

Colour polymorphism torn apart by opposing positive frequency-dependent selection, yet maintained in space

Swanne P. Gordon^{1*}, Hanna Kokko², Bibiana Rojas¹, Ossi Nokelainen^{1,3} and Johanna Mappes¹

¹Centre of Excellence in Biological Interactions, Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland; ²Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland; and ³Department of Zoology, University of Cambridge, Cambridge, UK

Summary

1. Polymorphic warning signals in aposematic species are enigmatic because predator learning and discrimination should select for the most common coloration, resulting in positive frequency-dependent survival selection.
2. Here, we investigated whether differential mating success could create sufficiently strong negative frequency-dependent selection for rare morphs to explain polymorphic (white and yellow) warning coloration in male wood tiger moths (*Parasemia plantaginis*).
3. We conducted an experiment in semi-natural conditions where we estimated mating success for both white and yellow male moths under three different morph frequencies.
4. Contrary to expectations, mating success was positively frequency-dependent: white morph males had high relative fitness when common, likewise yellow morph males had high relative fitness when instead they were common. We hence built a model parameterized with our data to examine whether polymorphism can be maintained despite two sources of positive frequency dependence. The model includes known spatial variation in the survival advantage enjoyed by the yellow morph and assumes that relative mating success follows our experimentally derived values. It predicts that polymorphism is possible under migration for up to approximately 20% exchange of individuals between subpopulations in each generation.
5. Our results suggest that differential mating success combined with spatial variation in predator communities may operate as a selection mosaic that prevents complete fixation of either morph.

Key-words: aposematism, coloration, mating success, modelling, predation, sexual selection, spatial mosaic

Introduction

A fundamental question in evolutionary biology is what processes drive the origin and maintenance of genetic polymorphisms in the wild. Polymorphisms indicate the potential for unusual types of selection because they deviate from simpler cases where one genotype has a consistent advantage over another and natural selection drives the winning type to fixation (Calsbeek, Bonvini & Cox 2010). If a genotype's relative fitness improves with its relative frequency (positive frequency dependence), polymorphisms typically cannot be explained: here selection acts

against rare alleles, and the population is consequently expected to evolve towards one of alternative stable states but not towards a stable polymorphism (Mallet & Joron 1999; Lehtonen & Kokko 2012). Negative frequency dependence on the other hand has the opposite effect: a genotype is selected against when common and selected for when rare, making it easier to maintain polymorphisms (Sinervo & Lively 1996). Gene flow, that links populations undergoing spatially heterogeneous selection, can also make alternative morphs persist within an environment: in this case, a locally suboptimally performing morph can avoid extinction because of continual immigration from elsewhere, where it is selected for (divergent selection, Gray & McKinnon 2007).

*Correspondence author. E-mail: swanne.gordon@jyu.fi

Being easily measured and tracked, animal colours represent some of the best-studied examples of trait polymorphisms (McKinnon & Pierotti 2010). Colours are often strongly linked to fitness in many different contexts such as crypsis (Steward 1977; Rothschild 1981; Endler & Greenwood 1988) and thermoregulation (Clusella-Trullas, Van Wyk & Spotila 2007; Hegna *et al.* 2013). Colour is also often used a warning signal of unprofitability (Poulton 1890; Cott 1940; Stevens & Ruxton 2012), and in this context, theoretical expectations (Servedio 2000; Endler & Mappes 2004) as well as empirical studies (review, Allen 1988; Mallet & Barton 1989; Borer *et al.* 2010) point to positive frequency-dependent survival selection. It is easier for predators to learn to discriminate against signals that are not only highly recognizable and memorable, but also common ('strength in numbers'); this selects against polymorphism (Müller 1878; Mallet & Joron 1999; Lindström *et al.* 2001; Rowland *et al.* 2007). This renders it enigmatic why numerous polymorphic aposematic colour signals from a range of taxa are known in nature (e.g. O'Donald & Majerus 1984; Nokelainen *et al.* 2012; Rojas & Endler 2013).

Recent research has investigated the effect of various processes that can maintain polymorphisms and hence also impact prey fitness. In many cases, these mechanisms are not mutually exclusive and instead work in concert to achieve polymorphisms (Sánchez-Guillén *et al.* 2011). These range from spatio-temporal variation in predation (Endler & Rojas 2009; Stevens & Ruxton 2012; Nokelainen *et al.* 2014) or sexual selection (Maan & Cummings 2009; Nokelainen *et al.* 2012) to frequency-dependent selection (Svensson, Abbott & Hardling 2005; Olendorf *et al.* 2006). Gene flow (Rosenblum 2006) and genetic drift (Hoffman *et al.* 2006; Gray & McKinnon 2007) add to the mix of effects that can impact the stability of polymorphisms.

Research on colour polymorphisms in aposematic species has focused on the role of predator learning and signal evolution, due to the strong selection exerted by the high cost of conspicuousness when predators are naive (Lindström *et al.* 2001; Mappes *et al.* 2014). Here, we show that full understanding into the maintenance of warning colour polymorphisms may require considering multiple forces and their interaction. Our particular interest lies in understanding potential trade-offs between natural and sexual selection (Kotiaho *et al.* 1998; Maan & Cummings 2009; Nokelainen *et al.* 2012; Cummings & Crothers 2013; Finkbeiner, Briscoe & Reed 2014) and in the arguably understudied interaction of frequency-dependent selection and gene flow (Joron & Iwasa 2005; Calsbeek, Bonvini & Cox 2010).

Colour in aposematic organisms may serve multiple functions: thermoregulation, defence against predators and mate attraction. This highlights the possibility that both natural and sexual selection can be frequency-dependent (Gray & McKinnon 2007; Sinervo & Calsbeek 2006; Roulin & Bize 2007), and it is known from gen-

eral theory that when different types of frequency dependence interact, the diversity of dynamic outcomes can be far greater than predicted under one source of selection only (Sinervo & Calsbeek 2006). Here, we combine information from a mating experiment with published estimates of survival selection in order to examine their interaction and the potential role of frequency-dependent selection on the maintenance of colour polymorphism in an aposematic organism. The results aim to confirm an idea previously expressed by Roulin & Bize (2007): it is difficult to explain polymorphisms if the common morph derives a mating advantage, but in a spatially structured population, some degree of gene flow could potentially counteract the predicted loss of genetic variation. Interestingly, a modelling study of a particular colour polymorphism – Müllerian mimicry – shows that too high gene flow can again break down a polymorphism, because the system then effectively behaves like one unit with averaged parameter values (Joron & Iwasa 2005). We therefore also include a modelling component to investigate the range of dispersal values that could protect a polymorphism in a system where natural and sexual selection interact and also vary spatially.

Our study species is the aposematic wood tiger moth (*Parasemia plantaginis*). Males of this species exhibit discrete wing coloration on both local and on a broad geographical scale (Hegna, Galarza & Mappes 2015). European populations feature two distinct genetic male morphs, yellow and white (Galarza *et al.* 2014). In Finnish populations, the more conspicuous yellow male morph has been shown to have greater warning signal efficacy, such that viability selection mediated by predation favours this morph (Nokelainen *et al.* 2012, 2014). White adult males, on the other hand, sometimes appear to have better mating success than yellow males, especially in instances where mating incurs high costs (Nokelainen *et al.* 2012). However, mating success has so far only been evaluated in a context of equal frequency of both morphs, and under small laboratory conditions as opposed to more natural settings. To test for the effects of potential frequency dependence in the context of sexual selection, we conducted an experiment in semi-natural conditions where we estimated mating success under three different morph frequencies. We next integrate our data from this experiment with known estimates of predation against both morphs in a simple spatial model. The combination of both mating experiment and model is based on the suggested trade-off between natural and sexual selection described above driving the evolution and maintenance of colour polymorphism in our system. This trade-off is particularly interesting and unexplored in a case where both natural and sexual selection are potentially frequency-dependent in aposematic organisms, as never before has variation in frequency dependence across fitness components been considered in this context.

Materials and methods

STUDY SYSTEM

The wood tiger moth system (*Parasemia plantaginis*) is well studied in terms of the evolution of warning signals (e.g. Ojala, Lindström & Mappes 2007; Lindstedt, Lindström & Mappes 2009; Lindstedt *et al.* 2010; Nokelainen *et al.* 2012, 2014; Nokelainen, Lindstedt & Mappes 2013). Wood tiger moths are easy to rear and maintain in a laboratory setting and can be collected and manipulated with relative ease in semi-wild or wild conditions. They overwinter as larvae in the wild and naturally go through one generation a year per life cycle. Adult wood tiger moths are capital breeders and hence do not feed (but drink) during adulthood. Although laboratory-reared female-male ratio at eclosion is very close to the 50 : 50 ratio, in the field the operational sex-ratio is much more male biased (Gordon *et al.* In prep.). This is likely based on the fact that females in the laboratory live on average only a few days, whereas males can live up to two weeks (Santostefano & Mappes, In prep.), which likely reflect their life span differences in wild. Adult females are also conspicuous, avoided by bird predators, and, unlike the discrete coloration in males, exhibit a continuous range of wing coloration from yellow to orange and to red (Lindstedt *et al.* 2011). Like most moths, mature female moths emit a sexual pheromone call to attract males. Once fertilized, wood tiger moth females can lay approximately 300 eggs in one clutch (Ojala, Lindström & Mappes 2007). Females appear choosy, as males attempting to mate can be rejected if not preferred by females (pers. obs.). All wood tiger moths used in the enclosure experiment (see below) were laboratory-raised from stock collected from wild moths caught in Finland in 2010 and reared under greenhouse conditions at the University of Jyväskylä for multiple generations (for more details on laboratory rearing see Lindstedt, Lindström & Mappes 2008).

ENCLOSURE EXPERIMENT

In order to examine the effect of frequency-dependent selection on mating in wood tiger moths, we ran a large-scale outdoor enclosure experiment in Konnevesi research station of the University of Jyväskylä (Finland), from 4 July 2012 until 24 July 2012. A 20 m × 30 m (3 m high) enclosure was divided into six separate cage compartments, approximately 6 m × 10 m each. The floor of each compartment was open with natural foliage, and the top and two of the sides were made of white mesh, open to natural light. The other two sides were covered with a plastic green tarp to block visual contact between the cages and to limit the spread of mating pheromones exuded by the calling females.

We performed 10 individual runs, each involving three treatments: a balanced morph ratio (12 white males and 12 yellow males); a white-biased treatment (16 white males and eight yellow males); and a yellow-biased treatment (16 yellow males and eight white males). Every male and female used in the entire experiment was individually marked using a paint dot on the underside of both the top and bottom portion of the fore and hind wing. Each run used 24 virgin males and five virgin females per treatment (in order to closely mimic natural male-biased sex ratios in the wild), totalling 72 males and 15 females per run. Treatments within each run were randomly assigned to three of the six compartments, and all males and females used in each run were randomly assigned to their treatment. However, to achieve the

desired sample sizes, some unmated males had to be re-used in the next subsequent run. To see whether this logistical constraint had an effect, we analysed the mating success of each male categorized as naïve vs. re-used; we found no effect (for white male morphs: effect = -0.294 ± 0.436 , $z = -674$, $P = 0.500$; and for yellow morph: effect = 0.018 ± 0.460 , $z = 0.038$, $P = 0.970$). In total, we used 720 male measurements of which 590 involved naïve males and 130 were re-used.

The five females were tethered to a string in an open Styrofoam box (one female per box), which allowed them to move and fly but not escape the box. Behavioural assessments of the females prior to and during the experiment did not reveal any changes in behaviours that would limit them from escaping harassment from males if they so chose (Rojas *et al.* In prep.), and tethering them allowed us the opportunity to count all their eggs reliably. If free female choice was, however, impacted negatively by tethering, this impact should be spread equally across all three treatments.

The boxes, that offered protection from the wind and rain, were placed at equidistant points around the corners of every cage. Females were allowed to acclimatize for approximately one hour. Given past information on ideal mating times in these moths (S. Gordon pers. obs.), males were then released in the middle of each cage at approximately 4:00 pm. All moths were collected again at 8:00 am the next morning. The time window hence includes the time when matings naturally occur in this species (Nokelainen *et al.* 2012). Cages were watched from the beginning to the end of each run by one observer in each cage. Data about general hourly weather conditions, moth behaviours and the ID's of mated pairs upon the onset of mating were collected by observers. Because females may mate multiply, the total number of matings was not fixed. Males, however, were only found to mate maximally once, with a few exceptions: there were two double matings involving each one Y and one W male mated to same female, and one mating of the same white male to two females. As males were individually marked and identified after the onset of every mating, we could confidently assign paternity to all matings (except for the doubly mated pair, where we were able to assign paternity to the most likely father based on the coloration of the adult offspring for the basis of analyses). At the end of each experimental run, all mated females were brought to the greenhouse at the University of Jyväskylä and their egg number and hatching success were measured.

STATISTICAL ANALYSES

Male age and size

As females may choose males based on differences in age and size, we first calculated the spread of male characteristics across treatments to assess any bias the assignment of individuals into treatments may have produced. We used pupal weight as a good proxy for adult weight (as adults do not eat). We ran an ANOVA with age (days since pupation) and pupa weight as response variables and treatment by morph interaction as explanatory factors (using function 'aov' in Program R v2.15, 2013).

Mating success

We measured the mating success of the W and the Y morph using generalized linear mixed effect models, first combined, and then separately (one for each morph). In each model, the number

of matings in each cage was the response variable (binomial variable where 1 is mated and 0 means not mated in all cases). Pupa weight, male age and treatment (and their interaction) were the explanatory variables. We included enclosure cage as a random effect to account for the non-independence of matings within each enclosure.

Fitness

The relative fitness of individual males (estimated morph siring success) was calculated as simply a measure of the proportion of larval recruits (out of the total number of enclosure recruits) that were sired by each male (i.e. proportional to its contribution to the next generation) relative to the total number of males in the focal enclosure. It was obtained by dividing the number of hatched larvae sired by a given male by the total number of hatched larvae in each enclosure. This led to every male in each of the 30 unique enclosures having a measure of relative fitness. We then again ran a binomial generalized linear mixed effect model (morphs were not also separated as above because here there were no significant three-way interactions) where fitness was the response variable. Morph frequency (treatment), morph and their interaction were the explanatory variables, age and pupa weight were added covariates, and enclosure compartment was a random effect.

Results

EFFECT OF MALE AGE AND SIZES

Yellow males used in the experiment tended to be younger than white males (ANOVA $F_{1,701} = 14.610$, $P < 0.001$), yet this effect was not different between treatments ($F_{2, 701} = 0.223$, $P = 0.800$). The same pattern was found for pupal weight, which was smaller for yellow males (ANOVA $F_{1,706} = 9.367$, $P = 0.002$), yet equally so for all treatments ($F_{2,706} = 0.273$, $P = 0.761$). Due to this, both male age and pupa weight were included as covariates in all analyses.

MATING SUCCESS

The model for mating success of white male moths showed a significant positive effect of weight (model effect size = 0.016 ± 0.006 , $z = 2.722$, $P = 0.006$) but none of age (effect size = 0.011 ± 0.044 , $z = 0.242$, $P = 0.809$) or treatment ($P > 0.7867$). The model for the mating success of yellow males likewise showed a significant effect of weight (effect size = 0.0323 ± 0.011 , $z = 0.011$, $P = 0.002$), treatment (driven by this morph's significantly lower mating success in the white-biased treatment, see below; effect size = 6.404 ± 3.181 , $z = 2.013$, $P = 0.044$) and a significant interaction between weight and treatment: yellow males had lower mating success in the white-biased treatment (no difference in the yellow-biased treatment), with large males suffering more from this effect (effect size = -0.038 ± 0.017 , $z = -2.191$, $P = 0.028$). Age was also an insignificant covariate in the model for mating success of yellow males (effect size = -0.040 ± 0.050 , $z = -0.803$, $P = 0.422$).

We should note that the model combining both morphs model showed comparable results for age and size, but had no overall effect of morph (effect size of White morph vs. Yellow morph = -4.640 ± 2.692 , $z = -1.728$, $P = 0.085$). The only significant interactions were a white-biased treatment by morph interaction (model effect size = 8.886 ± 3.995 , $z = 2.224$, $P = 0.026$); that is, the yellow morph had significantly lower mating success in the white-biased treatment, driven mainly by larger males doing worse in this regard (significant weight by white-biased treatment by yellow morph interaction: effect size = -0.052 ± 0.022 , $z = -2.403$, $P = 0.0163$). More importantly, however, a Type III Chi-square ANOVA shows an overall significant effect of the three-way interaction between morph, weight and treatment (Chi-sq = 5.85, $P = 0.05$), implying that morphs responded differently to all aspects we were evaluating (treatment and weight). Therefore, for greater clarity, we will only focus on the separated morph model results in the discussion.

FITNESS

Fitness, which consider the mating success as well as the ultimate offspring sired hatching success, show that the yellow male morph have significantly lower fitness in the white-biased treatment compared to the other treatments (Fig. 1). This result is opposite for the white male morph. Analyses found no significant differences between treatments ($t = -1.318$, $P = 0.188$), but a significant treatment \times morph interaction ($t = 2.214$, $P = 0.027$). Age was again insignificant in the model ($t = -0.206$, $P = 0.837$), whereas weight was significant ($t = 3.195$, $P = 0.002$).

In order to better examine whether the differences between morphs in both high-frequency treatments were significant, we ran a similar model but excluded the balanced ratio treatment. Results show that in the white-biased treatment, yellow males have marginally lower fitness compared to white males (-1.408 ± 0.792 , $t = -1.777$, $P = 0.076$), while this pattern was reversed in the yellow-biased treatment (2.292 ± 0.998 , $t = 2.295$, $P = 0.022$).

MODELLING THE CONSEQUENCES

Our results surprisingly indicate *positively* frequency-dependent mating success: males of the white morph had high relative fitness when common, likewise yellow males had high relative fitness when they formed the majority of the males (Fig. 2). As both natural and sexual selection therefore appear positively frequency-dependent, the most straightforward theoretical prediction is that a population should evolve towards one of two possible alternative equilibria (Lehtonen & Kokko 2012): either white only or yellow only. The existence of large geographic areas with polymorphisms therefore led us to consider the role of spatially varying selection, for which there is evidence based on variation in bird communities (Nokelainen *et al.*

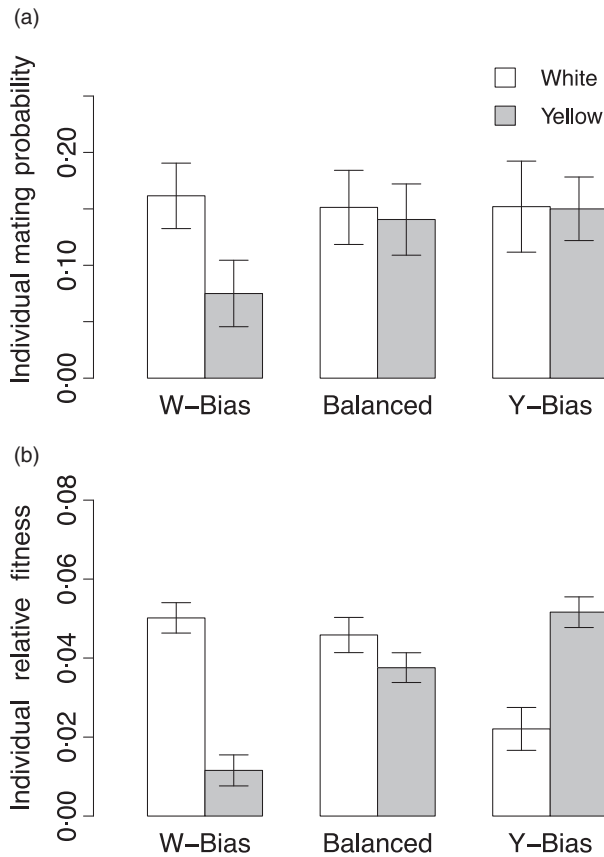


Fig. 1. Graph showing in panel a: the individual mating probabilities of both white (clear bars) and yellow (grey bars) male morphs in all three treatments white-biased (16 white males and 8 yellow males), balanced ratios (12 each white and yellow males) and yellow-biased (16 yellow males and 8 white males). Panel b shows the individual relative mating fitness of each morph in the same treatments defined as the proportion of hatched (laboratory reared) larvae that were sired by a given male relative to the total number of males in the focal enclosure. Error bars indicate standard errors.

2014). In a different context that also has spatial variation in predation-related processes (Müllerian mimicry), it has been shown that dispersal that is not too strong can produce cases of coexistence even if locally acting selection is positively frequency-dependent (Joron & Iwasa 2005); too high dispersal, however, will destroy the polymorphism. We therefore examine if this idea could also promote coexistence in our case where natural and sexual selection interact.

NATURAL VARIATION IN MORPH SURVIVAL AGAINST PREDATION

The features of our model are mainly based on an experiment examining predation upon different male colour morphs in wild (follows data published in Nokelainen *et al.* 2014). Therein, artificial prey models resembling the two male colour morphs (white and yellow) were crafted from waterproof-paper wings and plasticine bodies. The

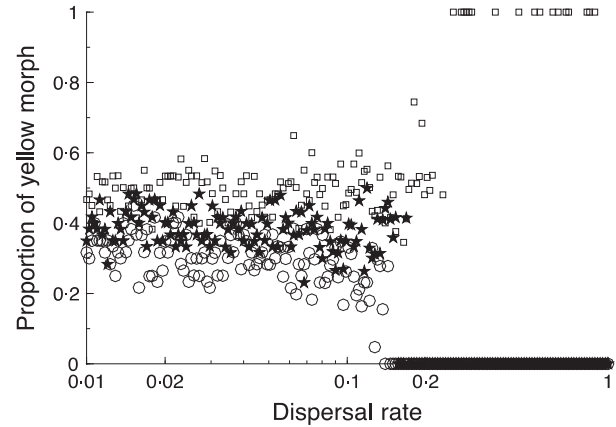


Fig. 2. The total frequency of yellow morph males in each simulation run at generation 1000, for 200 different log-spaced values for dispersal between 0.01 and 1, each computed three times [see options (i), (ii) and (iii) in main text]: first with a white-favouring interpretation of our data set (open circles), then a yellow-favouring interpretation (open squares) and a neutral interpretation (filled stars).

resemblance between real moths and artificial prey specimens was confirmed using image manipulation software and spectrophotometry. In the wild, specimens were put on green cardboard platforms to be preyed upon and placed along nonlinear transects. Bird attacks (indicated by diagnostic beak markings) were observed for five days in five geographical populations (spanning 60 sites) with known varying natural colour morph frequencies (Nokelainen *et al.* 2014). All possible avian predators were counted from experimental transects (see details from Nokelainen *et al.* 2014). However, only birds within 25 m of each transect (to emphasize the most potential predators in the close proximity) and passerine species (non-passerines such as hawks and waterfowl are not likely predators for moths) were included. Principal component analyses, that grouped birds by feeding tactics, were used then used to describe variability in predator community and were linked to morph survival. The results showed that *Parus*-species-dominated communities tend to attack yellow males less, with the opposite pattern being true in dunnock- and starling-dominated communities that tended to attack white males less (Nokelainen *et al.* 2014).

It is plausible that species differences in predators' responses to the morphs arise from species-specific feeding behaviour and also from physical properties of the habitats the species prefer. Although somewhat speculative, these differences could easily impact the efficiency of warning signals in each visual environment. For example, tits (*Paridae*), who tend to avoid yellow male moths, forage in bushes where yellow, being more contrasting in variable backgrounds, may be a better warning colour than white (Nokelainen *et al.* 2012). Indeed, previous studies have shown that colours such as yellow are very effective warning signals against passerines and chicks, while shorter wavelength colours (with greater UV-re-

flectance) have been found to actually attract attacks by these predators (Lyytinen *et al.* 2001; Lyytinen, Lindström & Mappes 2004). On the other hand, dunnocks (Prunellidae) tend to avoid white male moths and generally forage closer to the ground level in shaded boreal forests. White coloration might be a more effective warning signal in these circumstances because they have much higher luminance values compared to the yellow morph (Galarza *et al.* 2014).

MODEL USING BOTH PREDATION AND MATING RESULTS

We consequently built our model assuming that the 60 sites sampled above represent the range of naturally occurring spatial variation in morph-specific viabilities, as estimated in Nokelainen *et al.* 2014. We considered a world consisting of 60 patches, each producing 50 offspring in each generation. We only tracked the dynamics of males, that is we assume that male offspring morph frequencies are proportional to the frequency with which males of each morph become sires (this only requires assuming that inheritance via females does not bias patterns of inheritance in either direction).

Given that differential predation based on wing colour occurs during the adult time period, morph ratios are not necessarily constant and the model needs to track morph frequencies at the time when siring occurs. Having no precise information of this time dependency, we use the following logic to simulate the most likely pattern. Denote by b_i the survival advantage of the yellow morph in a current patch i (e.g. if site 13 has $b_{13} = 1.21$ then in this site a yellow individual's daily risk of dying is $1/1.21$, that is a fraction 0.89 of that of a white individual). Consequently, if a local population starts with w_0 whites and y_0 yellows (the subscript 0 denoting that no male has had the time to die yet), the probability that the next death targets a yellow male is $(y_0/b)/[(y_0/b)+w_0]$, and the complementary probability that the next removed male is white is $w_0/[(y_0/b)+w_0]$. Thus, if a uniformly distributed random number in the range $[0, 1]$ falls below $(y_0/b)/[(y_0/b)+w_0]$, we subtract one individual from the local y (thus $y_1 = y_0 - 1$, $w_1 = w_0$), and otherwise from the local w (thus $y_1 = y_0$, and $w_1 = w_0 - 1$). We repeat this procedure until no males of either morph are alive.

We then assume that females largely mate when mates are numerous (i.e. mostly when the season is not nearly over yet) and choose the mating time index of each female by rounding down a random number that is exponentially distributed with mean $(y_0+w_0)/10$. This result, denoted t , describes that a female mates at a point in time when t males have died and y_t and w_t are consequently still available for matings. The female mating time is thus designed to follow mate availability, and the factor 10 was set to make the proportion of females that attempt to mate when no males are alive negligibly low (for $y_0 + w_0 \approx 50$ which according to our assumptions is the approximate size of

local populations, this probability – i.e. the probability that the exponential distribution with mean $(y_0+w_0)/10$ produces a value that exceeds y_0+w_0 , the number of deaths it takes for all males to have died – is approximately 0.45×10^{-5}). We assume that these exceedingly rare late females in reality mate with the last available male, while noting that our results remain virtually unchanged if we instead assume that they completely fail to mate, as the siring success this late in the season contributes minimally to the mating success of either of the male morphs.

Populations were initiated such that 250 yellow and 250 white males were distributed randomly across the 60 patches, to yield initial patch-specific y_0 and w_0 values. The above procedure was then used within each subpopulation to determine siring times for 50 surviving offspring per patch (for those patches that had at least one male; we thus implicitly evoke density dependence: each occupied patch is equally productive). We equate survival with maturation.

The morph identity of each of the 50 sires was determined based on y_t and w_t at the siring time t , such that

$$\text{Prob \{offspring is yellow\}} = \frac{x}{x + (1-x)e^{ax+b}}.$$

Here, $x = y_t/(y_t + w_t)$ is the proportion of yellows in the current population of potential sires. The parameters a and b define the relative mating success advantage of whites, determined by a least-squares regression of $\log(\frac{\text{mean eggs sired by white}}{\text{mean eggs sired by yellow}})$, values taken from the empirical part of this paper against the treatment x values ($x = 1/3, 1/2$ and $2/3$) used in the experiments. Because a few females mated once with a yellow and once with a white male, we computed three different values for a and b , to cover the entire range of uncertainty caused by this behaviour: (i) the first regression assumed that these females all contributed only to their white mate's success, (ii) the second assumed that these females all contributed only to their yellow mate's success, and (iii) the third assumed that these females had half their eggs sired by the white mate, and half by the yellow mate.

This procedure of determining offspring morph was repeated across all patches. We then assumed that a proportion d of offspring disperse, and others stay in their current patch. A dispersed offspring was assumed to land in any other than its natal patch. Dispersal completed a generation, and the entire procedure was then run for 1000 generations for a variety of different values of d , for each of the assumptions (i), (ii) and (iii).

MODEL RESULTS

We depict a large collection of single-run outcomes at generation 1000 rather than averaging over several replicates per parameter value. The latter approach would not be able to distinguish between a scenario with alternative non-polymorphic equilibria and a protected polymor-

phism (e.g. 0.5 could then, confusingly, mean either equally many 'yellow fix' and 'white fix' cases, or all runs having stabilized at a yellow-white polymorphism). Our approach shows that a polymorphism would never be maintained if the populations were assumed panmictic ($d = 1$) or, more generally, if the dispersal rate d exceeded approximately 0.2 (Fig. 2). The model then mostly predicts that the yellow morph disappears, but this conclusion proved sensitive to how we assumed paternity to be divided in those cases when a female mated multiply. If we made the yellow-favouring assumption, then sometimes the yellow morph fixed (Fig. 2), and polymorphism also prevailed at slightly higher values of d than in the other cases.

When dispersal was less frequent ($d < 0.2$), the model's behaviour was consistent across a large range of dispersal rates and also across the different assumptions regarding multiple mating. Both morphs persisted, with no clear trend other than that if a morph was assumed to gain more paternity with doubly mated females in our estimation of a and b , this morph was able to shift its frequency somewhat upwards from the scenario of equal paternity (the difference between the open and filled symbols in Fig. 2). Thus, we predicted an approximate 1 : 1 morph ratio when we made the yellow-favouring assumption, approximately 40% yellow males when we made the equal paternity assumption, and approximately 30% yellow males when we made the white-favouring assumption.

Discussion

Aposematism is assumed, and in many organisms shown, to be under positive frequency-dependent selection leading to monomorphism (Mallet & Barton 1989; Lindström *et al.* 2001). In our study system, the wood tiger moth, the more conspicuous yellow male morph, survives as a whole better against predators than the white male morph (Nokelainen *et al.* 2012). However, it appears that this survival advantage is dependent on the local predator community (Nokelainen *et al.* 2014). In our case, this is important because male fitness proves to have not only one (predation), but potentially two (predation and sexual selection) positively frequency-dependent components.

Our enclosure experiment shows that white males have an advantage over yellow males in mating probabilities, but this appears to occur only at high frequencies of white males. However, when we analyse overall individual fitness, both morphs have an advantage when common that disappears when at a balanced ratio. These results may overall suggest some sampling artefact where there is a higher probability of finding higher quality males in the high-frequency morph group. However, this reason seems unlikely because aspects of male quality shown to affect mating success in the past (male age and size) were already included in our analyses to test their effects; and only male age had a small treatment effect (and only in the Y-biased treatment). Instead, two plausible and not

mutually exclusive reasons could be responsible for our results.

First, both morphs could trade off on certain aspects of reproduction. For example, yellow males have low mating success when in white male-dominated environments, but compensate for this by siring more eggs and having a higher hatching success of offspring compared to white males (see Appendix S1, Supporting Information). This difference in strategy could be related to morph-related differences in fitness-related traits we have not measured here. For example, a recent study found that white wood tiger moth males fly for longer stretches of time, whereas yellow males tend to focus their flying around the peak calling activity of female moths (Rojas *et al.* In review). This is probably due to the white males compensating their less efficient signal to predators by increased flight to escape predation attempts, or yellow males being limited in flight because of physiological limitations from producing their more efficient, yet costly chemical defence (Rojas *et al.* unpublished data). Regardless of the specific reason that difference in flight activity between the two morphs may lead to disproportionate chances for the white morph to achieve more matings, especially when common, and in return select over time for the yellow males who leave more sired eggs once they achieve mating.

Secondly, wood tiger moths may display flexible mating preferences, depending on morph frequencies, which would have led to our results. Sampling of natural populations shows that most populations of wood tiger moths have an admixture of both morphs (Hegna, Galarza & Mappes 2015). For example, in Finland alone, there are populations that are white dominated (such as in Central Finland where much of our laboratory stock in this experiment was taken from) and others that are yellow dominated (Southern Finland), and much gene flow occurs between populations (Nokelainen *et al.* 2012; Galarza *et al.* 2014). Adaptive coloration also varies spatially as well as temporally in our system, and viability strongly favours one morph over the other via predation, immunocompetence response to disease (Nokelainen, Lindstedt & Mappes 2013) and potentially other pressures. This may make it adaptive to avoid mating with rare morphs, potentially explaining why female wood tiger moths in our experiment appear to either choose mates (or lay more eggs) for the male morph with the more common phenotype (i.e. the morph with seemingly higher viability). Short-lived polyandrous species like wood tiger moths may easily track male morph frequencies via the social interactions that occur between calling females and courting males (McLain 2005; Westerman *et al.* 2014).

Some examples of flexible mate choice have been confirmed experimentally, and in each case, its occurrence is linked with systems that have natural variations in morph frequencies. For example, female soldier beetles (*Chauliognathus pennsylvanicus*) possess flexible mate preferences

for the most common male beetle of varying elytron spot lengths (McLain 2005). Likewise, male damselflies *Ischnura elegans* exhibit flexible mate choice for the most frequent female colour morph (Van Gossum, Stoks & De Bruyn 2001), and African butterflies *Bicyclus anynana* exhibit flexible learned preferences depending on past social experiences. Specifically, male *B. anynana* have been shown to learn preferences to mate with particular phenotypes of females with dorsal hind wing eyespot numbers, depending on pre-mating experience with that phenotype (Westerman *et al.* 2014). It is suggested that in this system plastic mate preferences, rather than pre-existing mating biases, allow male *B. anynana* butterflies the opportunity to choose the best mate in a highly fluctuating environment season, with social experiences providing the basis for learning.

Given the mating results, we modelled the consequences of two positively frequency-dependent components – predation and sexual selection – for the maintenance of the colour polymorphism in the wild. Our model results hint at an intriguing geographic mosaic of morphs, where variation in selective factors (e.g. predator/parasitic community or vegetation structure) and the level of gene flow can determine whether either one of the morphs spreads to fixation in a population, or whether the selective mosaic maintains a polymorphism indefinitely. If we assume that the yellow male survival advantage is not fixed but varies spatially – as our extensive natural sampling suggests – our results add a quantitative component to the ideas expressed by Roulin & Bize (2007) and Calsbeek, Bonvini & Cox (2010). More specifically, we show that both low and quite strong gene flow between sites forming the selection mosaic can keep the polymorphism intact. However, when the population is close to panmictic, the argument will fail as the morph with the better net success will self-reinforce its own success over the global landscape. This finding is extremely well in line with those of Joron & Iwasa (2005), who found the same qualitative effect of dispersal in another context of positive frequency dependence (that of Müllerian mimicry).

Our findings provide important insight into understanding the maintenance of genetic variation in the wild. We show how spatial mosaics of genetically based male morphs can arise under the mechanisms of frequency-dependent selection and gene flow for aposematic species (see also Sherratt 2006). Our results also show the general added value of empirical parameterizations of models. That said, it clearly would be too optimistic to interpret our rather good success at predicting morph frequencies (in Finnish populations as well as in Fig. 2, yellow morphs are the minority morph) as proof that our model captures all there is to this system. This could be for various reasons. First, we were able to measure spatial variation in natural selection, but logistical constraints have not allowed us to test for spatial variation in sexual selection in wild. Other known complications are that our

model assumes that morphs only differ in survival after maturation and mating success. To our knowledge thus far, wood tiger moth larvae are indistinguishable as to whether they will become white or yellow adult male morphs. However, if morph identity or morph-related differences in the fight against disease and or pathogens (Nokelainen, Lindstedt & Mappes 2013) impacts survival of larvae in so far unknown ways, this could lead to initial adult differences in morph ratio before adult predation risk or mating behaviour occurs.

Secondly, we did not include variation in behaviour between the morphs in our model. However, enclosure experiments have indicated morph differences in flight and mating attempts (Rojas *et al.* unpublished results), which suggests possible differences in fitness components other than those included in the current model (e.g. survival during dispersal). Thirdly, because the dispersal behaviour of moths is currently unknown, we simply allowed dispersing males to move to any non-natal patch in our model. In reality, environmental barriers (including distance, but also geographical features such as lakes or mountains) can impede movement and reduce gene flow. However, recent population genetic analysis shows that there is dispersal among populations (Galarza *et al.* 2014), and thus, it is reassuring that our results were quite robust to surprisingly different levels of gene flow. Fourthly, we intentionally ignored gene flow through females, as we do not know if females express morph genotypes to any degree. Lastly, temporal environmental variations in mating probability or predation pressure remain unstudied and consequently unmodelled. Given that we know that predator community variation impacts moth survival (Nokelainen *et al.* 2014), temporal variation in communities could also potentially strengthen the stability of polymorphisms.

In general, trait polymorphisms are puzzling to understand, because if one form has any type of advantage over another, selection should maximize fitness and drive that morph to fixation (Fisher 1930). Maintaining polymorphisms within populations therefore requires certain mechanisms that provide similar fitness benefits across multiple morphs over time (Maynard Smith 1982; Pryke 2007). Here, we show, using the aposematic wood tiger moth, that under positive frequency-dependent mating selection and for up to an approximate 20% gene flow of individuals between subpopulations, spatial variation in predator pressure is able to operate as a selection mosaic that prevents complete fixation of either morph.

Acknowledgements

We are indebted to Mari Jääntti, Elviira Kuusinen and Jyrki Kalliokoski for helping us in the field experiments, and Kaisa Suisto and Jimi Kirvesoja for help with the care and maintenance of laboratory organisms. Janne Valkonen and Atsushi Honma helped with both building the enclosure and the experiment. We also thank staff at the Konnevesi Research Station, especially Helinä Nisu for the use of their facilities, and scientific advice. Andrés Lopez-Sepulcre helped with analyses and edited versions of the manuscript. Members of the Mappes laboratory group at the Univer-

sity of Jyväskylä were also kind to read and comment on earlier versions of this manuscript. Funding was provided by the Academy of Finland via the Centre of Excellence in Biological Interactions (Project: 210000256 to J.M. and Project: 21000027441 to S.G.).

Data accessibility

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.nn493> (Gordon *et al.* 2015).

References

- Allen, J.A. (1988) Frequency-dependent selection by predators. *Philosophical Transactions of the Royal Society B*, **319**, 485–503.
- Borer, M., Van Noort, T., Rahier, M. & Naisbit, R.E. (2010) Positive frequency-dependent selection on warning colour in alpine leaf beetles. *Evolution*, **64**, 3629–3633.
- Calsbeek, R., Bonvini, L. & Cox, R.M. (2010) Geographic variation, frequency-dependent selection, and the maintenance of a female-limited polymorphism. *Evolution*, **64**, 116–125.
- Clusella-Trullas, S., Van Wyk, J.H. & Spotila, J.R. (2007) Thermal melanism in ectotherms. *Journal of Thermal Biology*, **32**, 235–245.
- Cott, H.B. (1940) *Adaptive Colouration in Animals*. Methuen, London.
- Cummings, M.E. & Crothers, L.R. (2013) Interacting selection diversifies warning signals in a polytypic frog: an examination with the strawberry poison frog. *Evolutionary Ecology*, **27**, 693–710.
- Endler, J.A. & Greenwood, J.J.D. (1988) Frequency-dependent predation, crypsis and aposematic coloration. *Philosophical Transactions of the Royal Society B*, **319**, 505–523.
- Endler, J.A. & Mappes, J. (2004) Predator mixes and the conspicuousness of aposematic signals. *American Naturalist*, **163**, 532–547.
- Endler, J.A. & Rojas, B. (2009) The spatial pattern of natural selection when selection depends upon experience. *American Naturalist*, **173**, E62–E78.
- Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*. Oxford University Press, New York, USA.
- Finkbeiner, S.D., Briscoe, A.D. & Reed, R.D. (2014) Warning signals are seductive: relative contributions of colour and pattern to predator avoidance and mate attraction in *Heliconius* butterflies. *Evolution*, **68**, 3410–3420.
- Galarza, J., Nokelainen, O., Ashrafi, R., Hegna, R.H. & Mappes, J. (2014) Temporal relationship between genetic and warning signal variation in the aposematic wood tiger moth. *Molecular Ecology*, **23**, 4939–4957.
- Gordon, S., Kokko, H., Rojas, B., Nokelainen, O. & Mappes, J. (2015) Data from: Colour polymorphism, torn apart by opposing positive frequency-dependent selection, yet maintained in space. *Dryad Digital Repository*: <http://dx.doi.org/10.5061/dryad.nn493>
- Gray, S.M. & McKinnon, J.S. (2007) Linking colour polymorphism maintenance and speciation. *Trends in Ecology and Evolution*, **22**, 71–79.
- Hegna, R.H., Galarza, J.A. & Mappes, J. (2015) Global phylogeography and geographic variation in warning coloration of the wood tiger moth (*Parasemia plantaginis*). *Journal of Biogeography*, doi:10.1111/jbi.12513.
- Hegna, R.H., Nokelainen, O., Hegna, J.R. & Mappes, J. (2013) To quiver or to shiver: increased melanisation benefits thermoregulation, but reduces warning signal efficacy in the wood tiger moth. *Proceedings of the Royal Society of London B*, **280**, 20122812.
- Hoffman, E.A., Schueler, F.W., Jones, A.G. & Blouin, M.S. (2006) An analysis of selection on a colour polymorphism in the northern leopard frog. *Molecular Ecology*, **15**, 2627–2641.
- Joron, M. & Iwasa, Y. (2005) The evolution of a Müllerian mimic in a spatially distributed community. *Journal of Theoretical Biology*, **237**, 87–103.
- Kotiaho, J., Alatalo, R.V., Mappes, J., Parri, S. & Rivero, A. (1998) Male mating success and risk of predation in a wolf spider: a balance between sexual and natural selection? *Journal of Animal Ecology*, **67**, 287–291.
- Lehtonen, J. & Kokko, H. (2012) Positive feedback and alternative stable states in inbreeding, cooperation, sex roles and other evolutionary processes. *Philosophical Transactions of the Royal Society B*, **367**, 211–221.
- Lindstedt, C., Lindström, L. & Mappes, J. (2008) Hairiness and warning colours as components of antipredator defence: additive or interactive benefits? *Animal Behaviour*, **75**, 1703–1713.
- Lindstedt, C., Lindström, L. & Mappes, J. (2009) Thermoregulation constrains effective warning signal expression. *Evolution*, **63**, 469–478.
- Lindstedt, C., Reudler Talsma, J.H., Ihalainen, E., Lindström, L. & Mappes, J. (2010) Diet quality affects warning coloration indirectly: excretion costs in a generalist herbivore. *Evolution*, **64**, 68–78.
- Lindstedt, C., Eager, H., Ihalainen, E., Kahilainen, A., Stevens, M. & Mappes, J. (2011) Direction and strength of selection by predators for the colour of the aposematic wood tiger moth. *Behavioural Ecology*, **2**, 580–587.
- Lindström, L., Alatalo, R.V., Lyytinen, A. & Mappes, J. (2001) Strong antiapostatic selection against novel rare aposematic prey. *Proceedings of the National Academy of Science*, **98**, 9181–9184.
- Lyytinen, A., Lindström, L. & Mappes, J. (2004) Ultraviolet reflection and predation risk in diurnal and nocturnal Lepidoptera. *Behavioral Ecology*, **15**, 982–987.
- Lyytinen, A., Alatalo, R.V., Lindström, L. & Mappes, J. (2001) Can ultraviolet cues function as aposematic signals? *Behavioral Ecology*, **12**, 65–70.
- Maan, M.E. & Cummings, M.E. (2009) Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *Proceedings of the National Academy of Science*, **106**, 19072–19077.
- Mallet, J. & Barton, N. (1989) Inference from clines stabilized by frequency-dependent selection. *Genetics*, **122**, 967–976.
- Mallet, J. & Joron, M. (1999) Evolution of diversity in warning colour and mimicry: polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics*, **30**, 201–233.
- Mappes, J., Kokko, H., Ojala, K. & Lindström, L. (2014) Seasonal changes in predator community switch the direction of selection for prey defences. *Nature Communications*, **5**:5016.
- Maynard Smith, J. (1982) *Evolution and the Theory of Games*. Cambridge University Press, UK.
- McKinnon, J.S. & Pierotti, M. (2010) Color polymorphism and correlated characters: genetic mechanisms and evolution. *Molecular Ecology*, **19**, 5101–5125.
- McLain, D.K. (2005) Female soldier beetles display a flexible preference for selectively favored male phenotypes. *Evolution*, **59**, 1085–1095.
- Müller, F. (1878) Über die Vortheile der Mimicry bei Schmetterlingen. *Zoologischer Anzeiger*, **1**, 54–55.
- Nokelainen, O., Lindstedt, C. & Mappes, J. (2013) Environment-mediated morph-linked immune and life-history responses in the aposematic wood tiger moth. *Journal of Animal Ecology*, **82**, 653–656.
- Nokelainen, O., Hegna, R.H., Reudler, J.H., Lindstedt, C. & Mappes, J. (2012) Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of the Royal Society of London B*, **22**, 257–265.
- Nokelainen, O., Valkonen, J., Lindstedt, C. & Mappes, J. (2014) Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. *Journal of Animal Ecology*, **83**, 598–605.
- O'Donald, P. & Majerus, M.E.N. (1984) Polymorphism of melanic ladybirds maintained by frequency-dependent sexual selection. *Biological Journal of the Linnean Society*, **23**, 101–111.
- Ojala, K., Lindström, L. & Mappes, J. (2007) Life-history constraints and warning signal expression in an Arctiid moth. *Functional Ecology*, **21**, 1162–1167.
- Olendorf, R., Rodd, F.H., Punzalan, D., Houde, A.E., Hurt, C., Reznick, D.N. *et al.* (2006) Frequency-dependent survival in natural guppy populations. *Nature*, **441**, 633–636.
- Poulton, E.B. (1890) *The Colours of Animals: Their Meaning and Use Especially Considered in the Case of Insects*. Kegan, Trench, Trubnes and Co, NY, USA.
- Pryke, S.R. (2007) Fiery red heads: female dominance among head colour morphs in the Gouldian finch. *Behavioural Ecology*, **18**, 621–627.
- R Core Team (2013) R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Rojas, B. & Endler, J.A. (2013) Sexual dimorphism and intra-population colour pattern variation in the aposematic frog *Dendrobates tinctorius*. *Evolutionary Ecology*, **27**, 739–753.
- Rosenblum, E.B. (2006) Convergent evolution and divergent selection: lizards at the White Sands ecotone. *American Naturalist*, **167**, 1–15.
- Rothschild, M. (1981) The mimics must move with the times. *Biological Journal of the Linnean Society*, **16**, 21–23.
- Roulin, A. & Bize, P. (2007) Sexual selection in genetic colour-polymorphic species: a review of experimental studies and perspectives. *Journal of Ethology*, **25**, 99–105.

- Rowland, H.M., Ihalainen, E., Lindstrom, L., Mappes, J. & Speed, M.P. (2007) Co-mimics have a mutualistic relationship despite unequal defences. *Nature*, **448**, 64–67.
- Sánchez-Guillén, R.A., Hansson, B., Wellenreuther, M., Svensson, E.I. & Cordero-Rivera, A. (2011) The influence of stochastic and selective forces in the population divergence of female colour polymorphism in damselflies of the genus *Ischnura*. *Heredity*, **107**, 513–522.
- Servedio, M.R. (2000) The effects of predator learning, forgetting, and recognition errors on the evolution of warning coloration. *Evolution*, **4**, 751–763.
- Sherratt, T.N. (2006) Spatial mosaic formation through frequency-dependent selection in Mullerian mimicry complexes. *Journal of Theoretical Biology*, **240**, 165–174.
- Sinervo, B. & Calsbeek, R. (2006) The developmental and physiological causes and consequences of frequency dependent selection in the wild. *Annual Review of Ecology and Systematics*, **37**, 581–610.
- Sinervo, B. & Lively, C. (1996) The Rock-Paper-Scissors Game and the evolution of alternative male strategies. *Nature*, **380**, 240–243.
- Stevens, M. & Ruxton, G.D. (2012) Linking the evolution and form of warning coloration in nature. *Proceedings of the Royal Society of London B*, **279**, 417–426.
- Steward, R.C. (1977) Melanism and selective predation in three species of moths. *Journal of Animal Ecology*, **46**, 483–496.
- Svensson, E.I., Abbott, J. & Hardling, R. (2005) Female polymorphism, frequency dependence, and rapid evolutionary dynamics in natural populations. *American Naturalist*, **165**, 567–576.
- Van Gossum, H., Stoks, R. & De Bruyn, L. (2001) Reversible frequency-dependent switches in male mate choice. *Proceedings of the Royal Society of London B*, **268**, 83–85.
- Westerman, E.L., Chirathivat, N., Schyling, E. & Monteiro, A. (2014) Mate preference for a phenotypically plastic trait is learned, and may facilitate preference-phenotype matching. *Evolution*, **68**, 1661–1670.

Received 15 January 2015; accepted 22 June 2015

Handling Editor: Stewart Plaistow

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Egg number and hatching success.