Similar predator aversion for natural prey with diverse toxicity levels

Mathieu Chouteau, Jules Dezeure, Thomas Sherratt, Violaine Llaurens,
Mathieu Joron

To cite this version:


HAL Id: hal-02148400
https://hal.archives-ouvertes.fr/hal-02148400
Submitted on 5 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Title: Similar predator aversion for natural prey with diverse toxicity levels

Running title: Prey defense in mimicry

Author: Mathieu Chouteau¹*, Jules Dezeure², Thomas Sherratt³, Violaine Llaurens², Mathieu Joron⁴

¹Laboratoire Ecologie, Evolution, Interactions Des Systèmes Amazoniens (LEEISA), USR 3456, Université De Guyane, CNRS Guyane, IFREMER Guyane, 275 route de Montabo, 97334 Cayenne, French Guiana.
²Institut de Systématique, Evolution, Biodiversité (ISYEB), UMR 7205 CNRS-MNHN-UPMC-EPHE, Muséum national d'Histoire naturelle, CP 50, 45 rue Buffon, 75005 Paris, France.
³Department of Biology, Carleton University, Ottawa, Ontario K1S 5B6, Canada
⁴Centre d’Ecologie Fonctionnelle et Evolutive (CEFE), UMR 5175 CNRS - Université de Montpellier – EPHE - Université Paul Valéry, 1919 route de Mende, 34293 Montpellier 5, France.

*Corresponding author: mathieu.chouteau@cnrs.fr
Abstract:

Müllerian mimicry between chemically-defended species arises from selection exerted by predators in which individuals benefit from higher survival when they share the same warning signal. However, despite sharing warning signals, co-mimetic species harbor a diversity of toxins at a range of different concentrations. This variation may affect the rate of predator avoidance learning and therefore the dynamics of mimicry. Here, to understand the nature of mimetic relationships in natural communities of butterflies and moths, we compared protection against predators induced by chemical defenses of 13 lepidopteran species belonging to 6 mimicry complexes. Protection was estimated by quantifying the extent of avoidance learning, using domestic chicks as model predators. We showed that most co-mimics were avoided at similar high rates, with the exception of two species eliciting markedly slower rates of avoidance. Assuming our model and natural predators behave similarly and cannot distinguish co-mimics visually, the similar avoidance learning they induce supports the contention that mutualistic relationships among these co-mimetic species might be predominant in natural communities, despite large variations in toxins concentrations. Indeed, by comparing our estimated avoidance learning rate to mean toxin concentrations, we showed that prey with two- to three-fold differences in toxin content generate similar avoidance learning indices. This lack of direct relationship between prey defense level and predator avoidance learning points to alternative evolutionary mechanisms promoting the evolution of high levels of toxins.

Keyword: aposematism, toxicity, predator aversion, Müllerian mimicry, quasi-Batesian mimicry
Introduction

Müllerian mimicry is an iconic example of evolutionary convergence in which multiple defended prey within a locality evolve similar warning signals. This convergence hinges on the survival benefit associated with the increased abundance of a warning signal. Sharing a warning signal effectively spreads the mortality involved in sampling prey during the education of predators over all co-mimics, resulting in a reduced per capita predation risk (Mallet, 1999; Sherratt, 2008).

The extent to which predators sample unprofitable prey with a given signal depends not just on their innate prior beliefs but also their rate of avoidance learning, which may be higher for prey of higher unprofitability (Skelhorn & Rowe 2006a). Therefore, the relative contributions of otherwise indistinguishable co-mimics to predator learning may vary depending on their respective defenses, resulting in either a parasitic or mutualistic relationship (Speed, 1993). The role of variation in prey defenses in the evolutionary dynamics of mimicry has been the focus of much experimental research on predator cognition and theoretical modelling of mimicry systems (Mallet, 2001; Sherratt, 2008; Skelhorn et al., 2016; Ruxton et al., 2018). Fully palatable species clearly act as parasites to their unpalatable (or otherwise unprofitable) models, but the situation is more complex when co-mimetic prey species show different levels of defenses (Rowland et al., 2007; Sherratt, 2008). Indeed, the fitness consequences of differences in defense levels among co-mimics are not trivial to determine, because avoidance learning is affected both by unpalatability (the perception of prey defenses by predators) and by the signal density associated with mimicry (Mallet, 1999; Chouteau et al., 2016; Aubier et al., 2017).

Some studies have suggested that prey species eliciting intermediate avoidance learning by predators act as parasites when mimicking better defended species because they reduce the overall rate of predator associative learning (i.e. quasi-Batesian mimicry (Speed, 1993; Rowland et al., 2010)). Other studies have instead suggested that mimics eliciting intermediate avoidance learning contribute more to predator education through increased warning signal abundance than they disturb the association through lower unpalatability, resulting in overall mutualistic interactions (Müllerian mimicry (Müller, 1878; Mallet, 1999; Rowland et al., 2007; Aubier et al., 2017; Pekar et al., 2017)).

To understand the consequences of variation in prey defences, laboratory experiments have been performed with model predators such as domestic chicks and great tits presented with prey items
containing toxins at varying concentrations (e.g. low vs. high concentrations of quinine sulfate, or dedicated animal repellents such as denatonium benzoate). The concentrations used for such experiments were chosen according to preliminary assays confirming that predators could discern them, as shown by fast vs. slow avoidance learning (Huheey, 1976; Speed, 1993; Mallet, 1999; Lindstrom et al., 2006; Skelhorn & Rowe, 2006b; Balogh et al., 2008; Ihalainen et al., 2008). Those experiments constitute a rich literature dissecting the relationships between prey with highly contrasted defense levels. However, the range of variation in the deterrents used in experiments are still difficult to relate to the chemical defenses shown by natural prey species engaged in mimicry relationships in nature. Indeed, experimental vs. natural toxins are different, but more importantly, we continue to ignore to what extent prey species from a natural mimetic community vary in their individual deterrence to predators.

In wild prey, chemical defenses may be synthesized de novo, or sequestered from various sources and used as key precursors or directly stored as toxic compounds. It is known that both the nature of defensive compounds and their amounts vary tremendously within and between mimetic species (Speed et al., 2012; Ruxton et al., 2018). For instance, Ithomiini butterflies are defended by pyrrolizidine alkaloids (Trigo & Brown, 1990), while co-mimics of the genus Heliconius contain cyanogenic glycosides (Engler-Chaouat & Gilbert, 2007). Whether this diversity in chemical defenses translates into different learning rates by predators, and hence unequal levels of protection against predators remains untested, but constitutes a key link between laboratory studies and our understanding of natural prey communities.

In the present study, we perform experiments using prey items made from wild-caught samples from species belonging to mimetic communities of chemically-defended butterflies and moths, to determine the distribution of their deterrent effect, through the perception of a model predator. We then investigate how such perception by model predators relates to their toxin content.
Methods

Species and mimicry rings

We compared the protection against predation enabled by chemical defenses carried by 13 mimetic species of butterflies and moths from North-Eastern Peru (in the surrounding of Tarapoto – San Martin department), participating in 6 distinct mimicry rings (five species display several morphs, each one mimicking a different community). These species belong to three distantly related clades (Heliconiinae, Danainae and Pericopinae) and are known to possess chemical defenses of different nature (see Table 1 for sample details). As a non-mimetic palatable reference, samples from Anartia amathea (Nymphalinae) were used (DeVries, 1987). Butterflies were field-collected and kept alive in a 4.0L x 3.0W x 2.5H meters outdoor insectary in Tarapoto with ad libitum access to flowers and sugar water. The butterflies were processed as soon as possible (within 12h after collection) and used fresh for the experiment described below.

Table 1: Variation in protection levels among defended prey species measured with domestic chicks. Here we list the butterfly species tested along with their systematic grouping and mimicry ring(s) (note that some species can display different wing phenotypes and therefore belong to several mimicry rings). We then list the nature of their chemical defense, the number of butterflies of each type tested (nb), their mean CG concentration per mg of dry weight and the estimated CG concentration of crumbs made with their bodies when applicable (mean ± standard error). We also list the number of chicks experimented on (nE), the estimated mean number of experimental crumbs eaten by these chicks before avodiance is complete (mean E +/- SE) and the mean total number of experimental crumbs actually eaten (mean T +/- SE).

<table>
<thead>
<tr>
<th>Species subfamily</th>
<th>Mimicry ring</th>
<th>Chemical nb</th>
<th>Concentration (μg/mg)</th>
<th>Crumbs GC (μg/crumb)</th>
<th>nE</th>
<th>E ± SE</th>
<th>T ± SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heliconius erato</td>
<td>Heliconiinae</td>
<td>Postman</td>
<td>38</td>
<td>0.75 ± 0.13</td>
<td>4.85 ± 0.18</td>
<td>6</td>
<td>26.90 ± 6.49</td>
</tr>
<tr>
<td>Heliconius numata</td>
<td>Heliconiinae</td>
<td>Tricolor</td>
<td>33</td>
<td>2.17 ± 0.12</td>
<td>14.03 ± 0.57</td>
<td>6</td>
<td>8.19 ± 4.16</td>
</tr>
<tr>
<td>Heliconius burneyi</td>
<td>Heliconiinae</td>
<td>Dennis-ray</td>
<td>8</td>
<td>2.50 ± 0.23</td>
<td>16.17 ± 1.34</td>
<td>6</td>
<td>10.72 ± 2.66</td>
</tr>
<tr>
<td>Heliconius melpomene</td>
<td>Heliconiinae</td>
<td>Postman</td>
<td>24</td>
<td>2.53 ± 0.16</td>
<td>16.36 ± 0.78</td>
<td>6</td>
<td>6.95 ± 1.79</td>
</tr>
<tr>
<td>Heliconius ethilia</td>
<td>Heliconiinae</td>
<td>Spotted</td>
<td>7</td>
<td>2.69 ± 0.51</td>
<td>17.39 ± 1.55</td>
<td>6</td>
<td>9.83 ± 1.85</td>
</tr>
<tr>
<td>Heliconius aeide</td>
<td>Heliconiinae</td>
<td>Dennis-ray</td>
<td>12</td>
<td>4.26 ± 0.38</td>
<td>27.55 ± 1.87</td>
<td>6</td>
<td>5.51 ± 2.45</td>
</tr>
<tr>
<td>Heliconius doris</td>
<td>Heliconiinae</td>
<td>Blue</td>
<td>16</td>
<td>4.33 ± 0.42</td>
<td>28.00 ± 1.65</td>
<td>6</td>
<td>7.55 ± 1.27</td>
</tr>
<tr>
<td>Heliconius sara</td>
<td>Heliconiinae</td>
<td>Blue</td>
<td>12</td>
<td>5.58 ± 0.48</td>
<td>36.08 ± 2.45</td>
<td>6</td>
<td>3.09 ± 0.69</td>
</tr>
<tr>
<td>Eueides isabella</td>
<td>Heliconiinae</td>
<td>Tricolor</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>80.14 ± 48.66</td>
</tr>
<tr>
<td>Mechanitis polymnia</td>
<td>Danainae</td>
<td>Spotted</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>46.78 ± 23.27</td>
</tr>
<tr>
<td>Melinea mothone</td>
<td>Danainae</td>
<td>PA</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>10.53 ± 2.93</td>
</tr>
<tr>
<td>Chetone histrio</td>
<td>Pericopinae</td>
<td>Tricolor</td>
<td>PA assumed</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>6.92 ± 1.56</td>
</tr>
<tr>
<td>Chetone hydra</td>
<td>Pericopinae</td>
<td>Bicolor</td>
<td>PA assumed</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>7.8 ± 2.55</td>
</tr>
<tr>
<td>Anartia amathea</td>
<td>Nymphalinae</td>
<td>non-mimetic</td>
<td>none assumed</td>
<td>na</td>
<td>na</td>
<td>6</td>
<td>no learning</td>
</tr>
</tbody>
</table>

CG = Cyanogenic Glycosides; PA = pyrrolizidine alkaloids.
E = Estimated crumbs eaten until learning
T = Total crumbs eaten during the 12 trials
Behavioural experiments

To understand the benefits associated with chemical defenses in natural prey, we sought to determine how variation in defense levels alone (i.e., independently from warning signal) translate into variations in survival through predator aversion learning. Previous experiments have shown that birds are important predators of these Lepidoptera (Mallet & Barton, 1989; Kapa, 2001; Langham, 2004; Finkbeiner et al., 2012; Chouteau et al., 2016). Therefore, we used the domestic chick (Gallus gallus domesticus) as a standardized predator model to compare aversion to defended prey (Speed, 1999; Skelhorn & Rowe, 2006a, b; Stuckert et al., 2014). Domestic chicks are known to modulate their food intake depending on concentrations of chemical defenses (Skelhorn & Rowe, 2006b). We thus estimated the rate at which chicks learn to avoid crumbs made from the bodies of toxic butterflies and moths. Experiments were carried out in 12 successive trials, during which, chicks were presented simultaneously with two types of familiar colored food crumbs: one containing extracts from a chemically defended prey and dyed with one color (hereafter referred to as experimental crumbs) and the other without any prey extract and dyed with another color (referred to as control crumbs). After a few trials, chicks generally learned to associate experimental crumb color with chemical defenses and tend to avoid it.

Crums were prepared with 2.5 g of starter chick crumbs (Purina Avemicyn-A®) and 1 mL of colored water (1 g of powdered colorants from Industria Lucerico S.A.C. for 100 mL of water). To prepare experimental crumbs containing butterfly extract of a given species, we ground the bodies of 10 wild caught live butterflies from the same species together, and added 1 g of this mix to the crumb mix. To prepare control crumbs, the butterfly extract was replaced by 1 g of starter chick crumb added to the mix. The mix was then air-dried for 12 h and cut into 1 x 1 x 1 mm crumbs weighting 23.09 ± 5.43 mg (see estimated toxin concentration of crumbs in table 1). Each crumb eaten corresponds to ~28% of toxins contained in an equal “bite size” of the natural prey (1 g of butterfly for 3.5 g of crumb mix). Both experimental and control crumbs were prepared in both green and orange colorings.

Ten-day-old chicks born and raised in commercial aviary farms, and fed from birth with Avemicyn chick crumbs, were kept in a 3 x 2 m outdoor enclosure in Tarapoto (Northeastern Peru) with ad libitum access to water. The three days before the experiment, chicks were fed three times daily with a
simultaneous presentation of orange and green control crumbs in a white tray in an experimental arena. This training enables the chicks to acclimate to the experimental arena and overcome any crumbs color preference initially present. The experimental arena was composed of two adjacent 45 x 35 cm subsections with a blue floor and water dispensers, enclosed with chicken mesh. The experimental chick was placed in one subsection and a companion chick in the adjacent subsection in order to limit stress. Tested chicks were deprived from food 2 h before the experiment.

For each trial, a 35 x 20 cm white tray containing 20 orange and 20 green crumbs of distinct colors positioned alternatively along an 8 x 5 grid was placed in the experimental section. The first trial (named Trial 0 in fig. 1) was performed by using only control crumbs (i.e. 20 orange and 20 green palatable crumbs) and enabled to verify that no preference for a particular color of crumbs was present for the tested chick. The tray was removed from the experimental section as soon as 20 crumbs were attacked by the chick (sum of the orange and green crumbs pecked or eaten). During each trial, a portion of food was also given to the companion chick. Experimental trials occurred the same way as trial 0 but with all 20 green control crumbs being replaced by green experimental crumbs. Each experimental chick performed 12 successive experimental trials, with each trial separated by 30 min intervals. Six different chicks were used for each prey species, so that the design was fully nested. Colors used for experimental and control crumbs (green and orange) were switched for half of the chicks (3 per butterfly species) to allow us to control for any effect of innate preference for one of the colors. The ethical committee for animal welfare of the National Museum of Natural History (ethical declaration n° 68-064) validated this experimental design.

**Ethical Note**

The experimental design was evaluated by the ethical committee Cuvier from the National Museum of Natural History (Paris, France), which provided a ‘Favourable’ opinion, describing the impact on the animals as ‘Light’ (see licence number 68-064 attached). Our experiment is thus conforming to EU legal requirements.

*Number of individuals involved in the study:* Following the general ‘three R rules’ (Replacement, Reduction and Refinement), we used the minimum number of chicks per butterfly species tested to
obtain a sufficient statistical power \((n = 6\) per butterfly species). Only 84 chicks were thus used in this study, with a balanced sex-ratio. Chicks were circa. 10 days of age.

*Origins of the chicks used in the experiments:* Domesticated chicken are bred in many homes in this area of Peru and were thus bought in the local market, located less than 10 minutes away from the location of the experiments.

*Disposal of the chicks used after the experiments:* After a couple of days after the experiments, the chicks were donated to nearby families already raising chickens.

*Disturbance:* These domesticated chicks were placed in conditions highly similar as those encountered in the breeding facilities they came from, so that we think these conditions did not induce any particular disturbance.

*Potentially harmful manipulations:* In this behavioural experiment, we did not perform any harmful manipulation, except that half of the food provided in the experiments contained butterfly extracts that may contain some chemical compound inducing deterrence. As shown in the results of our manuscript, most chicks avoided food containing chemically-defended butterfly very rapidly, and therefore ingested a very limited amount of food containing deterring compounds. Chicks thus generally exhibit no residual signs of the experiments.

*Distress or pain:* In rare occasions \((n = 4)\), some chicks did not immediately react to the chemical compounds of the butterflies and continues to eat food containing butterfly extracts, and could display signs of stress. We thus used the following rule: if a chick display sign of prostration (do not move, has the neck located within the shoulders and closed eyes) for more than 3 minutes, we immediately stopped the experiment and moved the tested chicks back to the large 3 x 2 m outdoor enclosure. These chicks then recovered within a couple of minutes after being stopped eating food with butterfly extracts and do not display any residual signs afterwards. Each chick was used only once for the experiment and chicks that displayed prostration behavior were not tested again.
Figure 1: Model predator association learning. Attacks by individual chicks on experimental butterfly crumbs through the experiment trials (beginning at trial 1, white area). Each curve represents the change in consumption of an individual chick presented with green (grey curves) or orange (black curves) experimental crumbs. Trial 0 (grey area) was performed using only control crumbs of both colours to confirm chicks did not have a preference for one colour.
Statistical analysis

Model fitting was performed in R version 3.1.3 (R Development Core Team, 2013). Our domestic chicks generally decreased the number of unpalatable crumbs sampled in successive trials, until reaching a stable minimum attack level, mirroring a classical avoidance learning curve (Speed, 1999; Skelhorn & Rowe, 2006a). A common way for analyzing avoidance learning is through fitting a generalized linear mixed model (GLMM) with the individual predator as a random effect, assuming a binomial error variance since each crumb presented was either attacked or not (e.g. see Kazemi et al., 2014; Sherratt et al., 2015). However, one of the properties of the standard logit model GLMMs is that they necessarily assume a long-term asymptote attack probability of 0 or 1, yet we wished to determine whether the long-term asymptotic consumption rate of chicks feeding on treated crumbs differs from 0. To obtain biologically meaningful avoidance learning parameters we therefore fitted a family of exponential consumption models by non-linear least squares. Model fitting was performed using the nlme function of the nlme package (Pinheiro et al., 2017) which allowed us to also estimate nested random effects. To fit meaningful models, the analysis was restricted to the consumption of crumbs made from those species butterfly that were increasingly avoided by chicks (i.e. all species but the A. amathea control). Here the number of experimental (unpalatable) crumbs eaten was treated as the dependent variable, trial number was a covariate predictor, crumb color was a fixed effect, and both chicken ID and butterfly species were treated as random factors, with chicken nested within butterfly species. In addition to the crumb colour term (which proved to be non-significant), the most general learning model took the form of experimental crumbs eaten per trial = $a + b e^{-c \cdot \text{trial}}$ where $a$, $b$ and $c$ are the estimable parameters such that $(a+b)$ is the y-intercept (when trial = 0), $a$ is the long-term feeding rate asymptote (when trial→∞, assuming $c$ is positive) and $c$ is the learning parameter. The self-starting function SSasymp was used to facilitate model convergence. For simplicity, the random effects component was assumed to affect either parameter $a$, $b$ or $c$. Normality and homogeneity of variance of residuals around the fitted model were evaluated through examination of the residual plots. The importance of terms in the model was assessed through log likelihood ratio (LR) tests and Akaike Information Criteria (AIC).
Figure 2: Variation in predator avoidance learning generated by species within and across mimicry rings. Species showing significant differences in the rate of avoidance learning (mean E and 95% confidence interval) from all of their co-mimics (or one mimic; horizontal bar) are indicated with asterisks: ** α = 0.05, * α = 0.10. The color squares delimit the species part of a mimicry rings, and the butterfly picture the mimetic warning signal. Chemical identity of defenses is indicated for each species: CG for cyanogenic glycosides and PA for pyrrolizidine alkaloids.

After fitting the above over-arching model to the consumption of crumbs derived from species believed to be unpalatable, we found that we could not reject the null hypothesis that the long-term feeding rate (a) of chicks was zero - see Results. Instead, our most parsimonious model was one in which the rate of consumption of an unpalatable crumbs simply declines exponentially over trials (following $b e^{-c_{trial}}$, with $b$ representing the initial rate of consumption at trial 0 and $c$ representing the rate of avoidance learning). Since the random component (chick nested within butterfly species) of the learning parameter ($c$) was also highly significant in the above model, we sought to obtain parameter estimates for individual birds feeding on the crumbs derived from given butterfly species. To do this, we separately fitted the same family of exponential models to the consumption data of each and every chick using the nlsList function of the nlme package. As with our overarching fitted models above, the long-term...
consumption asymptote \( a \) was significantly different from 0 in only 3 of the 78 chicks tested (well within the expected type I error rate if the null hypothesis were true) and these cases were distributed across different butterfly species. We therefore also fitted the most parsimonious model namely \( b e^{-c_{\text{trial}}} \) to the individual crumb consumption data. Since the estimate of \( c \) for each chick was consistently (and often significantly) positive, the estimated cumulative long term consumption of unpalatable crumbs of individual chicks after removing the number of palatable crumbs consumed at trial 0 \( (E = \sum_{0}^{\infty} b e^{-c_{\text{trial}}}) - b \) will be \( b \{e^c/(e^c-1) - 1\} \). We view this composite parameter \( E \) as an appropriate and biologically meaningful indicator of the rate of avoidance learning since (i) it is trial independent and can be applied to identify the avoidance rate of those treated crumbs that have not been completely avoided by the end of the experiment and (ii) it relates directly to Müller’s argument that birds should consume a fixed number of a given unpalatable prey type before avoidance learning is complete.

Naturally, in cases where avoidance learning was largely complete by the end of 12 trials, then one would expect that the estimate of \( E \) would closely match the total number of unpalatable crumbs consumed by a given bird \( (T) \) over the course of the experiment, which is precisely what we find (see Table 1). To test for variation in the unpalability parameter \( E \) between all butterfly species and also among species that fall part of a mimicry ring, we used one-way ANOVAs on the individual estimates followed by a Tukey honestly significant difference (HSD) post hoc test to identify pairwise differences in \( E \) among species. In the study area, some of the species show distinct morphs participating in distinct mimicry rings (Table 1). Colour pattern variation is controlled by simple genetic switches in those species, but morphs are otherwise not differentiated across the genome (Joron et al., 2006; Baxter et al., 2010; Counterman et al., 2010) and share common general life histories (Brown, 1981; Joron, 2005).

Therefore, we assumed that distinct morphs within species did not vary significantly in defense.

Since trial 0 involved entirely palatable crumbs so as to start with an intercept of approximately 10, we have fitted the same set of models to our data with trial 0 starting at the first presentation of unpalatable crumbs (this implies \( E = b e^c/(e^c-1) \)). This analysis is presented in our Supplementary Materials. Since the estimates of \( E \) with and without the first trial considered are highly correlated \( (R^2 = 0.98, P < 0.001) \) and the same general conclusions apply, we focus on the former analysis here.
**Relationship between prey defense level and avoidance learning by predators**

Protection from predation, estimated using the learning parameter $E$, was compared to the average concentration of defensive chemical compounds (hereafter referred to as defense level) measured in samples of the same species. This analysis was restricted to the eight studied species of the genus *Heliconius* since 1) they are defended by the same chemical compounds: cyanogenic glycosides (CG) (Merrill et al., 2015) and 2) the mean CG concentration (measured in μg of NaCN per mg of dry butterfly) is already available from different wild-caught individuals from the exact same localities we used for this study (see Arias et al., 2016b for more details). In brief, to measure the CG concentration of each individual, (Arias et al., 2016b) have extracted all CGs (regardless of their chemical diversity) by performing a hydrolysis under harsh acidic conditions. This step frees cyanides from biological tissue and traps them in solution as NaCN. NaCN concentration ($C_{NaCN}$), was then quantified using a colorimetric method (Arias et al., 2016b). The average concentration of CG for each of the eight *Heliconius* species was obtained using at least 5 individuals per species (see Table 1).

To assess the relationship between defense levels and aversion, we compared the fit of three candidate models linking measures of concentrations of defensive compounds ($C_{NaCN}$) to avoidance learning ($E$): an intercept-only model (no relationship), a linear model (proportional relationship) and an exponential model (geometric relationship). Models were fitted using the least squares method of the nls function in the R nlme package (Pinheiro et al., 2017) and their relative parsimony was assessed through log likelihood ratio tests and Akaike Information Criteria (AIC).

To assess whether differences in defense level cause differences in aversion by predators, we assessed whether there were significant differences between species of butterflies in defense levels and behavioural responses. One way ANOVAs using Welsh’s test, followed by a Games-Howell *post hoc* tests was used when $C_{NaCN}$ was ther response variable to account for the non-homogeneity of variances (Levene’s $W_{7,142} = 3.84; \ p \leq 0.001$) and unequal sample size between groups. One-way ANOVA followed by Tukey HSD post-hoc tests was performed on $E$. 
Figure 3: Relationship between defense level and avoidance learning in eight *Heliconius* species. Avoidance learning index in tested chicks (*E*) foraging on crumbs made of butterflies whose cyanogenic glycoside concentration (*C_NACN*) has been estimated. Means and 95% confidence intervals are shown for each species of butterfly tested. The dashed curve represents the best model describing the relationship between defence and avoidance learning.

Results

*Estimation of avoidance learning using chicks*

All butterfly species but *A. amathea* (the palatable control) induced in the chicks a clear decline in the consumption of experimental crumbs over time which appear to assympote towards zero (see Figure 1). A preliminary comparison of our fitted exponential models suggests that the most parsimonious models involve the fixed effects of trial (model 1 vs model 7, LR = 2074.42, *P* < 0.001) but not crumb colour (model 1 vs model 2, LR = 1.56, *P* = 0.212) while a two parameter learning model is superior to the three parameter involving a non-zero asymptote (model 1 vs model 3, LR = 18.87, *P* < 0.001) – see Table 2. In addition, there was evidence of significant variation in the rate at which individual chicks avoided treated crumbs made from different butterfly species (model 1 vs model 6, LR = 955.3, *P* < 0.001).
Table 2: Statistical analysis of avoidance learning experiments. A range of candidate non-linear models fitted to the consumption of crumbs over time by chicks. These models involved both fixed and random effects. The most parsimonious model was one in which consumption simply declined exponentially, but there was also significant variation in the rate of avoidance learning of experimental crumbs made from different butterfly species by individual chicks.

<table>
<thead>
<tr>
<th>Model number</th>
<th>Model for treated prey consumption over trials (a, b, c, f parameters)</th>
<th>Random effect parameter (~ species/bird)</th>
<th>Df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b<em>exp(-c</em>trial)</td>
<td>c</td>
<td>5</td>
<td>3166.7</td>
</tr>
<tr>
<td>2</td>
<td>b<em>exp(-c</em>trial) + f*colour</td>
<td>c</td>
<td>6</td>
<td>3170.3</td>
</tr>
<tr>
<td>3</td>
<td>a + b<em>exp(-c</em>trial)</td>
<td>c</td>
<td>6</td>
<td>3187.6</td>
</tr>
<tr>
<td>4</td>
<td>a + b<em>exp(-c</em>trial)</td>
<td>a</td>
<td>6</td>
<td>3647.5</td>
</tr>
<tr>
<td>5</td>
<td>a + b<em>exp(-c</em>trial)</td>
<td>b</td>
<td>6</td>
<td>4080.1</td>
</tr>
<tr>
<td>6</td>
<td>b<em>exp(-c</em>trial)</td>
<td>-</td>
<td>3</td>
<td>4118.0</td>
</tr>
<tr>
<td>7</td>
<td>a</td>
<td>a</td>
<td>4</td>
<td>5239.1</td>
</tr>
</tbody>
</table>

Having established evidence that avoidance learning occurs and that there is significant variation in the rate at which crumbs made from different butterfly species are avoided by chicks, we now focus on comparing the rates at which different butterfly species are avoided. Avoidance learning (E) - estimated from fitting the most parsimonious model \( b.e^{-c.trial} \) to the consumption data of each and every chick - was variable among defended butterfly species \( (F_{12,65} = 2.11; \, p = 0.027) \). The species eliciting the fastest learning, sometimes within the first trial, were *Heliconius sara* \( (E = 3.1 \pm 0.7 \, SE) \) and *H. aoede* \( (E = 5.5 \pm 2.4) \). Learning rate was much slower for *Mechanitis polymnia* \( (E = 46.8 \pm 23.2) \) and *Eueides isabella* \( (E = 80.1 \pm 48.6) \). The latter species was the only one for which unpalatability \( (E) \) was found significantly different from some others \( (*H. sara* and *H. aoede* at \( p = 0.05) \) when all butterflies are analysed jointly (i.e. mimicry rings were not considered).

**Variation in avoidance among species within mimicry rings**

Estimating avoidance learning with a chick predator enabled us to compare the protective effect induced by the toxin content of the different prey species. Here, we compared the strength of chemical defense in Heliconiinae, Danainae and Pericopinae relying on different toxins and belonging to the same mimicry rings.

When comparing among co-mimics, protection was overall found to be relatively similar (Figure 2), even when species belong to distantly-related clades. Only the slower levels of avoidance learning...
found for *H. erato* resulted in two mimicry rings showing significant variation in avoidance learning among species: ‘postman’ ring ($F_{1,10} = 8.782; \ p = 0.014$), and ‘dennis-ray’ ring ($F_{4,25} = 6.456; \ p < 0.001$).

**Relationship between the level of prey defenses and avoidance learning by predators**

Focusing solely on the Heliconiinae butterflies which all rely on cyanogenic glycosides for defense, we investigated the relationship between mean toxicity levels for each species (i.e. $C_{NACN}$) and the resulting protection from predator as estimated with $E$. By comparing different statistical models, we found that an exponential decay model (asymptotic decrease) provided the best fit to the data (AICs: intercept only model 51.50; linear model 44.84; exponential model 34.67) and that the exponential model explained more variation than the linear model with the same number of terms ($F_{1,5} = 37.191, \ p = 0.002$, see Figure 3).

As previously described in Arias et al.(2016b), variations in $C_{NACN}$ can be clustered into three significantly different groups ($F_{Welch,7,36.75} = 30.091; \ p < 0.001$; Figure 2): 1) *H. erato* was the least toxic, with a significantly lower $C_{NACN}$ than all other species ($C_{NACN} = 0.75 \pm 0.13 \mu g / mg$; intercluster post hoc $p \leq 0.001$); 2) *H. numata, H. melpomene, H. ethila* and *H. burneyi* formed a group with intermediate $C_{NACN}$ ($C_{NACN} = 2.17 \pm 0.12$ to $2.69 \pm 0.51 \mu g / mg$; intracluster post hoc $p \geq 0.900$; intercluster post hoc $p \leq 0.043$); 3) *H. aoede, H. doris* and *H. sara* were the most toxic, with highest $C_{NACN}$ ($C_{NACN} = 4.26 \pm 0.38$ to $5.58 \pm 0.48 \mu g / mg$; intracluster $p \geq 0.072$; intercluster $p \leq 0.043$).

By contrast, prey protection estimated here by chick learning behavior revealed only two main clusters of species ($F_{7,40} = 5.173; \ p < 0.001$; Figure 3) with a single species, *H. erato*, on the one hand, generating avoidance learning at a significantly slower rate than all other species (intergroup post hoc $p < 0.019$). All other species where $C_{NACN}$ was estimated were avoided at a similar speed (intragroup post hoc $p \geq 0.122$).

We showed that $E$ initially decreased with increasing toxicity $C_{NACN}$ (Figure 3). However, above a threshold of toxicity (here detectable with *H. numata* containing $C_{NACN} = 2.17 \pm 0.12 \mu g / mg$), $E$ no
longer decreased. This asymptotic non-linear relationship between $C_{\text{NACN}}$ and $E$ suggests that above a threshold of around $C_{\text{NACN}} = 2 \mu g / mg$, a two or three-fold increase in concentration of chemical defense level provides similar rate of avoidance learning.

**Discussion**

*Are diverse communities of mimetic prey influenced by their variation in toxicity?*

By studying the behavior of a model avian predator faced with the toxins derived from wild butterflies, we provide new insights into the relationships linking prey defense and predator learning in natural communities. First, our experimental data provide very little evidence for the existence of a non-zero asymptote for crumbs prepared from our butterflies. This is important because the theory of quasi-Batesian mimicry specifically depends on the assumption of a non-zero asymptote (Speed, 1993; Mallet and Joron, 1999), otherwise even moderately defended prey would ultimately not be parasitic on the better defended prey they resemble. Our results also suggest that despite large variation in chemical defences between prey types, most are avoided at similar rates. Indeed, the non-linear asymptotic relationship between toxin concentration and behavioural response results in a situation where several-fold variation in chemical concentrations in prey items does not translate into proportional variation in protection from our model predators. Only in the range of low to moderate defense level does an increase in defense level translate into faster avoidance learning for our model predator. This confirms the capacity of our chick predators to perceive and discriminate between some levels of toxicity. Yet for higher defense levels, avoidance learning quickly saturates and often occurs after a single trial.

Many dose-response relationships in toxicology are non-linear (often logit or probit; Demidenko et al., 2017), and so that a saturating behavioural response can be expected. However, the position of defenses contained in butterflies from natural communities with respect to the saturation threshold was unknown. The majority of the tested species fall above the saturation threshold, where variation in toxin titres does not generate much variation in predator response. Assuming that the behavior of domestic chicks is representative of natural predators, parasitic (quasi-Batesian) mimicry relationships (Rowland et al., 2010) may be expected to apply to at most a small proportion of the species composing this natural
community (those eliciting slow learning), while most species would be sufficiently toxic to elicit
efficient learning by predators.

This result leads us to suggest that species in the natural communities may have a role in their respective
mimicry ring that is not determined primarily by their relative toxicity, but by other ecological or
phenotypic differences. Indeed, because we purposefully prevented access to visual signals in our
experiments, our results reveal the response of predators to the taste (or toxicity perception) component
of associative learning, which amounts to assuming indistinguishable warning signals between prey
species. Under this assumption (which may hold for instance when warning patterns are widely
generalized among co-mimics or when mimics are extremely similar), then co-mimics with different
levels of toxicity make a similar per capita contribution to the predator’s avoidance learning. Assuming
our model predator behave consistently with natural predators, this suggests that relationships at the
community level are likely to be largely mutualistic, and that differences in mimicry roles (model vs.
mimic) would be determined primarily by species abundance in contrast to Batesian mimicry
expectations (Finkbeiner et al., 2018). Within these mimicry rings, selection on colour pattern is thus
likely to promote resemblance to the warning signal of the most abundant species, rather than of the
most toxic one (Mallet, 1999).

Nevertheless, co-mimetic species may differ in other traits influencing predator avoidance learning.
Phenotypic differences such as variation in signal quality (wing pattern, flight movements, and
imperfect resemblance), variation in the modalities of toxin presentation, in additional signaling
modalities, or in habitat preferences may also play a significant role in the evolution of mimetic color
patterns within natural communities (Beccaloni, 1997; Skelhorn & Rowe, 2005; Lindstrom et al., 2006;
Skelhorn & Ruxton, 2008). Although crumb color was unimportant in our experiments, we cannot rule
out interactive effects of toxin and phenotype on aversion learning. For example, classic work by
Gittleman & Harvey (1980) indicated that domestic chicks learned more quickly to avoid unpalatable
crumbs when they were conspicuous, so differences in toxicity might generate differences in aversion
learning only when associated with particular phenotypes. Finally, we acknowledge that spatial and
temporal variation in predation pressure (Chouteau & Angers, 2011; Mappes et al., 2014), in mimetic
community composition or in unpalatability (e.g. depending on hostplant secondary defenses and
availability) might all affect aversion learning by predators and could therefore impact the nature of mimetic relationships.

**What drives the evolution of variations in toxicity?**

Understanding how natural selection exerted by predators drives the evolution of defense levels in aposematic prey is not straightforward. In most cases, chemical defense cannot be detected prior to attack. Therefore, investment in individual defense level is not predicted to increase in response to natural selection on attack avoidance. However, aposematic insects can sometimes survive predation attempts (Wiklund & Järvi, 1982; Sillén-Tullberg, 1985; Chai, 1996) and predators can often taste and release prey selectively based on the perception of harmful chemicals (Pinheiro, 1996; Gamberale-Stille & Guilford, 2004; Skelhorn & Rowe, 2006a, c). For instance, beak marks are often seen on the wings of aposematic butterflies indicating previous attacks by predators (Smith, 1979; Mallet & Barton, 1989; Ohsaki, 1995) and many defended prey possess traits which increase the chance of surviving an attack, either by increasing tissue resistance or by increasing the chance of being taste-rejected without harm (Skelhorn & Ruxton, 2008). For instance, storing toxins in the wing integuments (Nishida, 1994, 2002), or excreting toxins when handled (such as the thick and bright yellow haemolymph secreted by labybirds (Holloway et al., 1991) or the voluminous foam produced by Chetone moths) may discourage handling by predators and improve survivorship. Therefore, surviving attacks can be enhanced by individual toxin contents, and selection may favor traits in the prey that increase the rapid perception of their defenses by predators. Here, the fact that most species display CG concentrations above the sufficient amount for associative learning (2 µg/mg in our experiment) might reflect the importance of being sufficiently defended to derive individual benefits from chemical defences.

The evolution of higher toxin contents above a certain threshold also points at other evolutionary forces shaping the level of defense. Increased toxicity may provide increased protection against toxin-resistant predators (Williams & Brodie, 2003; Trigo, 2011) or may be related to other process not related to predator deterrence (Speed et al., 2012). High toxin loads may be expected (1) if they provide protection against parasitoids (Campbell & Duffey, 1979; Barbosa et al., 1991; Sime, 2002; Lampert et al., 2008 but see 2010), (2) if they are a by-product of host-plant adaptation (Engler-Chaouat & Gilbert, 2007), or (3) if they are associated with a higher reproductive success, for instance when toxins are used as
precursors for male pheromone synthesis (Hartmann et al., 2005; Trigo, 2011) or contribute to nuptial
gifts (Cardoso & Gilbert, 2007). Disentangling these effects will improve our understanding of the
multiple forces shaping defense levels in prey, and hence their role in antipredator strategies such as
mimicry.

**Estimating protection induced by chemical defenses using domestic chicks**

While domestic chicks are not the natural predators of flying insects, both their capacity to modulate
avoidance learning, and their ability to rapidly learn to reject toxic crumbs support their use as a model
predator in many studies (Speed, 1999; Skelhorn & Rowe, 2005, 2006a, b, c; Amézquita et al., 2013).
We also observed relatively low individual variation among chicks in their avoidance learning (Figure
1), although variation was higher with the less protected species, as previously observed (Brower,
1958). Moreover, our findings using domestic chicks as a standardized predator are generally consistent
with previous studies evaluating the behavior of other predators toward mimetic butterflies. Like this
study, Arias et al. (2016a) found that *M. polymnia* was more palatable to birds (wild great tits),
compared to *H. numata* and *H. melpomene*, but they found *H. erato* to be avoided more readily than we
did, although still less than other species. Using caged silver-beaked tanagers (*Ramphocelus carbo*) and
laboratory-reared *Heliconius* butterflies, Brower et al. (1963) showed that *H. numata*, *H. melpomene*,
and *H. erato* were significantly better protected than *H. doris*. The discrepancy with our study lies in a
lower estimated protection for *H. doris* and equal protection in *H. erato* and *H. melpomene*. This may
stem from true differences in avoidance learning between studies, perhaps because of geographic
variation in toxicity (Brazil vs. Peru) or from differences between laboratory-reared vs. wild-caught
specimens. Alternatively, estimates might differ because variation in certain visual components of the
warning signal such as pattern, size, or behavior displayed by the live butterflies, as used by Brower et
al. (1963), could affect predator avoidance response. We believe that the number of "prey items"
required for domestic chicks to completely avoid a prey species should allow a more ecologically
relevant estimation of prey defenses (Rowland et al., 2017) than the use of disgust behavior as
previously used (Brower et al., 1963; Arias et al., 2016a). Although our methodology cannot enable us
to dissect the proximal causes of behavioural differences among chicks (sensitivity of gustatory
receptors, visual perception, toxin tolerance, cognitive capacities, etc.; see Ruxton et al., 2018 for a
review) it allows us to estimate the impact of variation in chemical defenses on prey survival, and
explore its implications for the evolution of defenses in mimetic communities. Finally, the method
presented relies on widely available domestic chicks and little equipment, and therefore offers the
advantage of being easily applicable to a variety of organisms and enables direct comparisons between
different studies when the same strain of chick is used (see Rowland et al., 2017).

Conclusions

Here, using a model predator, we provide evidence that the protection associated with the perception of
natural chemical defenses by predators does not increase proportionately with the concentration of
toxin. Avoidance learning plateaus above a certain threshold of toxin concentration. Above this level,
the range of toxin concentrations is not readily explained by increased protection benefits, but perhaps
by other ecological processes or even, in the absence of cost associated with toxin accumulation, simply
by neutral processes. Our results suggest that if prey species within a mimicry ring cannot be readily
distinguished by predators on the basis of their visual signals, then most well-defended species are
expected to make a similar per capita contribution to the avoidance learning of their warning signal by
predators. By contrast, only a small proportion of species with markedly lower toxin content elicited
the slower avoidance learning that might indicate a lower per capita contribution to warning signal
efficiency.

Data archival: Data are available in the linked Mendeley data repository.

Acknowledgments: We thank M. McClure, N. Rosser and M. Tuatama for their help with the logistics
of the project; M. Veillette for her comments on manuscript drafts. We also thank the Peruvian
government for providing the necessary research authorizations (permit: 002-2015-SERFOR-
DGGSPFFS) and the Gobiero Regional San Martin PEHCBM for the authorization to perform
research within the Cordillera Escalera Conservation Area (permit: 124-2016-GRSM/PEHCBM-
DMA/EII-ANP/JARR). This research was supported by a Marie Sklodowska-Curie fellowship
(FITINV, 655857) to MC, a French National Agency for Research (ANR) grant (DOMEVOL, ANR-
13-JSV7-0003-01) and an ‘Emergence’ grant from Paris city council to VL, and by a European
References


**Supplementary Information**

**Table S1:** Statistical analysis of avoidance learning experiments based entirely on unpalatable prey. Here we have discarded data on the total number of palatable crumbs consumed of the experimental colour in the initial presentation, so that trial 0 (formerly trial 1) begins with the first presentation of unpalatable crumbs. The same range of candidate non-linear models were fitted to the consumption of treated crumbs over time by chicks as in the main text (we have kept the same order for comparability; nc = no convergence). Once again most parsimonious models involve the fixed effects of trial (model 1 vs model 7, LR = 731.28, P < 0.001) but not crumb colour (model 1 vs model 2, LR = 1.93, P = 0.165). In addition, there was evidence of significant variation in the rate at which individual chicks avoided treated crumbs made from different butterfly species (model 1 vs model 6, LR = 545.42, P < 0.001).

<table>
<thead>
<tr>
<th>Model number</th>
<th>Model for treated prey consumption over trials ((a, b, c, f) parameters)</th>
<th>Random effect parameter ((~\text{species/bird}))</th>
<th>df</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(b\times \exp(-c\times\text{trial}))</td>
<td>(c)</td>
<td>5</td>
<td>3317.4</td>
</tr>
<tr>
<td>2</td>
<td>(b\times \exp(-c\times\text{trial}) + f\times\text{colour})</td>
<td>(c)</td>
<td>6</td>
<td>3321.3</td>
</tr>
<tr>
<td>3</td>
<td>(a + b\times \exp(-c\times\text{trial}))</td>
<td>(c)</td>
<td>6</td>
<td>nc</td>
</tr>
<tr>
<td>4</td>
<td>(a + b\times \exp(-c\times\text{trial}))</td>
<td>(a)</td>
<td>6</td>
<td>3526.2</td>
</tr>
<tr>
<td>5</td>
<td>(a + b\times \exp(-c\times\text{trial}))</td>
<td>(b)</td>
<td>6</td>
<td>3444.6</td>
</tr>
<tr>
<td>6</td>
<td>(b\times \exp(-c\times\text{trial}))</td>
<td>-</td>
<td>3</td>
<td>3858.8</td>
</tr>
<tr>
<td>7</td>
<td>(a)</td>
<td>(a)</td>
<td>4</td>
<td>4046.7</td>
</tr>
</tbody>
</table>

**Figure S1a,b**

Plot of the estimate of the total number of treated crumbs consumed by chicks when the first presentation of palatable crumbs was included or excluded from the analysis (in 62 cases where convergence was obtained). (a) entire data set, (b) the data set without five high values. Here the fitted curves were similar so that the two estimates were highly correlated \((F_{1,60} = 3107, R^2 = 0.98, P < 0.001)\).
Estimates of $E$ (when the first presentation was removed) were not different among defended butterfly species ($F_{12.49} = 1.203; p = 0.308$) when analysed conjointly. When analysed by mimicry rings, results are similar to the one presented in the main text with only the mimicry rings including $H. erato$ displayed significant variation of $E$: ‘postman’ ring ($F_{1.9} = 5.873; p = 0.038$), and ‘dennis-ray’ ring ($F_{4.18} = 3.583; p = 0.025$).

Finally, we found the same relationship between mean toxicity levels for each species (i.e. $C_{NACN}$) and the resulting protection from predator as estimated with $E$. The exponential decay model (asymptotic
decrease) provided the best fit to the data (AICs: intercept only model 51.09; linear model 42.75; exponential model 30.45) and explained more variation than the linear model with the same number of terms ($F_{1,5} = 65.581, p < 0.001$).