INTRODUCTION

Aposematic signals that warn predators about prey defences have long occupied biologists' attention as a prime example of adaptation (Poulton, 1890; Ruxton, Allen, Sherratt, & Speed, 2018; Wallace, 1867). Aposematism is subject to positive frequency-dependent selection (Fisher, 1958). When it is rare, predation will disfavour aposematism because few predators will have had the opportunity to learn to avoid the warning signal, so those few aposematic individuals present will suffer high per capita mortality (Lindström, Alatalo, 1867).
Lyytinen, & Mappes, 2001). On the other hand, when aposematic prey are common, their per capita mortality will be low, and an established aposematic signal will be very difficult to invade. Indeed, positive frequency dependence can be so strong that individuals of different defended species evolve to resemble one another, in a phenomenon called Müllerian mimicry (Müller, 1879).

Although general evidence for Müllerian mimicry is strong (Ruxton et al., 2018), there are several issues with the theory. Müller’s original hypothesis did not incorporate realistic predator behaviour (Rowland, Ihlanainen, Lindström, Mappes, & Speed, 2007; Sherratt, 2008) and perhaps as a consequence fails to predict diversity within or between mimicry rings (Joron & Mallet, 1998; Sherratt, 2008). New theories envision predator behaviour more realistically (Aubier, Joron, & Sherratt, 2017; Balogh, Gamberale-Stille, Tullberg, & Leimar, 2010; Balogh & Leimar, 2005). Many verbal hypotheses have also been put forth to explain diversity in warning coloration (reviewed in Briolat et al., 2019). However, in most systems we still do not understand why diversity in warning signals occurs, or how actual predators behave. Additionally, many of the most striking examples of aposematism and Müllerian mimicry come from the tropics (Davis Rabosky et al., 2016; Dumbacher & Fleischer, 2001; Huey & Pianka, 1977; Jiggins, 2017; Joshi, Prakash, & Kunte, 2017; Mochida, Zhang, & Toda, 2015; Norman, Finn, & Tregenza, 2001; Padula et al., 2016; Stuckert, Venegas, & Summers, 2014; Summers & Clough, 2001; Symula, Schulte, & Summers, 2001), while diversity of warning signals at temperate latitudes has received less attention (Hegna, Galarza, & Mappes, 2015; Marek & Bond, 2009; Williams, 2007; Wilson, Williams, Forister, von Dohlen, & Pitts, 2012). This could be due to sampling bias or differences in the evolution of warning signals between temperate and tropical communities. At present, though, we lack sufficient data from temperate systems to address this question.

We tested Müller’s hypothesis in a system of beetles from temperate latitudes. Oreina leaf beetles are colourful, chemically defended insects with a range centred in the European Alps (Dobler, Mardulyn, Pasteels, & Rowell-Rahier, 1996). The majority of species are polymorphic and glossy, with most morphs having a basal colour of either blue or green (Kippenberg, 1994; Waldron et al., 2017). Within the Alps, communities of three to six species can be found in patchily distributed locations (Figure 1). They rely heavily on the availability of their host plants—Adenostyles and Petasites (Apiaceae) and Chaerophyllum, Heracleum or Peucedanum (Apiaceae). They also carry chemical defences in specialized exocrine glands (Pasteels, Eggenberger, Rowell-Rahier, Ehmke, & Hartmann, 1992).

Borer et al. (2010) found positive frequency-dependent selection on coloration among local populations of Oreina gloriosa where either blue or green morphs predominated. This supports the hypothesis that predators stabilize the most abundant local colour morph in O. gloriosa and that coloration functions as a warning signal in that species. Taken with the results of Waldron et al. (2017), who showed that specular highlights on the dorsum of Oreina enhance predator avoidance learning, this supports the conjecture that multiple Oreina species may form Müllerian mimicry rings across central Europe. Although this genus is common and its chemical ecology has been well studied (Dobler et al., 1996; Dobler & Rowell-Rahier, 1994), curiously there are no quantitative data on local colour morph frequencies.

If Oreina beetles in fact form local Müllerian mimicry rings, Müller’s hypothesis predicts that (1) colour morphs in each species should exhibit positive covariance among different local communities, indicating that local frequency-dependent selection drives convergence in coloration. We tested this prediction by analysing the frequencies of Oreina colour morphs across eight communities in the Alps. We also tested the behaviour of predators in response to different Oreina colour morphs and species. Specifically, Müller’s hypothesis predicts that (2) predators will be able to learn to avoid either the blue or green morphs; (3) after learning to avoid blue or green morphs, predators will be more willing to attack the alternative colour morph, which would lead to positive frequency-dependent selection; and (4) learned avoidance of a colour morph of one species will be generalized to morphs of different species. Additionally, where these predictions are not met, they have the potential to explain deviations from perfect Müllerian convergence, such as polymorphism within populations.

**Figure 1** Six of the most common species represented in typical blue and green morphs. The species are aligned in alphabetical order; (1) Oreina alpestris, (2) Oreina bifrons, (3) Oreina cacaliae, (4) Oreina gloriosa, (5) Oreina speciosissima and (6) Oreina speciosa.
As thermoregulation may limit the evolution of warning signals in temperate regions (Hegn, Nokelainen, Hegn, & Mappes, 2013; Lindstedt, Lindström, & Mappes, 2009), we tested the additional hypothesis that variation in colour morph among populations of Oreina could be linked to a thermoregulatory advantage of the darker, blue morph. This hypothesis predicted a positive correlation between the frequency of blue morphs in different populations and elevation.

2 | METHODS

2.1 | Oreina frequencies in the field

2.1.1 | Collection

We chose eight locations in the Alps at which to characterize communities of Oreina (Figure S1). These locations were a minimum distance of 50 km apart, far larger than the territory sizes of any individual avian insectivores. At each location, we sampled multiple sites across early, mid- and late summer, ranging from 3 June to 7 August in 2012 and 2013 (Figure S2). Sites were distributed over an elevational gradient ranging from 961 to 2,338 m. Due to logistical constraints, we were not able to visit each site every season or year (sites have been consistently dominated by the same colour morph—either blue or green—for over 20 years of visits, SD, pers. obs.). Sites measured 5 × 5 m square, and like the locations in which they were nested, were far enough apart that they were unlikely to be shared between any passerine bird territories. We scanned the surface of the vegetation and attempted to collect all of the Oreina that could be found in 10 min of searching. Oreina are long-lived or relatively stationary and tend to be on the visible surfaces of the broad leaves of their host plants, so our methods were unlikely to be biased towards collecting one morph or the other even as their frequencies varied between sites.

In our study location, a total of eight species can be identified according to the most recent key for the genus Oreina and its sister genus Chrysolina (Bourdonné & Douget, 1991; Kippenberg, 1994): O. alpestris, O. bifrons, O. cacaliae, O. gloriosa, O. speciosissima, O. speciosa, O. virgulata and Chrysolina coerulans.

2.1.2 | Statistical analysis

We assigned each beetle to a colour morph category (blue or green). Intermediate or alternative forms were rare, making morph assignment straightforward. We also assigned them to a taxonomic group: O. cacaliae and “other.” We adopted this very simple taxonomic classification because although males of all species can be unambiguously identified by dissection, morphological identification of females is often difficult or impossible. Particularly, O. alpestris and O. speciosa are extremely difficult to tell apart and recent genetic analyses revealed a high amount of gene flow and introgression among them (Triponez et al., 2011). Oreina cacaliae, on the other hand, can be easily recognized by its flat and elongate body shape. Another reason to consider O. cacaliae separate from other Oreina is that the latter use autogenously produced cardenolides for defence, whereas O. cacaliae uses pyrrolizidine alkaloids sequestered from its host plants (Pasteels et al., 1992; Rowell-Rahier, Pasteels, Alonso-Mejia, & Brower, 1995). The pooling among Oreina spp. and C. coerulans was further justified by our behavioural experiments (see Results). We emphasize that species-level identification was not necessary for testing the predictions of our hypotheses.

We used logistic regression to test the prediction (1) that the most frequent colour morph of one species would be correlated with the most frequent colour morph of other species across sites. We modelled the ratio of blue:green beetles of “other” species at each site, with the proportion of blue O. cacaliae as a predictor, and the location of each site included as a random effect.

In a separate model, we tested the prediction of the thermoregulatory hypothesis for differences in colour between different sites: that colour morph frequency would be correlated with elevation. To test this, we used a logistic regression to model the ratio of blue:green among all beetles at each site, with elevation as a continuous predictor and location (which contained multiple sites) included as a random effect. Latitudinal variation was not included, as a relatively small variation in temperature due to its effect was likely given the spatial distribution of our sites. Both of these analyses were performed using the lme4 package (Bates, Maechler, Bolker, & Walker, 2019) in R 3.4.4 (R Core Team, 2019). The significance of predictors in each model was assessed using the likelihood ratio test (LRT; Bolker, 2008).

2.2 | Laboratory behaviour assays and appearance quantification

2.2.1 | Ethical note

Experiments were carried out at Konnevesi Research Station, Finland, between January and February 2014. The Central Finland Centre for Economic Development, Transport and Environment (KESELY/1017/07.01/2010) and National Animal Experimental Board (ESAVI-2010-087517Ym-23 & ESAVI/9114/04.10.07/2014) provided permission to keep wild blue tits (Cyanistes caeruleus) in captivity and to use them in the behavioural study.

2.2.2 | Birds

Blue tits (Cyanistes caeruleus) are generalist avian insectivores found at our research station, and also in the Alps where they are sympatric with Oreina. Importantly, however, blue tits from our research station are nonmigratory, making them naive to Oreina. Wild-caught birds were housed in purpose-built aviaries
(Lindström, Lyytinen, Mappes, & Ojala, 2006) and were ringed and released at the place of capture after the experiments were complete. In two distinct experiments, we first tested avoidance learning and colour generalization between blue and green morphs of *O. cacaliae* and, second, generalization between green morphs of *O. cacaliae* and *O. alpestris*. Each bird participated in only one of the two experiments.

### 2.2.3 | Prey

We collected green beetles, *O. cacaliae* and *O. alpestris* (formerly known as *O. variabilis*; Dobler et al., 1996; see Kippenberg, 1994) near Zastler in the Black Forest, Germany (June 2013), and blue *O. cacaliae* from Tschiertschen, Switzerland (June 2013). Both beetle species release chemical defensive secretions from exocrine glands located laterally on the pronotum and the elytra when they are disturbed or attacked. As the chemical defences of the two species differ, we made sure that the beetles we used of each species contained their maximum supply of chemicals, as well as we could manage given our experimental equipment. We did this by maintaining beetles in environmental chambers (day: night temperature regime of 16:16°C and a light schedule of 16:8 hr) for 15 days with food ad libitum before freezing them at −80°C. When we could manage given our experimental equipment. We did this by maintaining beetles in environmental chambers (day: night temperature regime of 16:16°C and a light schedule of 16:8 hr) for 15 days with food ad libitum before freezing them at −80°C. Species-appropriate host plants were gathered at the site of collection and later provided by the Botanical Garden, Universität Hamburg, Germany.

### 2.2.4 | Aviaries

Experimental aviaries (50 × 65 × 45 cm) were lit to simulate daylight using a 26W, Repti Glo 5.0 Compact UVB light bulb positioned in the centre of the aviary roof. Each aviary had a perch, a water bowl (access ad libitum) and a small mesh-covered observation hole. A vertical metal barrier (ca. 15 cm) was used to obscure prey items during behavioural experiments. A moveable floor tray was used to insert a Petri dish, fitted with green cardboard, to present sunflower seeds and prey during training and experimental trials.

### 2.2.5 | Pretraining

Pretraining included habituation into the experimental cage and training birds to find food behind the metal barrier, which obscured it from view while they were on their perches. Birds had to hop on top of the barrier to observe the food, giving an obvious point at which to begin recording latency to attack prey items. For further description, see Rojas et al. (2017). Pretraining took between 1 and 4 hr to complete. All birds completed pretraining successfully.

### 2.2.6 | General protocol for experimental trials

Both the colour and the species generalization experiments followed the same protocol (Figure 2a). After pretraining was complete, a two-hour food deprivation period was started, to ensure that all birds were in a similar state of hunger. After two hours, a last instar larvae of *Tenebrio molitor* (mealworm) was offered to check that each bird was hungry and motivated to forage. After eating the larvae, a second, shorter food deprivation period (20 min) began before we offered an adult control beetle of *T. molitor*. The palatable *T. molitor* prevented birds from learning to avoid all beetles per se. Two to five minutes after the *T. molitor* had been consumed, we introduced an *Oreina*. We began recording attack latency as soon as the bird landed on the metal barrier and could see the beetle. Birds had a maximum of four minutes (240 s) to attack the *Oreina* before the trial was terminated. If the bird did not attack within this time, the trial was recorded as a nonattack. An attack was recorded if the beetle was touched. Finally, to verify that birds remained motivated to forage throughout the experiment, we offered a final mealworm that concluded the experimental trial.

### 2.2.7 | Colour morph generalization experiment

This experiment tested prediction (2) that predators could learn to avoid either blue or green colour morphs and prediction (3) that learning to avoid one colour morph would not result in protection for the alternative morph. It consisted of a total of four experimental trials for each bird. First, during avoidance learning, we distributed 76 birds into two treatment groups (Blue and Green) that were exposed to a single colour morph of *O. cacaliae* (blue or green, respectively) in three trials that were administered on three different days. On the fourth day, we administered a generalization trial to see whether learned avoidance of one colour morph transferred to the alternative colour morph. For the generalization trial, Blue and Green groups were subdivided into two smaller groups, two control (BB and GG) and two treatment groups (BG and GB) (Figure 2b). The first letter indicates the colour morph that birds were trained to avoid, whereas the second letter indicates the colour morph presented to birds during the generalization test (trial 4). Thus, the two control groups (BB and GG) received the same colour morphs for all four days of the experiment. The treatment groups (BG and GB) experienced either blue or green morphs during the three days of avoidance learning before receiving the alternative morph on the fourth day.

### 2.2.8 | Species generalization experiment

Prediction (4) is critical for the evolution of Müllerian mimicry: learned avoidance of a colour morph of one species should be generalized to different species of the same colour. To test this, we used the most distinctive species: *O. cacaliae* and *O. alpestris*. If these
beetles are generalized, then it is safe to assume that even more similar species will also be generalized. We trained birds to avoid one colour morph (green) of either *O. cacaliae* or *O. alpestris*, and then measured their response to the same colour morph of the other species. A total of 56 birds were initially split into two groups (each receiving either *O. cacaliae* or *O. alpestris*) for three trials of avoidance learning. The birds were then split again into four groups, two control groups (CC and AA) and two treatment groups (CA and AC) (Figure 2c). For each group, the first letter indicates the species offered during avoidance learning, and the second indicates the species offered during the generalization trial. ‘C’ refers to *O. cacaliae* whereas ‘A’ refers to *O. alpestris*.

Due to a low number of blue tits captured and a restricted number of green *O. cacaliae*, data for treatment CC, where the birds experienced green *O. cacaliae* throughout the experiment, were taken from a previous experiment (Waldron et al., 2017). Despite minor differences in the beetle preparations before they were offered to the birds, this produced no differences in initial latency to attack or in learning (see Supporting information).

### 2.2.9 Statistical analysis

For the colour generalization experiment using *O. cacaliae*, we analysed birds’ willingness to attack during the first encounter using Cox proportional hazard regression. The attack risk per unit time was modelled as a function of colour. Then, we analysed how birds learned to avoid blue and green colour morphs across trials 2–3. We used a mixed-effects Cox proportional hazard model with attack latency as the dependent variable; colour, trial and their interaction as predictors; and bird identity as a random effect (package “coxme”; Therneau, 2019). Finally, we compared attack latency between treatment and control groups using Cox proportional hazard models in trial 4 (generalization).

We used an identical analysis for the species generalization experiment between *O. cacaliae* and *O. alpestris*, except that wherever colour was used as a predictor variable, we substituted species. In addition, we analysed learning across trials 1–3 instead of trials 2–3, as most learning took place between trials 1 and 2 (see Results).

All analyses were performed in R 3.4.4 (R Core Team, 2019).
3 | RESULTS

3.1 | Oreina frequencies in wild populations

We collected 966 beetles for analysis. Forty-six per cent were O. cacaliae. The remainder were split among six other species of Oreina (O. alpestris, O. bifrons, O. gloriosa, O. speciosissima, O. speciosa and O. virgulata) and Chrysolina coerulans. We found that the ratio of blue:green individuals at each site among “other” species was significantly predicted by the proportion of blue O. cacaliae (Figures 3 and 4; LRT: $\chi^2 = 8.4, p = .003$). We did not find any significant relationship between the ratio of blue:green Oreina and elevation (LRT: $\chi^2 = 0.75, p = .39$), falsifying the prediction of the hypothesis that divergence in the colour morph among Oreina communities can be explained by different mean temperatures.

3.2 | Generalization between colour morphs

3.2.1 | Colour morph generalization

Cox proportional hazards regression revealed that birds initially hesitated ~50 s longer in trial 1 to attack green beetles than blue beetles (Table 1a, Figure 5). During learning (trials 2–3), the effect of colour was nonsignificant (Table 1b, Figure 5), but the effect of trial was significant (Table 1c). The colour:trial interaction was nonsignificant (Table 1d). Therefore, birds learned to avoid both colour morphs at the same rate. This confirmed prediction 2 that birds should learn to avoid both Oreina colour morphs.

In trial 4, we tested to see whether birds generalized their learned avoidance of one colour onto the other. Birds did not entirely generalize their learned avoidance of green beetles onto blue beetles (treatment GG versus GB; Table 1f, Figure 5). This agreed with prediction 3 that the colour morph that birds have learned to avoid should receive more protection. However, we found that birds strongly generalized learned avoidance of blue beetles onto green beetles, such that there was no difference in their attack rates (treatment BB versus BG; Table 1f, Figure 5). Thus, prediction 3 was rejected for birds that learned to avoid blue.

3.2.2 | Species generalization

Birds hesitated longer to attack O. cacaliae than O. alpestris in trial 1 (Table 2a, Figure 6). This hesitation carried over into our analysis of learning across trials 1–3: species significantly affected attack latency (Table 2b). However, birds learned to avoid both species across the learning trials (Table 2c), and there was no difference in learning rate for O. cacaliae or O. alpestris (Table 2d). In trial 4, there were no differences between the control and treatment groups (Table 2e,f), indicating that birds completely generalized their learned avoidance of one species onto the other. This result confirmed prediction 4 that predators should not distinguish between species after learning.

4 | DISCUSSION

We tested the predictions of the hypothesis that Oreina leaf beetles are classical Müllerian mimics (Müller, 1879). Field data confirmed the prediction that the ratio of blue:green morphs in different species should covary across different local communities. Behavioural experiments confirmed two additional predictions: birds learned to avoid either blue or green beetles, and they generalized their learned avoidance of green morphs of one species to green morphs of another, quite different species. The prediction that avoidance of one colour should not be generalized to another colour, which would produce positive frequency-dependent selection, was not entirely true, however. Avoidance of green was not well generalized to blue, but avoidance of blue was entirely generalized to green. Therefore, green communities of Oreina meet all the criteria for Müllerian mimicry, but the evidence in blue communities is more equivocal. If green morphs
benefit more from predator generalization than blue morphs, what prevents green from invading blue populations and spreading to fixation?

The persistence of blue morphs could be explained if they were favoured in environments with cooler temperatures where their coloration might let them become active more quickly than green morphs (Hegna et al., 2013; Lindstedt et al., 2009). This hypothesis predicts a positive relationship between the frequency of blue morphs in local communities and elevation (higher elevations are of course on average cooler). We did not find such a pattern, rejecting the simple hypothesis that blue beetles are favoured at cooler temperatures, although other, more complex interactions between beetle colour and environment are possible.

A more plausible explanation lies in predator behaviour. In the wild, the positive frequency-dependent selection that characterizes Müllerian mimicry may be more evident in blue Oreina communities than our data imply. Borer et al. (2010) found that 65% of wild populations of O. gloriosa experience positive frequency-dependent selection on colour, in both green and in blue populations. This suggests that at their naturally occurring frequencies, both colour morphs experience an advantage when they are most abundant. Therefore, although our bird experiments show an advantage to green morphs under our laboratory conditions, a change in bird behaviour with frequency might help explain why blue morphs are not driven to extinction. This raises the additional question of how an advantage to green morphs gives way to an advantage to blue ones.

Our data on morph frequencies in the field show an asymmetry. In green communities, blue morphs are not present among any species (Figure 3, top row). In contrast, in blue communities, green morphs are often a noticeable minority of individuals (Figure 3, bottom row). Green morphs may have an advantage when they are rare that stems from asymmetric predator generalization. This may allow green to spread until it achieves a noticeable frequency, which would produce the pattern in the bottom row of Figure 3. We hypothesize that above a certain threshold of rarity, but before they become more common than blue morphs, predators may begin to recognize green beetles as a distinct category of prey and start to sample them (i.e. overcoming neophobia; Sherratt, 2011). This would allow blue communities to resist invasion by green morphs, yet at the same time explain the persistence of green morphs in blue communities. To visualize the scenario that we describe in terms of a selective surface, see Figure 7, which depicts both negative frequency-dependent selection at low frequencies and positive frequency-dependent selection at high frequencies. Traditionally, negative frequency-dependent selection and positive frequency-dependent selection are considered separately,
but in mimicry, predator psychology may produce irregular selective surfaces.

If the hypothesis that we describe above is correct, it would be a psychological twist on the evolution of Müllerian mimicry. It is well understood that the simple mathematical model proposed in 1879 by Fritz Müller does not accurately describe predator behaviour (Rowland et al., 2007; Sherratt, 2008). Müller assumed that predators had to sample a fixed number of prey with a given phenotype to learn avoidance. This assumption has been shown to be false by experimental studies, which instead find that as the abundance of defended prey increases, they are increasingly likely to be sampled (Beatty, Beirinckx, & Sherratt, 2004; Lindström et al., 2001; Rowland, Hoogesteger, Ruxton, Speed, & Mappes, 2010; Rowland, Wiley, Ruxton, Mappes, & Speed, 2010). Indeed, predators should display this behaviour if they learn about prey optimally (Sherratt, 2011). Optimal learning by predators also leads to neophobia, which can allow diversity to persist in Müllerian mimicry rings (Aubier & Sherratt, 2015). In a Müllerian mimicry ring, neophobia permits rare morphs to persist at low frequencies. When they increase above a certain threshold, however, they become worthwhile for predators to sample, which depresses their numbers. Only if they rise to become the predominant morph in a community through some demographic accident does positive frequency-dependent selection work in their favour (Aubier & Sherratt, 2015). Generalized avoidance of green morphs from learned avoidance of blue morphs could have a similar effect to neophobia, maintaining green morphs at low frequencies in blue communities, if it attenuates as predators encounter more and more green morphs (Figure 7). This scenario is entirely consistent with results from this study and Borger et al. (2010), and implies that predator biases may play a role in maintaining diversity in warning coloration within mimicry rings. A rigorous test of this hypothesis would require measuring the fitness of green morphs over a range of frequencies and would be an important addition to our understanding of frequency-dependent selection on invertebrate coloration (Ajuria Ibarra & Reader, 2013). To illustrate mechanism, it would also be necessary to show that as more green beetles are encountered by predators, predators become more likely to sample them. Furthermore, this hypothesis does not explain why there is an asymmetry in predator generalization between blue and green morphs in the first place. It is also too simplistic for systems where trait variation between morphs is more continuous.

Asymmetric generalization has been found among different coloured seed bugs (Lygaeidae: Gamberale-stille & Tullberg, 1999) and firebugs (Pyrrhocoridae; Svádová et al., 2009), in shiny versus dull O. cacaliae (Waldron et al., 2017), and in poison frogs (Lawrence et al., 2019). Why it occurs is not entirely clear. One possibility is that morphs that benefit from asymmetric generalization are a supernormal stimulus for predators (Gamberale-Stille & Tullberg, 1996). Supernormal stimuli occur within the same dimension as a stimulus that an animal has been conditioned to, but at a greater intensity, eliciting a correspondingly greater response (Baddeley, Osorio, & Jones, 2007; Ghirlanda & Enquist, 2003). If the green coloration of Oreina is indeed a more intense stimulus than blue along a single dimension of predator perception (e.g., hue), it should be possible to predict predator responses to more and less intense stimuli along the same dimension.
An alternative hypothesis to explain asymmetric generalization is that predators learn different traits when they are conditioned to different prey types. This could occur through a "stimulus salience" effect, where exceptionally salient stimuli tend to outcompete others during the associative learning process (Kazemi, Gamberale-Stille, & Leimar, 2015; Kazemi, Gamberale-Stille, Tullberg, & Leimar, 2014; Kikuchi & Dornhaus, 2018; Kikuchi, Mappes, Sherratt, & Valkonen, 2016; Mackintosh, 1976; Sherratt, Whissell, Webster, & Kikuchi, 2015). For example, if green is exceptionally salient, it may be the only trait that a predator learns to associate with unpalatability during training, so if the predator subsequently encounters a blue beetle, the blue beetle will be attacked. In contrast, if blue is not particularly salient, during training a predator may learn to avoid blue to some degree, but also learn to avoid beetles on the basis of curvature, specular reflectance and antenna length. Then, a green beetle that is otherwise similar to blue beetles may receive protection despite having a different colour. This hypothesis produces the clear prediction that asymmetric generalization will occur only when different kinds of prey vary in a trait that differs in salience between them. Our species generalization experiment only included green beetles, but examining blue beetles should reveal a lower degree of protection to blue. This is shown by the increased latency to attack blue beetles in the test trial of the GB group compared to birds’ initial latency to attack blue beetles (Figure 5). Birds were very reluctant to attack Oreina at all even during the first trial, suggesting some degree of innate avoidance or general learned aversion to colourful beetles. Juvenile birds that are more exploratory in their dietary choices might therefore be more important agents of selection on warning signals in Oreina. Indeed, such naive predators are typically considered to be the driving force behind Müllerian mimicry (Doktorovová et al., 2019; Exnerová, Svádová, Fucíková, Drent, & Sty, 2010; Rowe, Lindström, & Lytinen, 2004; Ruxton et al., 2018; but see Ihalainen, Lindström, Mappes, & Puolakkainen, 2008).

We still do not know why some Oreina communities are blue and others are green to begin with, although we rejected thermoregulation as an explanation (sampling at other times of the year might still reveal a thermoregulatory benefit, however). There are many other hypotheses to explain diversity in warning signals that our present data do not speak to, but which might prove worthwhile avenues for future investigation (Briolat et al., 2019). Different signals may have differential efficacy among habitats due to different predator communities (Mappes, Kokko, Ojala, & Lindström, 2014; Nokelainen, Valkonen, Lindstedt, & Mappes, 2013; Valkonen et al., 2012; Willmott, Willmott, Elias, & Jiggins, 2017) or light environments (Endler, 1991), or the local prey community may change how signals are selected (Ihalainen, Rowland, Speed, Ruxton, & Mappes, 2012; Kikuchi, Dornhaus, Gopeechund, & Sherratt, 2019). There may also be fitness trade-offs for different colour morphs that are unrelated to their function in antipredator signals—they could be targets of sexual selection (Estrada & Jiggins, 2008; Finkbeiner, Briscoe, & Reed, 2014; Jiggins, Naisbit, Cole, & Mallet, 2001; Nokelainen, Hegna, Reudler, Lindstedt, & Mappes, 2012), selection for parasite resistance (Fröman, Lindstedt, Hiltunen, Laakso, & Mappes, 2009; Vilmos & Kurucz, 1998), or require environmentally derived precursors (Lindstedt, Talsma, Ihalainen, Lindström, & Mappes, 2010). Finally, and perhaps most persuasively, shifting balance may promote the origin of novel mimicry rings as new populations of aposematic species expand across a landscape (Mallet, 2010; Mallet & Joron, 1999; Sherratt, 2006). We urge careful attempts to reject hypotheses to improve our understanding of this fascinating system.

In sum, we found that simple assumptions about predator behaviour could only explain Müllerian mimicry among green morphs of Oreina. More nuanced predator behaviour may explain the persistence of Müllerian mimicry among blue Oreina morphs. Due to its simple basis in a binary colour polymorphism that diverges in frequency among populations, Oreina mimicry rings may also be ideal to test evolutionary hypotheses about mimicry in the field. Oreina adds to our knowledge of warning signals at temperate latitudes, showing that even in comparatively species-poor communities such as alpine meadows, warning signals experience complex selective pressures.

**ACKNOWLEDGMENTS**

We thank staff at Konnevesi research station for logistical support and Jimi Kirvesjoa for help with collecting beetle spectra. We thank members of Mappes lab and Tom Reader for feedback.
on earlier versions of the manuscript. Research was jointly funded by the Centre of Excellence in Biological Interactions, University of Jyväskylä, Finland, University of Hamburg, Germany and the Templeton Foundation (Grant ID #41855 to SD). DWK was funded by NIH-K12GM000708.

AUTHOR CONTRIBUTIONS

DKW wrote the manuscript and performed the analyses. SJW did the experiments, aided with analyses and wrote a draft of the manuscript. JV aided with experiments, analyses and writing. SD aided with experiments and writing. JM conceived of the project, designed the methods and aided with experiments and writing.

DATA AVAILABILITY STATEMENT

The data sets of the article are available in dryad https://doi.org/10.5061/dryad.6q573n5w5

ORCID

David W. Kikuchi https://orcid.org/0000-0002-7379-2788
Janne K. Valkonen https://orcid.org/0000-0002-2177-6612
Susanne Dobler https://orcid.org/0000-0002-0635-7719
Johanna Mappes https://orcid.org/0000-0002-1117-5629

REFERENCES

Ajría Ibarra, H., & Reader, T. (2013). Reasons to be different: Do conspicuous polymorphisms in invertebrates persist because rare forms are fitter?: Maintenance of colour polymorphisms in invertebrates. *Journal of Zoology*, 290(2), 81–95. https://doi.org/10.1111/jzo.12034

Aubier, T. G., Joron, M., & Sherratt, T. N. (2017). Mimicry among unbalanced prey should be mutualistic when predators sample optimally. *The American Naturalist*, 189(3), 267–282. https://doi.org/10.1086/690121

Aubier, T. G., & Sherratt, T. N. (2015). Diversity in Mullerian mimicry: The optimal predator sampling strategy explains both local and regional polymorphism in prey. *Evolution*, 69(11), 2831–2845. https://doi.org/10.1111/evo.12790

Badeley, R., Osorio, D., & Jones, C. D. (2007). Generalization of color by chickens: Experimental observations and a Bayesian model. *The American Naturalist*, 169(January), S27–S41. https://doi.org/10.1086/510142

Balogh, A. C. V., Gamberale-Stille, G., Tullberg, B. S., & Leimar, O. (2010). Feature theory and the two-step hypothesis of Müllerian mimicry evolution. *Evolution*, 64(3), 810–822. https://doi.org/10.1111/j.1558-5646.2009.00852.x

Balogh, A. C. V., & Leimar, O. (2005). Müllerian mimicry: An examination of Fisher’s theory of gradual evolutionary change. *Proceedings of the Royal Society B: Biological Sciences*, 272(1578), 2269–2275. https://doi.org/10.1098/rspb.2005.3227

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2019). *Package lme4*. B. D. C., Beirnckx, K., & Sherratt, T. N. (2004). The evolution of mullerian mimicry in multispecies communities. *Nature*, 431(7004), 63–67. https://doi.org/10.1038/nature02818

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton, NJ: Princeton University Press.

Borer, M., Noort, T. V., Rahier, M., Naisbit, R. E., Van Noort, T., Rahier, M., ... Naisbit, R. E. (2010). Positive frequency-dependent selection on warning color in alpine leaf beetles. *Evolution*, 64(12), 3629–3633. https://doi.org/10.1111/j.1558-5646.2010.01137.x

Bourdonné, J. C., & Douget, S. (1991). Données sur la biosystématique des Chrysolina L.S. (Coleoptera: Chrysomelidae: Chrysomelinae). *Annales de la Société Entomologique de France (N.S.),* 27, 29–64.

Briolat, E. S., Burdfield-steel, E. R., Paul, S. C., Katja, H. R., Seymoure, B. M., Stankowich, T., & Stuckert, A. M. (2019). Diversity in warning coloration: Selective paradox or the norm? *Biological Reviews*, 94, 388–414. https://doi.org/10.1111/brv.12460

Balogh, A. C. V., & Leimar, O. (2005). Müllerian mimicry: An examination of Mullerian mimicry evolution. *Evolution*, 50(6), 2373–2386. https://doi.org/10.2307/2410706

Dobler, S., & Rowell-Rahier, M. (1994). Production of cardenolides versus sequestration of pyrrolizidine alkaloids in larvae of *Oreina* species (Coleoptera, Chrysomelidae). *Journal of Chemical Ecology*, 20(3), 555–568.

Dobler, S., & Rowell-Rahier, M. (1994). Positive frequency-dependent selection on warning color in alpine leaf beetles. *Evolution*, 68(2-018-9465-8

Dumbacher, J. P., & Fleischer, R. C. (2001). Phylogenetic evidence for colour pattern convergence in toxic pitohuis: Müllerian mimicry in birds? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1480), 1971–1976. https://doi.org/10.1098/rspb.2001.1717

Endler, J. A. (1991). Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Research*, 31(3), 587–608. https://doi.org/10.1016/0042-6999(91)90109-1

Estrada, C., & Jiggins, C. D. (2008). Interspecific sexual attraction because of convergence in warning colouration: Is there a conflict between natural and sexual selection in mimetic species? *Journal of Evolutionary Biology*, 21(3), 749–760. https://doi.org/10.1111/j.1420-9101.2008.01517.x

Exnerová, A., Svádová, K. H., Fucíková, E., Drent, P., & Stys, P. (2010). Personality matters: Individual variation in reactions of naive bird predators to aposematic prey. *Proceedings of the Royal Society B: Biological Sciences*, 277(1682), 723–728. https://doi.org/10.1098/rspb.2009.1673

Finkbeiner, S. D., Briscoe, A. D., & Reed, R. D. (2014). Warning signals are seductive: Relative contributions of color and pattern to predator avoidance and mate attraction in Heliconius butterflies. *Evolution*, 68, 3410–3420.

Fisher, R. A. (1958). The genetic theory of natural selection: A complete variourum edition. Oxford, UK: Oxford University Press.

Friman, V. P., Lindstedt, C., Hiltunen, T., Laakso, J., & Mappes, J. (2009). Predation on multiple trophic levels shapes the evolution of pathogen virulence. *PLoS ONE*, 4(8), 8–13. https://doi.org/10.1371/journal.pone.0006761

Gambarale-Stille, G., & Tullberg, B. S. (1996). Evidence for a peak-shift in predator generalization among aposematic prey. *Proceedings of the Royal Society B: Biological Sciences*, 263(1375), 1329–1334. https://doi.org/10.1098/rspb.1996.0195

Gambarale-stille, G., & Tullberg, B. S. (1999). Experienced chicks show biased avoidance of stronger signals: An experiment with natural colour variation in live aposematic prey. *Evolutionary Ecology*, 13, 579–589.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, 66, 15–36. https://doi.org/10.1006/anbe.2003.2174

Hegna, R. H., Galarza, J. A., & Mappes, J. (2015). Global phylogeography and geographical variation in warning coloration of the wood tiger.
moth (Parasemia plantaginis). Journal of Biogeography, 42(8), 1469–1481. https://doi.org/10.1111/jbi.12513

Hegna, R. H., Nokelainen, O., Hegna, J. R., & Mappes, J. (2013). To quiver or to shiver: Increased melanization benefits thermoregulation, but reduces warning signal efficacy in the wood tiger moth. *Proceedings of the Royal Society—Biological Sciences*, 280, 20122812.

Huey, R. B., & Pianka, E. R. (1977). Natural selection for juvenile lizard mimicking noxious beetles. *Science*, 195(4258), 201–203.

Ihalainen, E., Lindström, L., Mappes, J., & Puolakkainen, S. (2008). Can experienced birds select for Müllerian mimicry? *Behavioral Ecology*, 19(2), 362–368. https://doi.org/10.1093/beheco/arm151

Ihalainen, E., Rowland, H. M., Speed, M. P., Ruxton, G. D., & Mappes, J. (2012). Prey community structure affects how predators select for Müllerian mimicry. *Proceedings of the Royal Society B: Biological Sciences*, 279(1736), 2099–2105. https://doi.org/10.1098/rspb.2011.2360

Jiggins, C. D. (2017). The ecology and evolution of Heliconius butterflies. Oxford, UK: Oxford University Press.

Jiggins, C. D., Naisbit, R. E., Coe, R. L., & Mallet, J. (2001). Reproductive isolation caused by colour pattern mimicry. *Nature*, 411(6835), 302–305. https://doi.org/10.1038/35077075

Jorom, M., & Mallet, J. (1998). Diversity in mimicry: Paradox or paradigm? *Trends in Ecology and Evolution*, 534798, 461–466.

Joshi, J., Prakash, A., & Kunte, K. (2017). Evolutionary assembly of communities in butterfly mimicry rings. *The American Naturalist*, 189(4), E58–E76. https://doi.org/10.1086/690907

Kazemi, B., Gamberale-Stille, G., & Leimar, O. (2015). Multi-trait mimicry and the relative salience of individual traits. *Proceedings of the Royal Society B: Biological Sciences*, 282(1818), 20152127. https://doi.org/10.1098/rspb.2015.2127

Kazemi, B., Gamberale-Stille, G., Tullberg, B. S., & Leimar, O. (2014). Stimulus salience as an explanation for imperfect mimicry. *Current Biology*, 24(9), 965–969. https://doi.org/10.1016/j.cub.2014.02.061

Kikuchi, D. W., & Dornhaus, A. (2018). How cognitive biases select for imperfect mimicry: A study of asymmetry in learning with bumblebees. *Animal Behaviour*, 144, 125–134.

Kikuchi, D. W., Dornhaus, A., Gopeechund, V., & Sherratt, T. N. (2019). Community diversity determines predator generalization in mimicry. *Elife*, 8, e43965.

Kikuchi, D. W., Mappes, J., Sherratt, T. N., & Valkonen, J. K. (2012). Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), 257–265. https://doi.org/10.1098/rspb.2011.0880

Kokko, H., Ojala, K., & Lindström, L. (2014). Seasonal changes in predator community switch the direction of selection for prey defences. *Nature Communications*, 5, 1–7. https://doi.org/10.1038/ncomms6016

Lindstedt, C., Lindstedt, J. H., Ihalainen, E., Lindström, L., & Mappes, J. (2009). Thermoregulation contributes to variation in warning colour and mimicry. *Ecological Entomology*, 35, 90–104. https://doi.org/10.1111/j.1365-2311.2009.01137.x

Mallet, J., & Joron, M. (1999). Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics*, 30(1), 201–233. https://doi.org/10.1146/annurev.ecolsys.30.1.201

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, 1–7. https://doi.org/10.1038/ncomms6016

Marek, P. E., & Bond, J. E. (2009). A Müllerian mimicry ring in Appalachian millipedes. *Proceedings of the National Academy of Sciences of the United States of America*, 106(24), 9755–9760. https://doi.org/10.1073/pnas.0810408106

Mochida, K., Zhang, W. Y., & Toda, M. (2015). The function of body coloration of the hai coral snake Sinomicroplus juponicus boettgieri. *Zoolological Studies*, 54(Feb), e33. https://doi.org/10.1186/s40555-015-0110-2

Müller, F. (1879). Ituna and Thyridia: A remarkable case of mimicry in butterflies. *Proceedings of the Entomological Society of London*, 1879, xx–xxiv.

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 B: Biological Sciences*, 279(1727), 257–265. https://doi.org/10.1098/rspb.2011.0880

Nokelainen, O., Valkonen, J. K., Lindstedt, C., & Mappes, J. (2013). Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. *Journal of Animal Ecology*, 83(3), 598–605. https://doi.org/10.1111/j.1365-2656.2012.14169

Norman, M. D., Finn, J., & Tregenza, T. (2001). Dynamic mimicry in an Indo-Malayan octopus. *Proceedings of the Royal Society B: Biological Sciences*, 268(1478), 1755–1758. https://doi.org/10.1098/rspb.2001.1708

Padula, V., Bahia, J., Stöger, I., Camacho-García, Y., Malaquias, M. A. E., Cervera, J. L., & Schrödl, M. (2016). A test of color-based taxonomy in nudibranchs: Molecular phylogeny and species delimitation of the *Felimida clenchi* (Mollusca: Chromodorididae) species complex. *Molecular Phylogenetics and Evolution*, 103, 215–229. https://doi.org/10.1016/j.ympev.2016.07.019

Pastels, J. M., Eggenberger, F., Rowell-Rahier, M., Ehmke, A., & Hartmann, T. (1992). Chemical defense in Chrysomelid leaf beetles. *Naturwissenschaften*, 79(11), 521–523.

Poulton, E. B. (1890). The colours of animals: Their meaning and use, especially considered in the case of insects. London, UK: Kegan Paul.

R Core Team (2019). *R: A language and environment for statistical computing* (Version 3.6.1). Retrieved from https://www.R-project.org

Rojas, B., Burdfield-Steel, E., Pakkonen, H., Suisto, K., Maczka, M., Schulz, S., & Mappes, J. (2017). How to fight multiple enemies: Target-specific chemical defences in an aposematic moth. *Proceedings of the Royal Society B: Biological Sciences*, 284(1863), 20171424. https://doi.org/10.1098/rspb.2017.1424

Rowe, C., Lindström, L., & Lyytinen, A. (2004). The importance of pattern similarity between Müllerian mimics in predator avoidance learning. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1537), 407–413. https://doi.org/10.1098/rspb.2003.2615

Rowell-Rahier, M., Pastels, J. M., Alonso-Mejía, A., & Brower, L. P. (1995). Relative unpalatability of leaf beetles with either biosynthesized or sequestered chemical defense. *Animal Behaviour*, 49, 709–714.
Rowland, H. M., Hoogesteger, T., Ruxton, G. D., Speed, M. P., & Mappes, J. (2010). A tale of 2 signals: Signal mimicry between aposematic species enhances predator avoidance learning. *Behavioral Ecology*, 21(4), 851–860. https://doi.org/10.1093/beheco/arp071

Rowland, H. M., Ihalainen, E., Lindström, L., Mappes, J., & Speed, M. P. (2007). Co-mimics have a mutualistic relationship despite unequal defences. *Nature*, 448(7149), 64–67. https://doi.org/10.1038/nature05899

Rowland, H. M., Wiley, E. M., Ruxton, G. D., Mappes, J., & Speed, M. P. (2010). When more is less: The fitness consequences of predators attacking more unpalatable prey when more are presented. *Biology Letters*, 6(6), 732–735. https://doi.org/10.1098/rsbl.2010.0207

Ruxton, G. D., Allen, W. L., Sherratt, T. N., & Speed, M. P. (2018). Avoiding attack (2nd ed.). Oxford, UK: Oxford University Press

Sherratt, T. N. (2006). Spatial mosaic formation through frequency-dependent selection in Müllerian mimicry complexes. *Journal of Theoretical Biology*, 240(2), 165–174. https://doi.org/10.1016/j.jtbi.2005.09.017

Sherratt, T. N. (2008). The evolution of Müllerian mimicry. *Naturewissenschaften*, 95(8), 681–695. https://doi.org/10.1007/s00114-008-0403-y

Sherratt, T. N. (2011). The optimal sampling strategy for unfamiliar prey. *Evolution*, 65(7), 2014–2025. https://doi.org/10.1111/j.1558-5646.2011.01274.x

Sherratt, T. N., Whissell, E., Webster, R. J., & Kikuchi, D. W. (2015). Hierarchical overshadowing of stimuli and its role in mimicry evolution. *Animal Behaviour*, 108, 73–79. https://doi.org/10.1016/j.anbehav.2015.07.011

Stuckert, A. M. M., Venegas, P. J., & Summers, K. (2014). Experimental evidence for predator learning and Müllerian mimicry in Peruvian poison frogs (Ranitomeya, Dendrobatidae). *Evolutionary Ecology*, 28(3), 413–426. https://doi.org/10.1007/s10682-013-9685-4

Summers, K., & Clough, M. E. (2001). The evolution of coloration and toxicity in the poison frog family (Dendrobatidae). *Proceedings of the National Academy of Sciences of the United States of America*, 98(11), 6227–6232. https://doi.org/10.1073/pnas.101134898

Svádová, K., Exnerová, A., Štys, P., Landova, E., Valenta, J., Fuciková, A., & Socha, R. (2009). Role of different colours of aposematic insects in learning, memory and generalization of naïve bird predators. *Animal Behaviour*, 77, 327–336. https://doi.org/10.1016/j.anbehav.2008.09.034

Symula, R., Schulte, R., & Summers, K. (2001). Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1484), 2415–2421. https://doi.org/10.1098/rspb.2001.1812

Therneau, T. M. (2019). *Mixed effects Cox models* (Version 2.2.14). Retrieved from https://cran.r-project.org/web/packages/coxme/coxme.pdf

Triponez, Y., Buerki, S., Borer, M., Naisbit, R. E., Rahier, M., & Alvarez, N. (2011). Discordances between phylogenetic and morphological patterns in alpine leaf beetles attest to an intricate biogeographic history of lineages in postglacial Europe. *Molecular Ecology*, 20(11), 2442–2463. https://doi.org/10.1111/j.1365-294X.2011.05096.x

Valkonen, J. K., Nokelainen, O., Niskanen, M., Kilpimaa, J., Björklund, M., & Mappes, J. (2012). Variation in predator species abundance can cause variable selection pressure on warning signalling prey. *Ecology and Evolution*, 2(8), 1971–1976. https://doi.org/10.1002/ece3.315

Vilmos, P., & Kurucz, É. (1998). Insect immunity: Evolutionary roots of the mammalian innate immune system. *Immunology Letters*, 62(2), 59–66. https://doi.org/10.1016/S0165-2478(98)00023-6

Waldron, S. J., Endler, J. A., Valkonen, J. K., Honma, A., Dobler, S., & Mappes, J. (2017). Experimental evidence suggests that specular reflectance and glossy appearance help amplify warning signals. *Scientific Reports*, 7(October 2016), 1–10. https://doi.org/10.1038/s41598-017-00217-5

Wallace, A. R. (1867). No Title. *Proceedings of the Entomological Society of London*, March 4th, XXX-XXXX

Williams, P. (2007). The distribution of bumblebee colour patterns worldwide: Possible significance for thermoregulation, crypsis, and warning mimicry. *Biological Journal of the Linnean Society*, 92(1), 97–118. https://doi.org/10.1111/j.1095-8312.2007.00878.x

Willmott, K. R., Willmott, J. C. R., Elias, M., & Jiggins, C. D. (2017). Maintaining mimicry diversity: Optimal warning colour patterns differ among microhabitats in Amazonian clearwing butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 284, 20170744. https://doi.org/10.1098/rspb.2014.09.005

Wilson, J. S., Williams, K. A., Forister, M. L., von Dohlen, C. D., & Pitts, J. P. (2012). Repeated evolution in overlapping mimicry rings among North American velvet ants. *Nature Communications*, 3, 1272. https://doi.org/10.1038/ncomms2275

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.