ORIGINAL PAPER

Effects of predator associative learning and innate aversion on mimicry complexes

Oliver T. Heerwig[1] · Sofia M. R. Jain-Schlaepfer[2] · Thomas N. Sherratt[2] ·
David W. Kikuchi[1,3]

Abstract

Undefended or weakly defended prey species can evolve to resemble better-defended prey (models) in a potentially parasitic relationship called Batesian mimicry. However, some highly defended prey have lethal defenses that might prevent predators from learning to avoid them, which raises questions as to how and why warning signals evolve in these species. One solution is that the warning signals of lethal species have evolved to resemble those of less defended species, where avoidance learning is possible. To examine the general feasibility of this hypothesis, we modeled associative learning by predators foraging on prey species that were either weakly or highly defended. The highly defended prey had a fixed probability of killing an attacking predator. We found that the weakly defended species was more likely to be a parasitic Batesian mimic when its defenses were weaker. Weakly defended prey were more parasitic when the weakly defended prey was relatively common, and when highly defended prey were less likely to be lethal. Generally, mimicry was more mutualistic (i.e. Müllerian) as the highly defended prey increased in lethality. However, for a relatively lethal mimetic mutant that resembled a weakly defended species to invade a non-mimetic population of highly defended prey, lethality needed to be high, as benefits of mimicry did not accrue at low frequency. Moreover, when we created predators that had innate aversions to dangerous prey, weakly defended mimics were parasitic upon highly defended prey. Innate aversions also evolved in an individual based-simulation. When our analyses of prey lethality and innate aversions are taken together, it is likely that highly defended prey are most often models in Batesian mimicry systems.

Keywords Batesian mimicry · Müllerian mimicry · Parasitism · Mutualism · Predator–prey · Associative learning

---

Oliver T. Heerwig and Sofia M. R. Jain-Schlaepfer contributed equally to this work.

✉ David W. Kikuchi
  david.kikuchi@oregonstate.edu

1 Evolutionary Biology Department, Universität Bielefeld, Konsequenz 45, 33615 Bielefeld, Germany

2 Department of Biology, Carleton University, 1125 Colonel By Dr., Ottawa, ON K1S 5B6, Canada

3 Department of Integrative Biology, Oregon State University, 4575 SW Research Way, Corvallis, OR 97333, USA

⌀ Springer

Introduction

Mimicry is widespread among animals. Two forms are Batesian mimicry and Müllerian mimicry (Ruxton et al. 2018). In Batesian mimicry, undefended or weakly defended prey species dupe predators into avoiding them by resembling prey species with stronger defenses (Bates 1862). In Müllerian mimicry, two or more defended prey species share a warning signal, reducing the number of each that is attacked before predators learn to avoid them (Müller 1879; Sherratt 2008). Batesian mimicry has been hypothesized to increase attack risk for the model, making the relationship parasitic (Speed 1993; Gavrilets and Hastings 1998; Franks et al. 2009). Müllerian mimicry is often regarded as mutualistic (Bronstein 2015; Aubier et al. 2017; Anderson and De Jager 2019). Even if one accepts these simple definitions (as we do in this study), not all mimetic relationships fall into the category that one might naively suspect (Speed 1993; Speed and Turner 1999; for detailed discussion, see Anderson and De Jager 2019).

Often, if one prey is weakly defended and the other highly defended, it can be difficult to determine whether their relationship is Batesian or Müllerian (e.g. Winters et al. 2018). In such cases, the response of predators to mixes of weakly and highly defended species is a critical determinant of the relationship (Rowland et al. 2007, 2010). Intriguingly, it has been proposed that when prey defenses are so strong that they are lethal, predators cannot learn to avoid them (Dunn 1954; Brattstrom 1955). Thus, counterintuitively, lethal prey may actually experience selection to resemble more weakly defended species (Mertens 1956; Hecht and Marien 1956; Emsley 1966; Wickler 1968). This scenario—sometimes called "Mertensian mimicry" or "Emsleyan mimcry" (Wickler 1968)—has received little empirical or theoretical investigation. It was originally inspired by coral snakes (Dunn 1954; Brattstrom 1955; Mertens 1956; Hecht and Marien 1956; Wickler 1968), but differential predator avoidance of lethal prey is found in other groups such as toxic newts (Kuchta 2005), which have been reported to cause the deaths of predators (Mobley and Stidham 2000). The mimetic relationship resulting from this unusual scenario of convergence could either be Müllerian or Batesian, in the sense that it may be mutualistic or parasitic (with the lethal mimics increasing attacks on the weakly defended models).

In addition to lethal prey evolving to mimic less lethal prey, another reason why predators might avoid lethal prey is innate aversions. This sidesteps the problem of associative learning. Predators have evolved innate aversions to common warning colors that signal unpalatability (Coppinger 1970; Schuler and Hesse 1985; Exnerová et al. 2007). Indeed, the coral snake colour patterns that inspired the ideas of Mertens (1956) elicited innate fear responses among naïve Neotropical birds when applied to wooden dowels (Smith 1975, 1977). Here, we used a simple model based on the modification of earlier models to address how (1) prey lethality and (2) innate aversions affect the evolution of mimetic relationships between highly and weakly defended prey.

Methods

Theoretical approach

As a departure point for exploring predator behavior, we use Speed's (1993) model of a predator that exhibits Pavlovian learning. The Pavlovian algorithm assumes that the

predator's tendency to attack prey with a specific phenotype increases if it is more palatable than the predator expects, and decreases if the prey is less palatable than the predator expects. We chose this model for its tractability and its foundation in psychological research (Rescorla and Wagner 1972). Other approaches based on identifying the optimal combination of exploration and exploitation are possible (Sherratt 2011; Kikuchi and Sherratt 2015; Aubier et al. 2017), but rapidly expand in complexity when many prey phenotypes are involved. Indeed, many of the solutions to exploration–exploitation models are so complex it is inevitable that predators would use a rule of thumb to resolve them (Sherratt 2016). The Rescorla-Wagner model has not only been identified as a reasonable approximation of associative learning but performs relatively well as a heuristic (Trimmer et al. 2012). As the main focus of our study is the response of predators to lethal prey, and the potential evolution of innate aversion, Speed's approach captures the necessary features we wish to include.

In Speed's (1993) model, a predator encounters a sequence of prey items from a range of different species. Each prey species j has palatability λ_j, where $0 < \lambda_j < 1$, ranging from completely unpalatable ($\lambda_j = 0$) to completely palatable ($\lambda_j = 1$). Each prey species also has an appearance index i. When i is shared between two species, mimicry occurs. Initially, prey of appearance i are attacked by the predator with probability $P_i = x_i$, where x_i represents the initial probability of attacking prey with appearance i. The model describes only the process of Pavlovian learning as the predator encounters prey with particular values of palatability λ_j. The λ_j value gives the extreme limit (i.e., closest to 0 or 1) towards which sampling prey of appearance i will move the predator probability of attack P_i.

In our default simulations, we consider all $x_i = 0.5$, so initially, there is a 50% chance to reject or attack a completely unfamiliar prey type. After attacking, P_i is updated according to the palatability of the prey item.

$$\Delta P_i = \alpha_j \left(\lambda_j - P_i \right) \tag{1}$$

where α_i controls the speed of learning and is calculated as:

$$\alpha_j = 0.5 + \left| \lambda_j - 0.5 \right| \tag{2}$$

Consequently, attacks on prey that are less/more palatable than the predator's current estimate P_i cause a decrease/increase in P_i since λ_j gives the most extreme probability of attack that experience with a prey item will cause a predator to exhibit (ΔP_i is 0 when $\lambda_j = P_i$). More extreme values of λ_j lead to faster rates of learning α_j. When two prey species share an appearance i, both of their values of λ_j contribute to predator learning. For example, if one has $\lambda_j = 0.4$ and another has $\lambda_j = 0.1$, the predator's probability of attacking prey P_i will tend to be bounded between those two values over time.

The model also incorporates predator forgetting. We assume that at each time step, a predator encounters a prey. If the predator does not attack the prey, all P_i are updated towards their initial attack rate x_i due to forgetting. If the prey is attacked, learning about that prey appearance i occurs, and P_i for that prey is updated according to (1) and (2), while forgetting occurs for all other prey appearances. Thus, without reinforcement, learned behavior reverts towards the initial attack rate. To calculate ΔP_i due to forgetting, we use the following equation, where α_f is the rate of forgetting:

$$\Delta P_i = \alpha_f \left(x_i - P_i \right) \tag{3}$$

We use these equations to model a species of predator foraging on two prey species with different levels of palatability. One species is highly unpalatable (*H*) and the other is weakly unpalatable (*W*). Because our primary interest is analyzing the conditions under which an exceptionally dangerous prey species can be selected to mimic a less defended prey species, we allow *H* to have two appearances. One, which is completely distinct from *W*, is H_d (*d* for distinct). The other, which exactly resembles *W*, is H_m (*m* for mimic).

In addition to being highly unpalatable, *H* can be lethal. The probability of predator death on attacking *H* is represented by $0 < \beta < 1$. We assume that when a predator dies, a new predator enters the system. This could represent a population of territorial birds, where the number of territories is fixed so there is always a reservoir of "floaters" without territories that we do not consider in the population (Fedy and Stutchbury 2004; Sergio et al. 2009). Implicitly, we also assume that prey disperse between predator territories, as the entire prey population is eligible for attack. Naive predators enter the system with all attack probabilities $P_i = x_i$, so they must learn anew about the palatability of any prey they encounter. In keeping with other assumptions from Speed (1993), we assume that predators encounter one prey per time step, and that if a prey item is attacked, it is always killed. We assume lethality of prey defenses is maintained by an alternative function, e.g. prey capture (Hossie et al. 2013).

Analysis of lethality on mimetic relationships

To quantify the relationship between *H* and *W* as Batesian or Müllerian as the lethality *β* varied, we compared the attack rates on *H* and *W* in two different scenarios. In the first scenario, only prey with H_d and *W* and appearances occurred—in other words, there was no mimicry. In the second scenario, only prey with H_m and *W* appearances occurred. This represents a situation where a mimicry complex has evolved between *H* and *W*. Predators were not shared between these two different scenarios. Comparing the attack rates on *H* and *W* in each scenario lets us measure whether the relationship between *H* and *W* is parasitic or mutualistic across a range of values for lethality (*β*). Initially, we averaged the output of 500 replicates of 100 time steps, where at each time step the predator encountered one prey (if the predator died, it was replaced with one whose attack probabilities were equal to x_i). Following Speed (1993), we let $\alpha_f = 0.02$. We first assumed an equal probability of encountering highly or weakly defended prey (sampled with replacement). We also experimented by varying the relative abundance of highly and weakly defended prey so that they were present in either a 1:4 ratio or a 4:1 ratio (Supplementary Figs. 2–3). We further examined the effect of letting $\alpha_f = 0.1$ (Supplementary Fig. 4), and an additional case where half of the prey were alternative prey that had no effect on predator behavior other than increasing the amount of forgetting between encounters with highly and weakly defended prey (Supplementary Fig. 5). In this latter case, the palatability of the alternative prey was irrelevant. These experiments with relative abundance, α_f, and alternative prey did not qualitatively change how lethality of the highly defended prey affected whether mimicry was Batesian or Müllerian (compare Supplementary Figs. 2–5 with Supplementary Fig. 1). Therefore, unless otherwise specified, all results we present are for 500 replicates with 100 time steps where highly and weakly defended prey were encountered at equal abundance, $\alpha_f = 0.02$, and the frequency of alternative prey was 0.

Changes between mutualism and parasitism between *H* and *W* are critical to how their mimetic relationship is classified. To better understand why prey lethality impacted mutualism versus parasitism, we evaluated the amount of time that predators spent learning about their prey, as opposed to foraging at an asymptotic rate governed by λ_j and α_f. To find the

asymptotic rate at which an experienced predator would sample prey of a given appearance, i.e. if it had encountered many prey without dying, we set lethality equal to 0 and discarded the first 50 prey items a predator encountered in any run of the experiment. Examination of the output showed that this was more than sufficient for predators to become "educated." We found the middle 95% of values for P_i that these educated predators exhibited. We defined this range as the asymptotic rate of attack. We then evaluated the impact of higher levels of lethality how much time a predator spent learning about prey of each appearance. We did this by calculating the proportion of prey encounters for which a predator's P_i values were outside the asymptotic range for attacking prey of that phenotype.

We needed to know under what circumstances it was possible for a mutation that causes mimicry to spread within prey populations. We analyzed invasion of mimetic alleles in the highly and weakly defended prey when mimicry did not initially occur between the two species. We did this by estimating the attack rate on a very rare H_m or W_m mutant (relative to the H_d or W_d nonmimetic phenotype, where W_d is a distinct, non-mimetic phenotype of W). We performed corresponding analyses of attack rates on a very rare H_d or W_m mutant (these analyses were performed separately, so only one species exhibited variation at a time). Additionally, assuming H_m or W_m could spread, we needed to test whether the whole population of H or W would evolve mimicry. This required analyzing the attack rate on the mutant versus the nonmimetic phenotype at a wide range of frequencies to ensure it would spread to fixation.

To explore conditions that might favor invasion of a mimetic morph, we varied the relative frequency of H_m or W_m in the population of H and W (versus H_d or W_d) from 0.01 to 99. For this analysis, we considered scenario where all H_d, H_m, and W (or W_d, W_m, and H) were exposed to the same predators. To obtain precise estimates of differences in attack rates, we arbitrarily ran the model for 100,000 time steps, and assumed that this allowed predators to encounter all prey in the community. We performed this analysis at increasing values of lethality until H_m or W_m could invade. To calculate the relative fitness of H_m and W_m, the per capita attack rate on H_m or W_m on encounter was divided by the per capita attack rate on H_d or W_d on encounter. Values above 1 indicate that H_m or W_m would spread, whereas values below 1 indicate that H_m or W_m would decrease in frequency. We examined situations where H and W had equal relative abundance, where H outnumbered W 4:1, and where W outnumbered H 4:1.

Finally, we explored different assumptions about innate attack rates on prey of different appearances. We examined different a priori assumptions about the initial probability of attack x_i by setting it to lower values and repeating analyses of parasitism and mutualism. We also allowed values of x_i to evolve in a simulation (see Supplementary Methods for details).

Results

We compared the relative attack rates on different types of prey as lethality varied. We found that the mimetic relationship between the highly defended prey H and weakly defended prey W was either mutualistic, or parasitic with W benefitting from H (Fig. 1A; Figures S1–S5). The relationship transitioned from parasitism to mutualism as the lethality of H (β) increased (Fig. 1A). In Fig. 1A, this is depicted where the rate of attack on H involved in mimicry (red line) falls below the rate of attack of non-mimetic H (brown line). The mutualistic relationship arose because as H became more likely to kill their predators, the naive predators that entered the population initially attacked at the rate of $x_i = 0.5$. Highly defended prey increasingly benefitted from the deterrent effects of W, which trained the predators to attack at a lower rate without killing them.
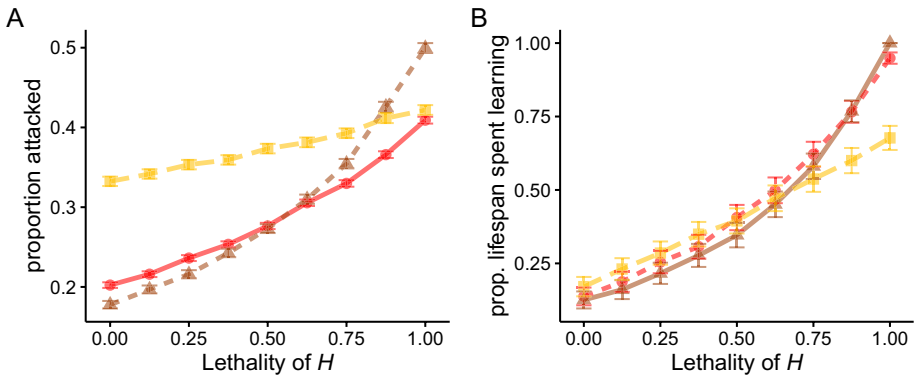
Fig. 1 Effect of lethality on predator attack rates on encounter with each type of prey and predator learning. The yellow squares/long dashed line (W) and brown triangles/short-dashed line (H_h) represent prey species co-occurring in the same environment, but without mimicry. The red circles/solid line represent the shared appearance of two species in a mimicry complex. The densities of the two species are equal in both situations. **A** Whenever the red line is below both brown and yellow lines, mimicry is mutualistic. When the red line is between the two other lines, the species described by the top line is a parasitic mimic of the species described by the bottom line (here, only the weakly defended prey is ever parasitic). Note that predation on W increases with lethality even without mimicry because predators that die after attacking H_h must learn about both prey anew from $x_{Hh} = x_W = 0.5$. **B** The amount of time a predator spends learning about each type of prey (as a proportion of the predator's lifespan) increases with prey lethality, because the predator population turns over more quickly and its initial attack rate is set to x_i. In both panels, $\lambda_W = 0.3$, $\lambda_H = 0.1$, and 100 highly and weakly defended prey are encountered in scenarios with and without mimicry. Bars are 95% binomial confidence intervals

We illustrate this effect in Fig. 1B, which shows that predators spent a greater proportion of their lifespans learning about prey phenotypes when β was higher (because they died more often), rather than attacking at rates consistent with educated foragers. The transition from parasitism to mutualism was robust across different levels of unpalatability λ_W, although as λ_W decreased, the area where W were parasitic vanished, leaving only a mutualistic relationship (Figure S1). Additionally, when H were much rarer than W, the relationship was always mutualistic (e.g. Figure S3).

Our invasion analysis revealed that the conditions for a rare mutant H_m to spread required a high value of β (Fig. 2A). The value of β required was higher than might be expected from examining Fig. 1, where attack rates on a non-mimetic population of H are compared with a completely mimetic population. This discrepancy arose because in the invasion analysis shown in Fig. 2A, the rare mutant H_m sacrifices the protection provided by the remainder of the highly defended population H_d, and obtains protection only from W. A sensitivity analysis revealed that as λ_W increased (λ_H held constant), the threshold value of β for H_m to invade increased. Recall that the effect of raising β was similar to that of increasing the rate at which predators forgot about H. When β was low and λ_W was fairly high, H was usually better protected by itself than in a mimicry complex (Fig. 1A). We found no evidence of selection against the mimetic H_m mutant as its frequency increased, so if it could invade, it would spread to fixation in H (Fig. 2A). Invasion thresholds were decreased by a lower relative abundance of H (Figures S6A, S7A).

Our corresponding invasion analysis of a mimetic mutant in W revealed a broader range of conditions where invasion and fixation could occur than in H (Figs. 2B, S6B, S7B). When both species were equally abundant or H outnumbered W, invasion of W_m was favored at lower
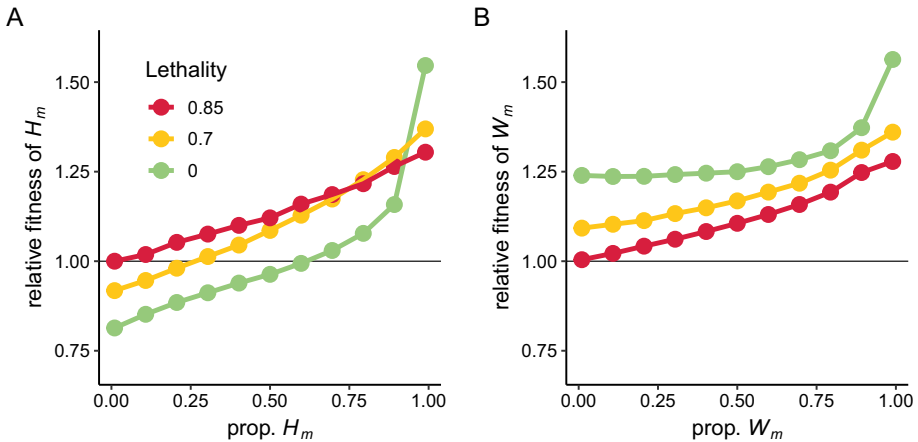
Fig. 2 Invasion and fixation analysis across different levels of lethality when H and W are equally abundant. **A** Only when lethality is relatively high is invasion by H_m possible, although if it can invade, it will fix. **B** Invasion and fixation of W_m is always possible across the range of lethality values examined. Fitness is calculated as the relative per capita attack rate on encounter of H_m/H_h and W_m/W_h. Palatabilities are $\lambda_W = 0.3$, $\lambda_H = 0.1$. 95% confidence intervals were smaller than the points, and so omitted

values of lethality β than was invasion of H_m (Figs. 2B, S6B). Only when W outnumbered H were there conditions where invasion of W_m was not favored (Figure S7B).

Innate aversion towards H (Fig. 3A) or towards prey in proportion to their unpalatability (Fig. 3B) caused W to be parasitic on H when they were in a mimicry complex. The stronger the innate aversion towards H, the less often predators died, and so the greater the benefits to H. When innate aversions towards the mimicry complex and H_d were very strong, increasing lethality actually *decreased* attack rates on those types of prey (Fig. 3C). This is an artefact of the modeling framework, because values of the initial attack rate x_i were less than values of λ_i for those prey types. However, it did not change the relationship between the prey species from parasitism to mutualism.

When we used individual-based simulation to see whether innate aversions that we selected above would arise, we found that lethal H reliably selected for x_i values near 0 in predators that were exposed to them (Figure S8). When this happened, W benefitted from resemblance to H, but H suffered increased attacks (Figure S9).

Discussion

In this study, we addressed the problem of how predators respond to warning signals when prey defenses can kill a predator. Using a model of associative learning (Speed 1993), we examined the roles of acquired and innate preferences on the relationships between highly and weakly defended prey species. We found that when predators have equal (i.e. $x_i = 0.5$) initial tendencies to attack both prey species, the relationship between the two prey can be mutualistic, i.e. Müllerian, particularly when the highly defended prey is lethal. However,
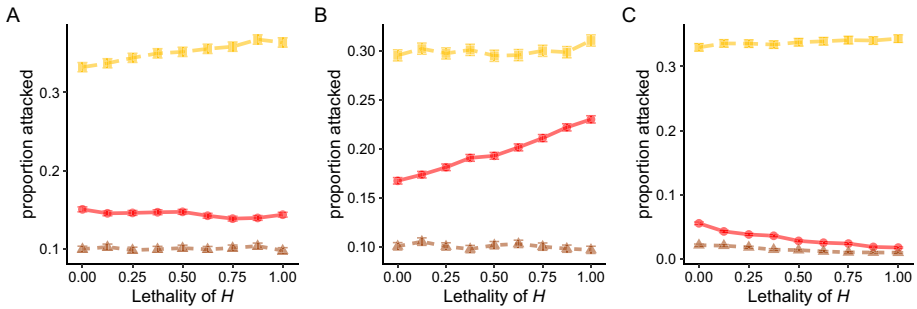
Fig. 3 The effect of different assumptions about innate aversions that predators may have to model-mimicry complexes, highly defended non-mimetic prey, and weakly defended non-mimetic prey. The yellow squares/long dashed line (W) and brown triangles/short-dashed line (H_h) represent prey species co-occurring in the same environment, but without mimicry. The red circles/solid line represent the shared appearance of two species in a mimicry complex. **A** The predator evolves an innate aversion to the mimicry complex and highly defended non-mimetic prey ($x_m = x_{Hh} = 0.1$; $x_W = 0.5$). **B** The predator evolves innate aversion in proportion to the average unpalatability associated with a prey's appearance. Here $x_m = 0.2$, $x_{Hh} = 0.1$, and $x_W = 0.3$. **C** The potential lethality of highly defended prey causes predators to evolve nearly complete avoidance of the mimicry complex and highly defended non-mimetic prey, such $x_m = x_{Hh} = 0.01$. No innate aversion towards the weakly defended non-mimetic prey occurs ($x_W = 0.5$). In all panels, $\lambda_W = 0.3$, $\lambda_H = 0.1$, and 100 highly and weakly defended prey are encountered in scenarios with and without mimicry. Bars are 95% binomial confidence intervals

a rare, mimetic mutant can invade the highly defended species are only favored at higher values of lethality. Furthermore, invasion of a mimetic mutant in the *weakly* defended species occurs under a wider range of conditions. The only exception to this is if the weakly defended prey greatly outnumber the highly defended prey, a common result in analyses of mimicry due to the safety in numbers provided by aposematism (Ruxton et al. 2004; Sherratt 2008). Thus, if mimicry evolves, it is more likely to happen with the weakly defended species initially evolving towards the appearance of the highly defended species except in more lethal cases. Predators that have evolved to innately avoid potentially lethal prey decrease the potential for a relationship between weakly and highly defended prey to be mutualistic. Instead, our work predicts that a conventional Batesian mimicry relationship prevails where the weakly defended prey benefits from resembling the highly defended prey since they are avoided from the outset of predator learning.

Coral snake mimicry complexes provided the initial impetus for discussing the relationship between weakly defended prey and prey with potentially lethal defenses (Dunn 1954; Brattstrom 1955; Mertens 1956; Hecht and Marien 1956; Wickler 1968). At present, however, evidence suggests that the deadly coral snakes are mimicked by weakly defended or undefended colubrid snakes, and not the other way around. Experiments with replicas show that local coral snake color patterns are avoided by sympatric predators (Brodie 1993; Hinman et al. 1997; Pfennig et al. 2001), yet in allopatry from coral snakes, attack rates on mimics often increase (Pfennig et al. 2001, 2007). Experiments have shown that some sympatric predator species exhibit innate aversions to coral snakes (Smith 1975, 1977), but not all (Beckers et al. 1996; Sherbrooke and Westphal 2006). Phylogenetic analyses find that coral snakes arrived in the New World before the evolution of warning signals in colubrid snakes (Davis Rabosky et al. 2016). These lines of evidence all support coral snakes as the contemporary models in Batesian mimicry complexes.

Some have used the term "Mertensian mimicry" to describe Mertens (1956) and Hecht and Marien's (1956) hypothesis that lethally defended prey that evolve to mimic weakly

defended prey (Wickler 1968), but not all have agreed on its utility (Greene and McDiarmid 1981). It is somewhat poorly defined because it was originally proposed as an explanation coral snake mimicry, so any discussion of it has been interwoven with that particular empirical system. Ideally, a hypothesis is a theoretical construct whose applicability is supported or rejected on the basis of evidence—the hypothesis should be conceptually separate from its application to a particular instantiation. In light of our modeling, we prefer to think of what Mertens (1956) and Hecht and Marien (1956) described as a process in which a highly defended prey evolves to resemble a more weakly defended one, specifically when highly defended prey is less effective at promoting associative learning by the predator because it causes lethality. The endpoints of this process could potentially be either Batesian mimicry (parasitic) or Müllerian mimicry (mutualistic), although our model predicts that if this process occurs, it results only in Müllerian mimicry.

A process like that envisioned by Mertens (1956) and Hecht and Marien (1956) has the potential to occur in some systems. Among poison frogs, it appears that predators more readily learn to avoid less-defended poison frog species (Darst and Cummings 2006), and the toxic *Dendrobates imitator* appears to have diverged to resemble less toxic models across its range (Symula et al. 2001). This system is thought to form a Müllerian mimicry ring (Rojas 2017). We note, however, that the pattern of convergence by *D. imitator* on other species could be due to pressure to resemble multiple models (Sherratt 2002), rather than less defended ones. Among some chemically defended nudibranchs that form mimicry rings, distastefulness of a compound (related to its efficacy in promoting associative learning) is poorly correlated with its toxicity (Winters et al. 2018, 2022). This also raises the possibility that highly defended nudibranchs may experience selection to mimic more weakly defended nudibranchs, although we do not know if nudibranchs could kill their predators prior to learning. Among neither poison frogs nor nudibranchs is lethality of toxins to relevant predators well-established.

A major advance by our model is considering the role of innate aversion in mimetic relationships. If a highly defended species evolves to resemble a less defended species when predators learn associatively, our model predicts that it could produce a Müllerian mimicry complex (Fig. 1; Figures S1-S5). Subsequently, predators experiencing selection for innate aversion could change the relationship to one of classical Batesian mimicry (as mimicry between unequally defended prey can be parasitic; this has been called "quasi-Batesian mimicry"; Speed and Turner 1999; Rowland et al. 2010; Fig. 3, S9). Alternatively, the evolution of innate aversion could proceed *prior* to the evolution of mimicry. Predators that detect, sample, and die from attacking deadly prey would experience selection to avoid them. Once innate aversion evolved in predators, it could catalyze the subsequent evolution of Batesian mimicry. The evolution of innate aversion is widespread and raises many questions, such as how warning signals evolve when they are not under selection for their function in associative learning. The caveats of the modeling framework we used here make it less than ideal for delving into the diversity of detailed ways in which innate aversion could evolve, but optimal foraging approaches hold much potential for future work.

There are a number of intriguing systems where innate recognition of deadly prey has been documented. They include venomous snakes such as coral snakes (Smith 1975, 1977), sea snakes (Caldwell and Rubinoff 1983), and vipers (Sherbrooke and Westphal 2006), which are innately avoided by naive birds. Predatory jumping spiders innately avoid ants (Nelson et al. 2006; Nelson and Jackson 2006). Ants carry a potent sting that could be lethal to spiders, and are mimicked by an entire group of salticids called Myrmarachnae (Edmunds 2006; Pekár 2014). However, other predators such as lizards learn to associate ant-like appearances with prey defense (Pekár et al. 2017). The relative

importance of associative learning versus innate aversion in the evolutionary dynamics of ant mimicry systems remains to be elucidated.

There are many important phenomena in predator psychology that fall beyond the scope of this model. One is predator neophobia (fear of unfamiliar stimuli), which may initially favor conspicuous warning color patterns (Lindström et al. 1999; Exnerová et al. 2007). This might interact with selection for innate aversion to help explain why lethally defended prey should evolve conspicuous rather than inconspicuous warning signals. Social learning provides yet another potential route to avoidance of lethal prey—if a predator observes another predator fall suffer or die from interacting with a prey, it may be less likely to attempt to attack the prey itself (Thorogood et al. 2018). Primates have been shown to exhibit biased social learning about snakes, learning to avoid snakes more quickly than other stimuli when conspecifics show fear (Cook and Mineka 1989).

Generally speaking, we need more research into the interplay of associative learning, generalized responses such as neophobia, and specific innate aversions to particular stimuli in the evolution of warning signals. Future work into the coevolution of predator behavior and prey phenotype is warranted.

Data availability Not applicable.

Code availability All code is available as an online supplement.

Declarations

Competing interests The authors declare no competing interests.

Conflicts of interest We have no conflicts of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Anderson B, De Jager M (2019) Natural selection in mimicry. Biolo Rev. https://doi.org/10.1111/brv.12564

Aubier TG, Joron M, Sherratt TN (2017) Mimicry among unequally defended prey should be mutualistic when predators sample optimally. Am Nat 189:267–282. https://doi.org/10.1086/690121

Bates HW (1862) XXXII. Contributions to an insect fauna of the Amazon Valley. Lepidoptera: Heliconidæ. Trans LinnSoc Lond 23:495–566. https://doi.org/10.1111/j.1096-3642.1860.tb00146.x

Beckers GJL, Leenders TAAM, Strijbosch H (1996) Coral snake mimicry: live snakes not avoided by a mammalian predator. Oecologia 106:461–463. https://doi.org/10.1007/BF00329702

Brattstrom BH (1955) The coral snake "mimic" problem and protective coloration. Evolution 9:217–219. https://doi.org/10.2307/2405591

Brodie EDI (1993) Differential avoidance of coral snake banded patterns by free-ranging predators in Costa Rica. Evolution 47:227–235

Bronstein JL (2015) The study of mutualism. In: Bronstein JL (ed) Mutualism. Oxford University Press, Oxford

Caldwell GS, Rubinoff RW (1983) Avoidance of venomous sea snakes by naive herons and egrets. Auk 100:195–198

Cook M, Mineka S (1989) Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. J Abnorm Psychol 98:448–459. https://doi.org/10.1037/0021-843x.98.4.448

Coppinger RP (1970) The effect of experience and novelty on avian feeding behavior with reference to the evolution of warning coloration in butterflies. II. Reactions of naive birds to novel insects. Am Nat 104:323–335. https://doi.org/10.1086/282666

Darst CR, Cummings ME (2006) Predator learning favours mimicry of a less-toxic model in poison frogs. Nature 440:208–211. https://doi.org/10.1038/nature04297

Davis Rabosky AR, Cox CL, Rabosky DL et al (2016) Coral snakes predict the evolution of mimicry across new world snakes. Nat Commun 7:11484. https://doi.org/10.1038/ncomms11484

Dunn ER (1954) The coral snake "mimic" problem in Panama. Evolution 8:97–102

Edmunds M (2006) Do Malaysian Myrmarachne associate with particular species of ant? Biol J Lin Soc 88:645–653. https://doi.org/10.1111/j.1095-8312.2006.00649.x

Emsley MG (1966) The mimetic significance of *Erythrolamprus aesculapii ocellatus* Peters from Tobago. Evolution 20:663. https://doi.org/10.2307/2406599

Exnerová A, Štys P, Fučíková E et al (2007) Avoidance of aposematic prey in European tits (Paridae): learned or innate? Behav Ecol 18:148–156. https://doi.org/10.1093/beheco/arl061

Fedy BC, Stutchbury BJM (2004) Territory switching and floating in the white-bellied antbird (*Myrmeciza longipes*), a resident tropical passerine in Panama. Auk 121:11

Franks DW, Ruxton GD, Sherratt TN (2009) Warning signals evolve to disengage Batesian mimics. Evolution 63:256–267. https://doi.org/10.1111/j.1558-5646.2008.00509.x

Gavrilets S, Hastings A (1998) Coevolutionary chase in two-species systems with applications to mimicry. J Theor Biol 191:415–427. https://doi.org/10.1006/jtbi.1997.0615

Greene HW, McDiarmid RW (1981) Coral snake mimicry: does it occur? Science 213:1207–1212

Hecht MK, Marien D (1956) The coral snake mimic problem: a reinterpretation. J Morphol 98:335–365. https://doi.org/10.1002/jmor.1050980207

Hinman KE, Throop HL, Adams KL et al (1997) Predation by free-ranging birds on partial coral snake mimics: the importance of ring width and color. Evolution 51:1011–1014

Hossie TJ, Hassall C, Knee W, Sherratt TN (2013) Species with a chemical defence, but not chemical offence, live longer. J Evol Biol 26:1598–1602. https://doi.org/10.1111/jeb.12143

Kikuchi DW, Sherratt TN (2015) Costs of learning and the evolution of mimetic signals. Am Nat 186:321–332. https://doi.org/10.1086/682371

Kuchta SR (2005) Experimental support for aposematic coloration in the salamander *Ensatina eschscholtzii xanthoptica*: implications for mimicry of Pacific newts. Copeia 2005:265–271. https://doi.org/10.1643/CH-04-173R

Lindström L, Alatalo RV, Mappes J et al (1999) Can aposematic signals evolve by gradual change? Nature 397:249–251. https://doi.org/10.1038/16692

Mertens R (1956) Das problem der mimikry bei korallenschlangen. Fischer, Waldachtal

Mobley JA, Stidham TA (2000) Great Horned Owl feath from predation of a toxic California newt. Wilson Bull 112:563–564. https://doi.org/10.1676/0043-5643(2000)112[0563:GHODFP]2.0.CO;2

Müller F (1879) *Ituna* and *Thyridia*: a remarkable case of mimicry in butterflies. Trans Entomol Soc Lond 770:20–29

Nelson XJ, Jackson RR (2006) Vision-based innate aversion to ants and ant mimics. Behav Ecol 17:676–681. https://doi.org/10.1093/beheco/ark017

Nelson XJ, Jackson RR, Li D et al (2006) Innate aversion to ants (*Hymenoptera*: *Formicidae*) and ant mimics: experimental findings from mantises (*Mantodea*). Biol J Lin Soc 88:23–32. https://doi.org/10.1111/j.1095-8312.2006.00598.x

Pekár S (2014) Is inaccurate mimicry ancestral to accurate in myrmecomorphic spiders (*Araneae*)? Biol J Linn Soc Lond 113:97–111. https://doi.org/10.1111/bij.12287

Pekár S, Petráková L, Bulbert MW et al (2017) The golden mimicry complex uses a wide spectrum of defence to deter a community of predators. eLife 6:e22089. https://doi.org/10.7554/eLife.22089

Pfennig DW, Harcombe WR, Pfennig KS (2001) Frequency-dependent Batesian mimicry. Nature 410:323–323. https://doi.org/10.1038/35066628

Pfennig DW, Harper GR, Brumo AF et al (2007) Population differences in predation on Batesian mimics in allopatry with their model: selection against mimics is strongest when they are common. Behav Ecol Sociobiol 61:505–511. https://doi.org/10.1007/s00265-006-0278-x

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF (eds) Classical conditioning II. Appleton-Century-Crofts, Norwalk, pp 64–99

Rojas B (2017) Behavioural, ecological, and evolutionary aspects of diversity in frog colour patterns. Biol Rev 92:1059–1080. https://doi.org/10.1111/brv.12269

Rowland HM, Ihalainen E, Lindström L et al (2007) Co-mimics have a mutualistic relationship despite unequal defences. Nature 448:64–67. https://doi.org/10.1038/nature05899

Rowland HM, Mappes J, Ruxton GD, Speed MP (2010) Mimicry between unequally defended prey can be parasitic: evidence for quasi-Batesian mimicry. Ecol Lett 13:1494–1502. https://doi.org/10.1111/j.1461-0248.2010.01539.x

Ruxton GD, Sherratt TN, Speed MP (2004) Avoiding attack. Oxford University Press, Oxford

Ruxton GD, Allen WL, Sherratt TN, Speed MP (2018) Avoiding attack, 2nd edn. Oxford University Press, Oxford

Schuler W, Hesse E (1985) On the function of warning coloration: a black and yellow pattern inhibits prey-attack by naive domestic chicks. Behav Ecol Sociobiol 16:249–255. https://doi.org/10.1007/BF00310988

Sergio F, Blas J, Hiraldo F (2009) Predictors of floater status in a long-lived bird: a cross-sectional and longitudinal test of hypotheses. J Anim Ecol 78:109–118. https://doi.org/10.1111/j.1365-2656.2008.01484.x

Sherbrooke WC, Westphal MF (2006) Responses of greater roadrunners during attacks on sympatric venomous and nonvenomous snakes. Southwest Nat 51:41–47. https://doi.org/10.1894/0038-4909(2006)51[41:ROGRDA]2.0.CO;2

Sherratt TN (2002) The evolution of imperfect mimicry. Behav Ecol 13:821–826

Sherratt TN (2008) The evolution of Müllerian mimicry. Naturwissenschaften 95:681–695. https://doi.org/10.1007/s00114-008-0403-y

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

Sherratt TN (2016) State-dependent optimization: a comment on Skelhorn et al. Behav Ecol 27:965–965. https://doi.org/10.1093/beheco/arw024

Smith SM (1975) Innate recognition of coral snake pattern by a possible avian predator. Science 187:759–760. https://doi.org/10.1126/science.187.4178.759

Smith SM (1977) Coral-snake pattern recognition and stimulus generalisation by naive great kiskadees. Nature 265:536–536

Speed MP (1993) Mullerian mimicry and the psychology of predation. Anim Behav 45:571–580

Speed MP, Turner JRG (1999) Learning and memory in mimicry: II. Do we understand the mimicry spectrum? Biol J Lin Soc 67:281–312. https://doi.org/10.1111/j.1095-8312.1999.tb01935.x

Symula R, Schulte R, Summers K (2001) Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. Proc R Soc Lond B 268:2415–2421. https://doi.org/10.1098/rspb.2001.1812

Thorogood R, Kokko H, Mappes J (2018) Social transmission of avoidance among predators facilitates the spread of novel prey. Nat Ecol Evol 2:254–261. https://doi.org/10.1038/s41559-017-0418-x

Trimmer PC, McNamara JM, Houston AI, Marshall JAR (2012) Does natural selection favour the Rescorla-Wagner rule? J Theor Biol 302:39–52. https://doi.org/10.1016/j.jtbi.2012.02.014

Wickler W (1968) Mimicry in plants and animals. George Weidenfeld and Nicholson Ltd., London

Winters AE, Wilson NG, van den Berg CP et al (2018) Toxicity and taste: unequal chemical defences in a mimicry ring. Proc R Soc b Biol Sci 285:20180457. https://doi.org/10.1098/rspb.2018.0457

Winters AE, Chan W, White AM et al (2022) Weapons or deterrents? Nudibranch molluscs use distinct ecological modes of chemical defence against predators. J Anim Ecol 91:831–844. https://doi.org/10.1111/1365-2656.13643