tackled in an extensive analysis of breeding phenology across wild populations of mammals and birds (de Villemereuil *et al.*, 2020). Using appropriate modelling of the fitness distribution (Poisson or Zero-Inflated Poisson distributions), the authors revealed substantial fluctuations in selection and detected lag loads for the majority of populations.

Breeding phenology is a particularly interesting case for fluctuating selection (i.e., timing of breeding) since this trait is both tightly linked to environmental fluctuations among years and is an important component of annual fitness and population persistence (Gienapp & Visser, 2006; Chuine, 2010). In the aforementioned study of 39 bird and mammal populations (de Villemereuil *et al.*, 2020), the considerable annual variation in selection for breeding date was generally resolved by phenological plasticity. However, in endangered populations, such variation in selection is expected to have an exacerbated impact. First, threatened populations are often exposed to extreme environments (Howard *et al.*, 2020) and environmental degradation (e.g. climate change, fragmentation, pollution) which can rapidly exacerbate the impact of environmental variation (Boer, 2009). Moreover, endangered populations are often small and with low genetic variability, hence having limited potential to constantly adjust to a moving phenotypic optimum. In other words, endangered populations are less likely to adapt their plastic response rapidly enough to constantly track optimal phenotypes and to resolve the chronic lag load generated by fluctuating selection

(Maynard Smith, 1976). If the rate of environmental change becomes too rapid for the population to keep up, the lag load can increase and impact demography, ultimately leading to higher extinction risk (Maynard Smith, 1976; Lande & Shannon, 1994). In this context, assessing the fluctuations in the selection pressures operating on threatened populations and evaluating their response is therefore not only crucial to better predict their ability to adapt to their current and future environment (Wigley *et al.*, 1998; Boer, 2009; Chevin *et al.*, 2017), but also to develop appropriate management policies.

In this study, we explored the pattern of temporal variation in selection on lay date in two populations of the endangered hihi (*Notiomystis cincta*). Once spread across the North Island of Aotearoa New Zealand, this endemic bird now only occurs naturally in Te Hauturu-o-Toi sanctuary (Little Barrier Island; 36°120 S, 175°050 E; Figure 2). In the past fifty years, hihi have been reintroduced to pest-free sanctuaries, spread across Aotearoa New Zealand and representing a large range of habitats and climatic conditions (Figure 2). The Tiritiri Mātangi and Kārori populations (the two populations of interest) occur at the two extremes of this range (i.e., the sanctuaries are 525 kms apart), with hihi monitored since reintroduction and lay date and reproductive success consistently recorded for both populations. Previous studies have demonstrated that both populations contain low levels of genetic diversity (Brekke *et al.*, 2011; de Villemereuil *et al.*, 2019b), low levels of additive genetic variance and therefore a low adaptive potential (de Villemereuil *et al.*, 2019a; Rutschmann *et al.*, 2020; Bonnet *et al.*, 2022). This low ability to adapt for hihi is in contrast to other species where higher adaptive potential has enabled a shift toward earlier lay dates in response to climate change (Charmantier & Gienapp, 2014 ; for a New Zealand species see: Teplitsky *et al.*, 2010). For the hihi, some plasticity of lay date in response to temperature has been detected in the Tiritiri Mātangi population (de Villemereuil *et al.*, 2019b), but this is not enough to resolve the discrepancy between the optimal and average lay date. Therefore, hihi populations appear very unlikely to track chronic environmental fluctuations, potentially challenging population persistence in the future. To better understand whether selection on lay date fluctuates over time and to assess whether the patterns of selection might differ between the Tiritiri Mātangi and Kārori populations, we modelled variation of fitness over time in both

environments using a moving optimum model framework. More specifically, in both populations we assessed (i) how environmental variation affects the shape of natural selection, (ii) whether the resulting natural selection fluctuates over time and (iii) how comparable those features are between populations.

## **METHODS**

### ***a. Study populations and data collection***

Tiritiri Mātangi sanctuary is located on an offshore island in Tīkapa Moana / the Hauraki Gulf (36°36'8"S, 174°53'13"E), a sub-tropical climatic zone (Figure 2). Seventy-one birds were introduced from Te Hauturu-o-Toi over three translocation events in 1995, 1996 and 2010. No natural migration to or from the island has been observed. Since 2005-2006, up to 20% of the fledglings have been translocated from the population to other sanctuaries to artificially maintain carrying capacity and for conservation purposes. The Kārori population is located in Zealandia urban eco-sanctuary, in Wellington city (41°17'26"S, 174°45'10"E, oceanic climatic zone, Figure 2). The sanctuary is protected by a pest-proof fence and has been free of mammalian predators since 2000. Sixty-four hihi were reintroduced in 2005 from Tiritiri Mātangi and Pūkaha National Wildlife Centre. Subsequently, fifty-seven birds were reintroduced over six translocations between 2005 and 2012. No natural immigration has ever been recorded but birds may emigrate outside of the sanctuary (Brekke *et al.*, 2011; Ewen *et al.*, 2013).

In both sanctuaries, nesting attempts and reproductive success are recorded every season for all individuals in the population. Hihi reproduce during the austral spring and summer (late September to January). Within a season, females can lay multiple clutches of between three and five eggs, with large variation in fledgling success (from 0 to 100%). In both sanctuaries, birds are provided with supplementary food and nest boxes. In total, the Tiritiri Mātangi dataset includes 1,204 whole-season breeding attempts, from 804 unique females, over 23 consecutive reproductive seasons (1997 to

2019). For Kārori, the dataset contains 606 breeding events from 168 unique females, over 15 consecutive reproductive seasons (from 2005 to 2019).

For our analyses, fitness was considered as the total number of fledglings produced by a female over the entire reproductive season and can therefore include the output of one, two or three reproductive events a season (the success in second or third clutches being usually much lower than the reproductive success of the first clutch; see Table S.1 for detailed numbers). The start of the breeding season was chosen as the timing of the lay date of the first breeding attempt of the season, with earlier initial breeding increasing the chances of re-clutching. For all analyses, lay date has been centred using the mean across years and scaled to a within-year variance of one.

### ***b. Models of fluctuating selection***

To analyse the shape and intensity of selection in both populations, we compared the statistical fit of three fitness models as described in de Villemereuil *et al.* (2020). In each model,  $z$  represents the first breeding attempt of the season and fitness ( $W$ ) is computed as the total number of fledglings produced over the entire reproductive season.

The first model (“NoSel”) assumes no selection: for each female, the fitness  $W$  is estimated as the expected number of fledglings ( $a$ ) but does not depend on the lay date ( $z$ ):

$$\text{(Eqn 1)} \quad W = \exp(a) \quad \text{“NoSel”}$$

The second model (“Dir.”) assumes directional selection, i.e., the expected fitness  $W(z)$  depends on the number of fledglings ( $a$ ) and a linear function of slope  $b$  of the lay date:

$$\text{(Eqn 2)} \quad W(z) = \exp(a + bz) \quad \text{“Dir”}$$

The third model (“Opt.”) assumes the existence of a Gaussian-shaped optimum lay date, i.e., the expected fitness  $W(z)$  is estimated assuming an optimum lay date ( $\theta$ ) for which the number of fledglings is maximised ( $W_{max}$ ). The parameter  $\omega$  describes the width of the fitness function, with smaller  $\omega$  causing stronger stabilizing selection:

$$\text{(Eqn 3)} \quad W(z) = W_{max} \exp\left(\frac{-(z - \theta)^2}{2\omega^2}\right) \quad \text{“Opt”}$$

All models included random effects of female identity to account for repeated measures. Because previous studies found a strong age structure in the start of the breeding season for Tiritiri Mātangi population (i.e., young and old females tending to start breeding later; see de Villemereuil et al., 2019b), we also included an intercept-modifier effect into the models based on three age categories: reproductive females were considered either young (if age = 1), old (if age > 6) or middle-aged (otherwise).

We first assumed selection to be constant between seasons and only included female identity as a random effect on the intercept of the models (i.e.,  $a$  or  $W_{max}$ ). These models are referred to as “ConstDir” and “ConstOpt”. Then, we allowed seasonal fluctuations in fitness parameters of selection ( $b$  or  $\theta$ ) by adding a yearly random effect to them (“FluctDir” and “FluctOpt” models). Finally, temporal autocorrelation ( $\varphi$ ) was introduced between consecutive “seasonal” values for the slope  $b$  (“FluctCorrDir”) or the optimum  $\theta$  (“FluctCorrOpt”) as a first-order auto-regression (AR1) from each year to another (Eqn 4). In other words, the value of  $\theta$  for year  $t$  depended on the value of  $\theta$  for year  $t-1$ :

$$\text{(Eqn 4)} \quad \theta_t = \mu + \varphi \theta_{t-1} + \varepsilon$$

In contrast to the ConstDir and ConstOpt models that assume that selection changes randomly each year, the models with autocorrelation imply that selection is to some degree predictable from the selection in the previous year (i.e., suggesting some degree of temporal auto-correlation in the environment). The combination of fitness functions and patterns of fluctuations resulted in seven models per population (Table 1).

### **c. Statistical analyses**

Models were computed with the Hamiltonian Monte Carlo framework as provided in Stan (Hoffman & Gelman, 2014). We ran 10 chains, with 3000 iterations each including a 1000-iteration burn-in and

thinning every five iterations. In total, we obtained 4000 iterations for each model to ensure an effective sample size above 200 for all parameters. Convergence was checked with graphical inspection as well as using the potential scale reduction factor diagnostic (Vehtari *et al.*, 2019). Models were compared using an information criterion derived from a Leave-One-Out (LOO) procedure with Pareto smoothed importance sampling, from which we extracted LOOIC, which can be interpreted in a similar manner to a classical information criterion.

With a smaller number of females in Kārori (n=168) compared to Tiritiri Mātangi (n=804), we were expecting models may lack power in this population, notably to distinguish between selective regimes. To test this, we used a random subsampling approach. We subsampled the Tiritiri Mātangi lay date dataset (n=1,204) to match the Kārori sample size (n= 606), and tested the fit of all models to this dataset. We repeated this subsampling ten times, fitted each of the seven models of selection to each of the ten datasets, and calculated the average support for each of the models by computing the ratio of the summed support for each model over all subsamples to the total support for all subsamples (See Supplementary Information 2). Then we compared the average support for each of the selection models to those using the full Tiritiri Mātangi dataset to see whether a smaller sample size provided a consistent result. We also visually compared the predicted shape of selection for each of the ten subsampled dataset models, and notably inspected the shape of selection between the *ConstDir* (i.e., directional selection that does not vary across years) and *ConstOpt* (i.e., stabilizing selection that does not vary across years) models. Doing this allowed us to better assess the shape of selection and notably to look for the presence or absence of a clear optimum within the observed phenotypic range, in the *ConstOpt* models. Finally, the shape of these models was contrasted to the *ConstDir* and *ConstOpt* models for the full Kārori dataset.

## **RESULTS**

The observed first breeding event (i.e., lay date) was similar between populations, with a median of November 2<sup>nd</sup> (range October 20<sup>th</sup> to November 11<sup>th</sup>) for Tiritiri Mātangi, and October 31<sup>st</sup> (range



October 4<sup>th</sup> to December 1<sup>st</sup>) for Kārori. For both populations second and third clutches often continued into the late austral summer (See Figure S.1.a). The average annual number of fledglings per female was 2.73 ( $\pm$  standard error of 1.93) for Tiritiri Mātangi and 3.30 ( $\pm$  2.37) for Kārori (See Figure S.1.b). In Tiritiri Mātangi both intercept parameters for age (young or old females) were negative, suggesting lower fitness for immature (meta-estimate weighted across models: median -0.27, credible interval [-0.36; -0.18]) or older females (-0.37 [-0.61; -0.15]). By contrast, although they were also negative, age parameters were not significant in Kārori (young: -0.044 [-0.19; 0.048]; old: -0.51 [-1.52; 0.37]).

### **a. Selection shape and fluctuations**

To study the selection operating on lay date, we contrasted models of selection where fitness (i.e., the total number of fledglings produced over the entire reproductive season) was predicted by the first breeding event of the season. We did not find a unique model that best fitted to the Tiritiri Mātangi population dataset. Instead, we found the three models including an optimum in lay date (*FluctCorrOpt*, *ConstOpt*, *FluctOpt*) to be within a close  $\Delta$ LOOIC of each other ( $\Delta$ LOOIC < 2, Table 1). All other models had very weak support ( $\Delta$ LOOIC > 7) and were not considered. From the three top models, the presence of an optimum in lay date in Tiritiri Mātangi is well supported (Figure 3.a): estimators indicate a significantly negative optimum lay date (meta-estimate weighted across models for  $\theta$ : -1.95 [-2.51, -1.42]) and a sharp peak width ( $\omega$ : 2.63 [2.15, 3.17]). Therefore, the optimal lay date for Tiritiri Mātangi is earlier than the average lay date observed in the population.

Based on LOOIC, it is difficult to validate the presence or absence of fluctuations in the optimum lay date. Yet a close inspection of the lay date optimum ( $\theta$  varies between -2.18 and -1.14; also see Figure 3.a) and of the magnitude of fluctuations in the lay date optimum ( $\sigma_\theta$ : 0.43 [0.10, 0.83]; see Table S.3.a-c for tables of model estimates) clearly indicates that there are substantial lay date optimum fluctuations between years, suggesting the presence of fluctuating stabilizing selection. In other words, the variation among years in the optimal lay date amounts to 42% of the phenotypic variation within years, meaning that changes in the timing of the optimum are relatively strong compared to

the variation occurring between individuals. Finally, the correlation between  $\theta$  and the average lay date in the population was significant (0.47 [0.07, 0.79]; associated Bayesian posterior p-value = 0.05), suggesting that, although the population shifts their lay date in response to an earlier optimum, there is a rather constant mismatch between both parameters (also see online animated Figure).

For the Kārori population, the best supported model was a model with constant directional selection (Figure 3.b, Table 1.b), with no change in the strength of selection (*i.e.*, the slope of the relationship between lay date and fitness) across years. However, both the *ConstOpt* and *FluctCorrDir* models were within a close  $\Delta$ LOOIC (<2). Furthermore, all models (except *NoSel*) were within  $\Delta$ LOOIC range of 3.5 (Table 1.b) and were hard to dismiss completely. That is, in Kārori, our results provide mixed support on the selection shape and whether selection fluctuates over time. Yet, visual (Figures S.2.b:c) and parametrical (meta-estimate weighted across models for  $\theta$ : -8.17 [-25.9, -1.01];  $\omega$ : 7.70 [3.91, 13.31];  $\sigma_\theta$ : 0.42 [2.12e<sup>-4</sup>, 1.55]; see Tables S.3.d-f for all model estimates) inspections of the models with fluctuating optimums reveals that they all present characteristics of directional selection (very early lay date optimum  $\theta$ ; very wide peak width  $\omega$ ; non-significant fluctuations in optimum lay date). Finally, note that the slope  $b$  of all models suggesting directional selections are significantly negative ( $b$ : -0.15 [-0.22, -0.076]), suggesting that, similar to Tiritiri Mātangi, earlier lay dates lead to higher fitness in the Kārori population.

### ***b. Power analyses***

To determine whether the smaller Kārori dataset would be able to distinguish between different types of selection, we subsampled the Tiritiri Mātangi dataset to the same number of lay date records as the Kārori dataset. Across 10 replicate samples, we found comparable support for directional (42%) or stabilizing selection (41%), the model with no selection receiving the lowest support (16%) (See Table S.2). Thus, not all subsets of the data support stabilizing selection (*Opt.* models). However, across all sub-samples, optimum models consistently return a narrow  $\omega$  parameter and therefore exhibit bell-shape curves, which suggests an optimum. This clearly contrasts with the Kārori dataset

(Figures S.2.a:c), where both the directional selection and stabilising selection models suggests the absence of optimum (i.e., no peak), with earlier lay date consistently leading to higher fitness. These findings suggest that neither the smaller dataset in Kārori nor the potential resulting lack of power are solely driving our results.

## **DISCUSSION**

Although once widespread across Aotearoa New Zealand, the hihi became almost extinct in the wild by the 1880s. Due to successful reintroduction efforts, hihi have recovered from a single remnant source population and can now be found in sanctuaries across the North Island of New Zealand. Because of micro- or macro-geographical variation between sanctuaries, each population faces unique environmental and/or ecological conditions. As a result of this variation, we inferred different shapes of natural selection on a key fitness trait, lay date, between the two populations in this study, despite all individuals originating from the same remnant population. Our results highlight the importance of characterising selection for each population when considering the likely adaptive potential of a species across their range. They also have important implications for understanding the extinction risk faced by threatened species characterised by small population sizes.

### ***Variation in selection shape***

We found two different shapes of selection in the two populations. For the Tiritiri Mātangi population, in agreement with previous results on a smaller dataset (de Villemereuil *et al.*, 2019a, 2020), the best models all predicted the existence of a pronounced optimum lay date, implying a regime of stabilizing selection. This result suggests that in Tiritiri Mātangi, initiating reproduction too early or too late may drastically decrease fledglings' chances of survival, and hence maternal fitness. In contrast, results indicated an ambiguous support for directional or optimal models in Kārori. However, a close examination of the models with an optimum in Kārori reveals that the inferred shape is in fact largely directional (Figure S.2.c), such that earlier breeding may increase fitness. It is possible that, with a smaller number of individuals, results observed in Kārori may be the result of a lack of power to

accurately infer the shape of selection. Yet, this is not likely to be the case as our power analysis shows that the smaller sample-size alone cannot explain the selection trend observe in Kārori (See Supplementary Information 2). Despite differences in the shape of selection, earlier lay dates were associated with increased fitness for both populations (although in Tiritiri Mātangi the stabilizing selection suggests that being very early also leads to reduced fitness). It is therefore likely that, within physiological constraints (Dawson & Sharp 2007), early breeding allows birds to have multiple breeding attempts and hence increase their fitness. The absence of an optimum in Kārori is quite puzzling as it is highly unlikely that earlier reproduction is always better. The detection of directional selection in breeding phenology is often thought to be the result of competition for reproductive partners, nesting locations or resources, with healthier individuals reproducing earlier (Price *et al.*, 1988; Johansson & Jonzén, 2012). In Kārori sanctuary, intra-specific competition appears unlikely: natural emigration exists but is not common, and nesting competition does not seem to be a major issue as several nest boxes stay vacant each season and natural nesting remains rare. One possible explanation for directional selection in Kārori lay date lies in the existence of inter-specific competition with other native birds such as korimako (bellbirds, *Anthornis melanura*) and tūi (tui, *Prothemadera novaeseelandiae*) (Castro & Robertson, 1997). Although also present on Tiritiri Mātangi Island, co-occurrence with these species in Kārori could produce a harsher competitive environment, especially around supplementary food, sometimes monopolised by korimako or tūi (Castro & Robertson, 1997). In this context, 'less competitive' hihi may be suffering from the competition more than others (e.g., inexperienced or older individuals) and be forced to lay later. Only a closer examination of intra- and inter-specific interactions in Kārori would allow this hypothesis to be tested. A second non mutually exclusive hypothesis to explain directional selection could be that an evolutionary strategy evolved in response to trade-offs, shifting the evolutionary equilibrium to later lay dates (Price *et al.*, 1988). Such evolutionary displacement towards late reproduction has been demonstrated in different birds (under the effect of nutritional state, food availability, age or photoperiod; Visser *et al.*, 2010), including in the hihi for age (de Villemereuil *et al.*, 2019a) or variation in home range quality (Rutschmann *et al.*,

2020). Harsher conditions in Kārori (e.g., the habitat is on average 20% colder and 70% more humid than in Tiritiri Mātangi; see further details in Supplementary Information 4) could also prevent inexperienced birds or individuals in poor condition from tracking the environmental optimum (Salvante *et al.*, 2013). The absence of significant difference between young, mature and old individuals confirms that conditions may be stronger as no one is able to 'outplay' the others, however experienced or fit they are. Again, additional analysis of the relationships between lay date, individual quality and environmental conditions would be necessary to further explore this hypothesis.

### ***Variation in selection fluctuation and potential response to selection***

A second notable difference between the two populations lies in the fluctuations of the selection pattern over time. Despite the absence of a clear best model highlighting fluctuation in selection (i.e., *FluctCorrOpt*, *ConstOpt* and *FluctOpt* models have similar support), the parameters of the *FluctCorrOpt* model for Tiritiri Mātangi are indicative of strong fluctuations in the lay date optimum ( $\sigma_{\theta}$ : 0.43 [0.10, 0.83]). This amount of variation is relatively low compared to other bird populations ( $\sigma_{\theta}$ : 1.89 [0.33, 0.4.1] from a meta-analysis on 13 different species; de Villemereuil *et al.*, 2020) but are consistent with prior results on a more limited dataset for Tiritiri Mātangi population (de Villemereuil *et al.*, 2020).

In our previous study, we demonstrated that the Tiritiri Mātangi population shows an important discrepancy between the average lay date and its optimal value (de Villemereuil *et al.*, 2019a), indicating a maladaptive phenology and potentially a fitness burden. In the same study, we also demonstrated that this discrepancy was present despite the existence of a strong plastic response of lay date to environmental temperatures (average temperature 50 days before the mean of lay date over the years). Here, our results indicate that the population may suffer from a second fitness burden, this time resulting from the lag load generated by annual fluctuations in selection. Considering the low levels of genetic variation in hihi populations (de Villemereuil *et al.*, 2019b), this evolutionary load could have the potential to exacerbate (or reduce) the discrepancy between the optimal lay date and

the observed one. In this example, phenotypic plasticity seems to be “calibrated” enough to partly track the fluctuations of the optimum (i.e., as indicated by the significant and positive correlation between  $\theta$  and the average lay date: 0.47 [0.07, 0.79]), reducing the strength of the lag load and consequently, the negative impact on the population. On the other hand, the ability of this plastic response to track the optimum means that in years when the optimum lay date is later, phenotypic plasticity tends to be maladaptive by further delaying the lay date. In such a situation, a non-plastic response could potentially increase the fitness (Duputié *et al.*, 2015). [To further understand the demographic consequences of the lag load in threatened populations, our work could be extended in the future to evaluate the impact of the discrepancy between the average and the optimal lay date for each year on the population growth rate.](#)

In contrast to Tiritiri Mātangi, and despite the absence of a unique best model, it appears that the shape of the selection was constant over time in Kārori. This result suggests a rather constant environment, imposing less variation in natural selection. However, our results highlight a larger variation in start-date of the breeding season in Kārori (ranging from October 4<sup>th</sup> to December 1<sup>st</sup>) compared to Tiritiri Mātangi (October 20<sup>th</sup> - November 11<sup>th</sup>), which suggests higher levels of plasticity in the phenology in Kārori. In a complementary analysis ([Supplementary Information 5](#)), we found no support for plasticity based on the thermal cues found to correlate with lay date in Tiritiri Mātangi (de Villemereuil *et al.*, 2019a). These contrasting results (i.e., an absence of fluctuations in selection, yet the presence of phenotypic plasticity, unrelated to the well-established phenological temperature cue in birds) are unexpected and remain difficult to interpret. One possible explanation is linked to the species’ reintroduction history. The remnant hihi population on Te Hauturu-o-Toi Island is located only 46 km away from Tiritiri Mātangi. Therefore, the climatic conditions experienced by birds on Tiritiri Mātangi are very likely to be within the range experienced by their ancestors, and the population may still be able to track environmental fluctuations based on ancestral plastic reaction norms. However, in Kārori, reaction norms based on climatic fluctuations experienced on Te Hauturu-o-Toi (or Tiritiri Mātangi) are likely to be unsuitable, as Kārori’s environment is significantly colder and wetter

(temperature: Kārori -  $14.49 \pm 0.71^{\circ}\text{C}$  vs Tiritiri Mātangi -  $17.21 \pm 0.46^{\circ}\text{C}$ ; rainfall: Kārori -  $131.73 \pm 61.25$  mm vs Tiritiri Mātangi -  $78.37 \pm 25.06$  mm; See Figure 1, Figure S.4 and Supplementary Information 4 for more information). To say it differently, the temperature cues hihi experience in Kārori are potentially too far removed from the conditions they initially evolved in, and the behavioural responses exhibited by hihi in this population may therefore be disconnected from the optimum in their new environment. This hypothesis is highly speculative but has the advantage to explain why both patterns of selection and plasticity differ between populations.

### ***Conclusions***

Our results highlight the difficulty and consolidate the importance of characterising the shape and variation in selection for both populations, as inferences from one of them are unlikely to represent the selection pressures experienced by the whole species. In threatened species (which are likely to be highly fragmented, to present low adaptive potential and have been established from multiple reintroductions events), such fine-scale characterisation will be particularly important in determining local management strategies to help mitigate the impacts of climate change and other rapid changes to the environment. Further, characterising selection pressures operating on all current populations may enable assessment of the suitability of different local environments, which can help inform translocation strategies. We also demonstrate that, even with a relatively small sample size (which is often the case for threatened species), it is still possible to infer patterns in selection pressures in wild populations. We therefore encourage other studies to evaluate the fate of more than one population (even small) to better understand the capacity of the species to mitigate future fluctuations in their environment.

### **REFERENCES**

- Boer, G.J. 2009. Changes in interannual variability and decadal potential predictability under global warming. *J. Clim.* **22**: 3098–3109.
- Bonnet, T., Morrissey, M.B., de Villemereuil, P., Alberts, S.C., Arcese, P., Bailey, L.D., *et al.* 2022. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals.

*Science* (80-. ). **376**: 1012–1016.

- Bonnet, T. & Postma, E. 2018. Fluctuating selection and its (elusive) evolutionary consequences in a wild rodent population. *J. Evol. Biol.* **31**: 572–586.
- Brekke, P., Bennett, P.M., Santure, A.W. & Ewen, J.G. 2011. High genetic diversity in the remnant island population of hihi and the genetic consequences of re-introduction. *Mol. Ecol.* **20**: 29–45.
- Bürger, R. 1999. Evolution of genetic variability and the advantage of sex and recombination in changing environments. *Genetics* **153**: 1055–1069.
- Bürger, R. & Lynch, M. 1995. Evolution and extinction in a changing environment : a quantitative-genetic analysis. *Evolution (N. Y.)* **49**: 151–163.
- Castro, I. & Robertson, A.W. 1997. Honeyeaters and the New Zealand forest Flora: The utilisation and profitability of small flowers. *N. Z. J. Ecol.* **21**: 169–179.
- Charmantier, A. & Gienapp, P. 2014. Climate change and timing of avian breeding and migration: Evolutionary versus plastic changes. *Evol. Appl.* **7**: 15–28.
- Chevin, L., Visser, M.E. & Tufto, J. 2015. Estimating the variation , autocorrelation , and environmental sensitivity of phenotypic selection. *Evolution (N. Y.)* **69**: 2319–2332.
- Chevin, L.M., Cotto, O. & Ashander, J. 2017. Stochastic evolutionary demography under a fluctuating optimum phenotype. *Am. Nat.* **190**: 786–802.
- Chevin, L.M. & Lande, R. 2015. Evolution of environmental cues for phenotypic plasticity. *Evolution (N. Y.)* **69**: 2767–2775.
- Chaine, I. 2010. Why does phenology drive species distribution? *Philos. Trans. R. Soc. B Biol. Sci.* **365**: 3149–3160.
- Darwin, C. 1859. *The Origin of Species by Means of Natural Selection* (Murray, ed). London.
- de Villemereuil, P., Charmantier, A., Arlt, D., Bize, P., Brekke, P., Brouwer, L., *et al.* 2020. Fluctuating optimum and temporally variable selection on breeding date in birds and mammals. *Proc. Natl. Acad. Sci.* **117**: 31969–31978.
- de Villemereuil, P., Rutschmann, A., Ewen, J.G., Santure, A.W. & Brekke, P. 2019a. Can threatened species adapt in a restored habitat? No expected evolutionary response in lay date for the New Zealand hihi. *Evol. Appl.* **12**: 482–497.
- de Villemereuil, P., Rutschmann, A., Lee, K.D., Ewen, J.G., Brekke, P. & Santure, A.W. 2019b. Little adaptive potential in a threatened passerine bird. *Curr. Biol.* **29**: 889-894.e3. Elsevier Ltd.
- Duputié, A., Rutschmann, A., Ronce, O. & Chaine, I. 2015. Phenological plasticity will not help all species adapt to climate change. *Glob. Chang. Biol.* **21**: 3062–3073.
- Eshel, I. & Hamilton, W.D. 1984. Parent-Offspring Correlation in Fitness under fluctuating Selection. *Proc. Natl. Acad. Sci. U. S. A.* **222**: 1–14.
- Estes, S. & Arnold, S.J. 2007. Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *Am. Nat.* **169**: 227–244.
- Ewen, J.G., Renwick, R., Adams, L., Armstrong, D.P. & Parker, K.A. 2013. 1980-2012: 32 years of reintroduction efforts of the hihi (stitchbird) in New Zealand. In: *Global Re-introduction Perspectives: 2013. Further case studies from around the globe* (P. S. Soorae, ed), pp. 68–73. Abu Dhabi, UAE.



- Gamelon, M., Tufto, J., Nilsson, A.L.K., Jerstad, K., Røstad, O.W., Stenseth, N.C., *et al.* 2018. Environmental drivers of varying selective optima in a small passerine: a multivariate, multiepisodic approach. *Evolution (N. Y.)* **72**: 2325–2342.
- Gienapp, P. & Visser, M.E. 2006. Possible fitness consequences of experimentally advanced laying dates in Great Tits: Differences between populations in different habitats. *Funct. Ecol.* **20**: 180–185.
- Hoffman, M.D. & Gelman, A. 2014. The No-U-Turn Sampler : Adaptively setting path lengths in Hamiltonian Monte Carlo. **15**: 1593–1623.
- Howard, C., Flather, C.H. & Stephens, P.A. 2020. A global assessment of the drivers of threatened terrestrial species richness. *Nat. Commun.* **11**: 1–10. Springer US.
- Johansson, J. & Jonzén, N. 2012. Game theory sheds new light on ecological responses to current climate change when phenology is historically mismatched. *Ecol. Lett.* **15**: 881–888.
- Kopp, M. & Matuszewski, S. 2014. Rapid evolution of quantitative traits: Theoretical perspectives. *Evol. Appl.* **7**: 169–191.
- Lande, R. & Shannon, S. 1994. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution (N. Y.)* **50**: 434–437.
- Matuszewski, S., Hermisson, J. & Kopp, M. 2014. Fisher’s geometric model with a moving optimum. *Evolution (N. Y.)* **68**: 2571–2588.
- Maynard Smith, J. 1976. What determines the rate of evolution? *Am. Nat.* **110**: 331–338.
- McAdam, A.G., Boutin, S., Dantzer, B. & Lane, J.E. 2019. Seed masting causes fluctuations in optimum litter size and lag load in a seed predator. *Am. Nat.* **194**: 574–589.
- Price, T., Kirkpatrick, M. & Arnold, S.J. 1988. Directional selection and the evolution of breeding date in birds. *Science (80- )*. **240**: 798–799.
- Rutschmann, A., de Villemeireuil, P., Brekke, P., Ewen, J.G., Anderson, N. & Santure, A.W. 2020. Consequences of space sharing on individual phenotypes in the New Zealand hihi. *Evol. Ecol.* **34**: 821–839. Springer International Publishing.
- Salvante, K.G., Dawson, A., Aldredge, R.A., Sharp, P.J. & Sockman, K.W. 2013. Prior experience with photostimulation enhances photo-induced reproductive response in female house finches. *J. Biol. Rhythms* **28**: 38–50.
- Teplitsky, C., Mills, J.A., Yarrall, J.W. & Merilä, J. 2010. Indirect genetic effects in a sex-limited trait: The case of breeding time in red-billed gulls. *J. Evol. Biol.* **23**: 935–944.
- Tufto, J. 2015. Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution (N. Y.)* **69**: 2034–2049.
- Tufto, J. 2000. The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. *Am. Nat.* **156**: 121–130.
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. & Bürkner, P.C. 2019. Rank-normalization, folding, and localization: An improved Rb for assessing convergence of MCMC. *arXiv* 1–27.
- Visser, M.E., Caro, S.P., Oers, K. Van, Schaper, S. V. & Helm, B. 2010. Phenology, seasonal timing and circannual rhythms: Towards a unified framework. *Philos. Trans. R. Soc. B Biol. Sci.* **365**: 3113–3127.

Wigley, T., Smith, R. & Santer, B. 1998. Anthropogenic influence on the autocorrelation structure of hemispheric-mean temperatures. *Science* (80-. ). **282**: 1676–1679.

## **FIGURE CAPTIONS**

**Figure 1. The lag-load concept.** **A.** In a stable environment, the year-to-year fluctuations in the optimum of fitness  $\theta$  are relatively low. As a consequence, natural populations can exhibit phenotypes (yellow distribution) reasonably well adapted to their environment, and the average phenotype of the population ( $\bar{z}$ ) is close to  $\theta$ . The difference between the maximum absolute fitness  $W_{max}$  and the mean fitness of the population ( $W$ ) is small. **B.** In a stochastic and unpredictable environment, fluctuations in the optimum of fitness  $\theta$  are more important and it is more difficult for natural populations to be perfectly adapted to their environment: each year, the direction and intensity of selection changes. This fluctuating selection can increase the gap between  $\bar{z}$  and  $\theta$  and therefore increased the difference between  $W_{max}$  and  $W$ . This phenomenon creates a fitness burden called the 'lag load'. Figure modified from de Villemereuil et al., (2020).

**Figure 2. Hihi populations across the North Island of Aotearoa New Zealand.** Te Hauturu-o-Toi is the remnant population. Shape and fluctuations of natural selection were estimated in the Tiritiri Mātangi and Kārori populations, the six other reintroduced populations (grey circles) have insufficient data to be included in this study. Average climatic conditions are provided for both populations.

**Figure 3. The best fitness models for lay date in the two hihi populations:** (a) stabilising selection with a fluctuating optimum for Tiritiri Mātangi island (each colour represents a different year, earliest years of the survey being in blue, latest ones in yellow) and (b) directional selection towards earlier lay date for Kārori, with best support for a single model across years.

Figure 1

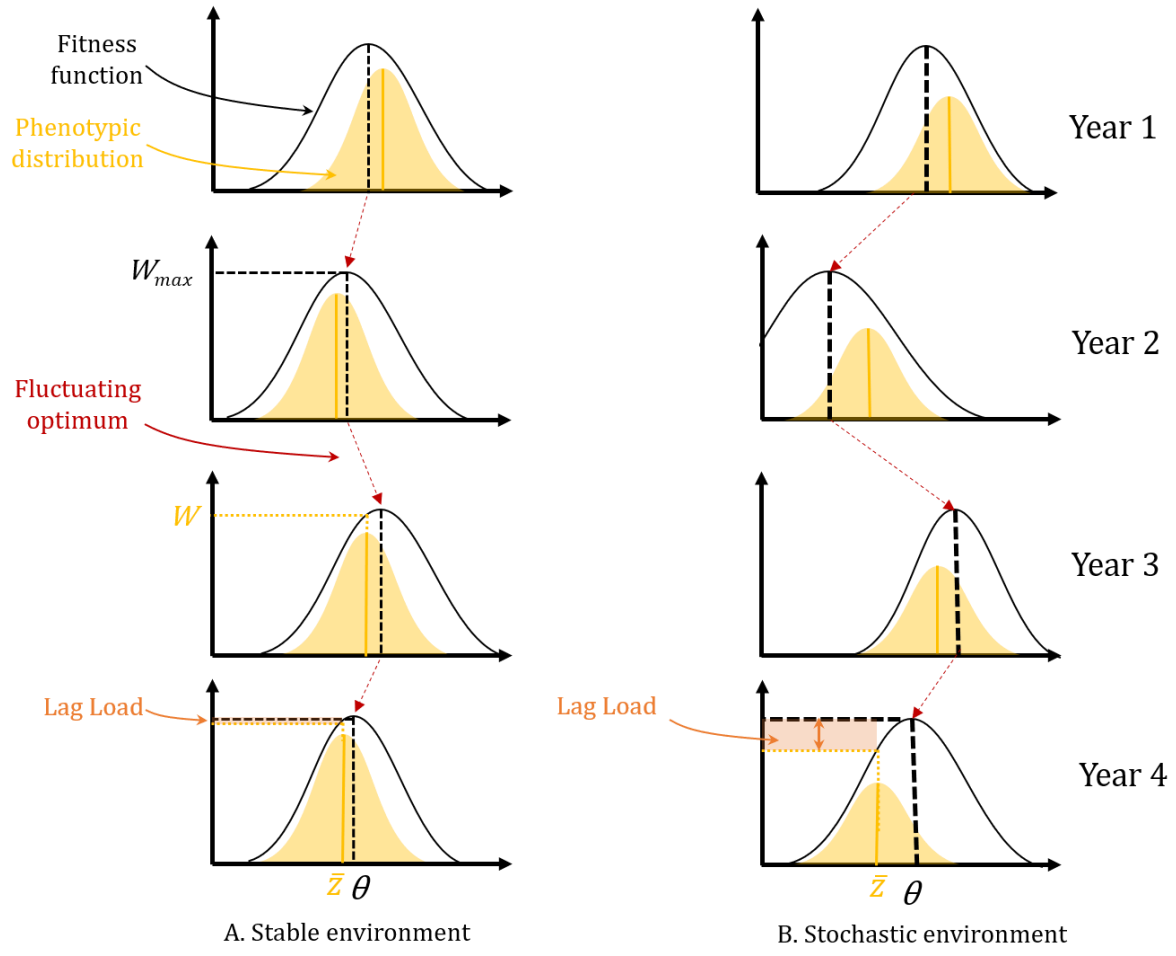
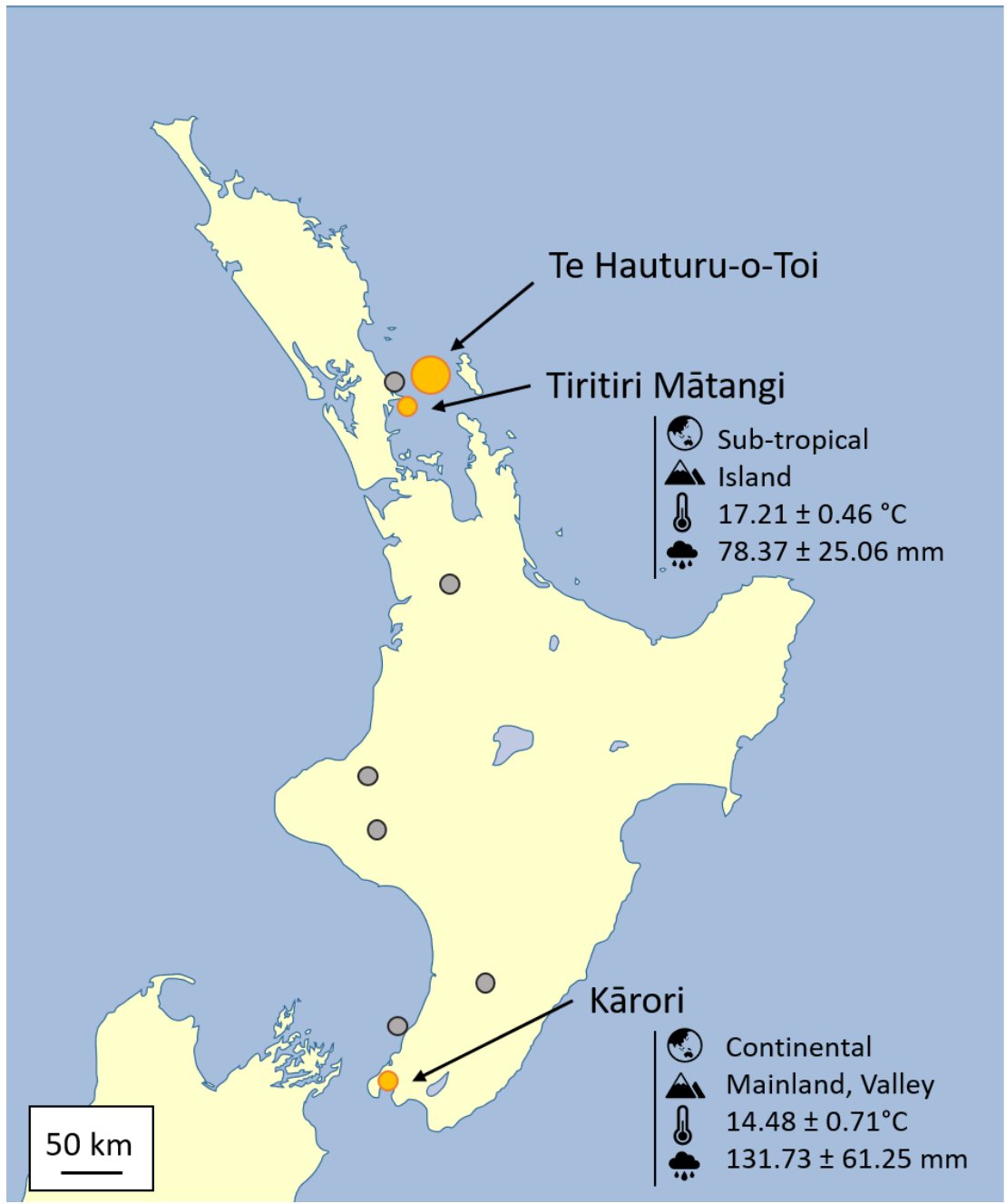


Figure 2



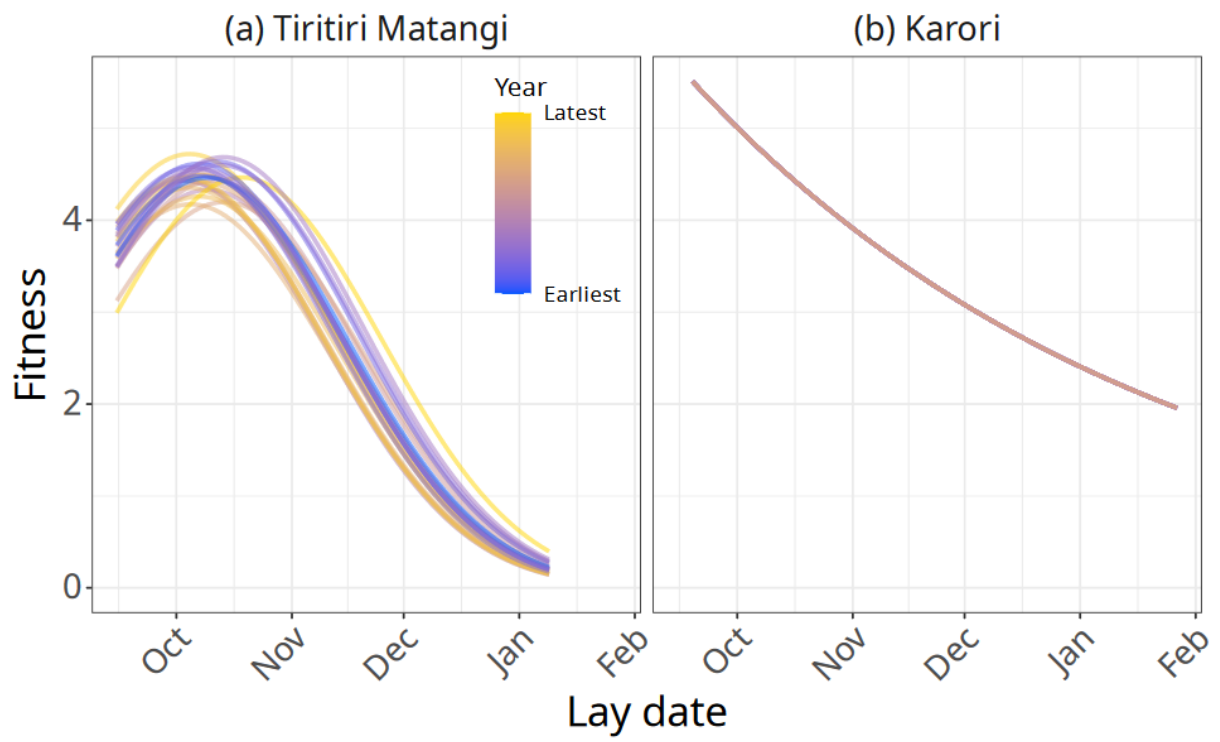


Figure 3

**Table 1: Fitness models for lay date.** Seven models, with three shapes of selection for lay date (None (*NoSel*), directional (*Dir*), stabilising (*Opt*)), are compared. Each selection shape was modelled as constant (*Const*), fluctuating (*Fluct*) and with temporal correlation in the fluctuations (*Corr*). Information content for each model is assessed via a leave-one-out procedure (LOOIC) and models are compared with  $\Delta$ LOOIC.

	Population	Model	LOOIC	$\Delta$ LOOIC
a.	<b>Tiritiri Mātangi</b>	FluctCorrOpt	4134.74	0
		ConstOpt	4135.63	0.89
		FluctOpt	4136.34	1.6
		FluctCorrDir	4142.16	7.42
		FluctDir	4142.27	7.53
		ConstDir	4152.11	17.38
		NoSel	4166.2	31.47
b.	<b>Kārori</b>	<i>ConstDir</i>	1091.44	0
		ConstOpt	1092.97	1.53
		FluctCorrDir	1093.16	1.72
		FluctDir	1093.56	2.12
		FluctCorrOpt	1093.98	2.55
		FluctOpt	1094.85	3.41
		NoSel	1104.71	13.27