

Weak warning signals can persist in the absence of gene flow

Lawrence, J. P. ‡*, **Rojas, B. ‡***, Fouquet, A., Mappes, J., Blanchette, A., Saporito, R., Bosque, R. J., Courtois, E., & Noonan, B. P. 2019. (‡Equal contribution).

<https://www.pnas.org/lookup/doi/10.1073/pnas.1901872116>

Assay 2: Clay model experiments

We constructed 45 mm long (snout to vent) replica frogs using Van Aken polymer modelling clay which does not harden and retains evidence of predation attempts. We poured melted clay into silicone molds made from plastic model replicas. In order to assess how predators responded to novel colors and patterns, we created models as follows. In the white population, we deployed models with the local color (white), that could have either the local pattern (stripes) or a new pattern (solid white), and also individuals with the local pattern (stripes), but a different color (yellow). Conversely, in the yellow population we had a novel color (white) with the local pattern (stripes), but also the local color (yellow) both with the local pattern (stripes) and a novel pattern (solid yellow). All models had blue legs (as do both populations sampled), and the stripes were placed on a black body. Both dendrobatid frogs and the polymer clay have been shown to lack UV reflectance. Both stripes and eyes were created with clay and affixed to models.

Models were placed along transects on which three model types were randomized with one model every 5 m. The three model types for each location were local color and pattern, local color and novel pattern, local pattern and novel color, ensuring that at least one component of the aposematic signal was familiar to predators. Transects were separated by at least 100m. We placed seven transects from 495m long to 1.5km long in the white population and 13 transects from 495m long to 1.5km long in yellow population. Models (1378 and 1136 in the yellow and white populations, respectively) were left for 72h to allow for predation attempts. Following this period, we collected models and determined which ones had been attacked by avian predators. Avian attacks were recognized by a characteristic U or V shape, as well as stab marks. We focused on attacks by avian predators because they are likely the most important in driving phenotypic

diversity in conspicuous signals given their ability to discern different colored phenotypes. The non-avian organisms that attack plasticine models, namely arthropods and mammals, are also capable of seeing at least differences in brightness, but mostly forage following chemical cues, making the biological interpretation of their attacks difficult. Similarly, snakes, which are some of the few known predators of poison frogs, can also see color, but they are highly motivated by movement and chemical cues, making it highly unlikely that they would attack clay models. If models had multiple bite marks, they were scored as a single predation attempt as we were not able to determine whether one predator attacked multiple times or multiple predators attacked once. Missing models were excluded from the analysis (43 of 1136 and 57 of 1378 for white and yellow, respectively). We used a General Linear Mixed Model with binomial distribution to compare predation attempts among phenotypes (novel/local pattern, novel/local color) at both sites to determine whether aposematic signal is a predictor of predation risk, using transect ID as a random factor. Thus, we used color novelty (y, n), pattern novelty (y, n) and population (w, y) as independent variables, and tested for their effect on whether each model was attacked (1) or not (0). Therefore, local or novel traits to the avian community are what are notable in this analysis rather than individual clay model phenotypes.

Data file name: LawrenceEtAl2019PNAS_Assay2Data

Variable description:

pop = population where the models were deployed

transect = as its names implies, transect number (important as models within the same transect are not to be considered fully independent from each other)

colpatt = colour patten of each of the models. ys = yellow stripes, ws = white stripes, sw = solid white, sy = solid yellow).

novpatt = says whether the pattern (striped/solid colour) is novel or local in the population where the model was deployed.

novcol = says whether the colour (white/yellow) is novel or local in the population where the model was deployed

pattern = says whether the models has stripes or a solid colour

colour = says whether the model is yellow or white

novelty = says whether the colour pattern as a whole (ys, ws, sw, sy) is novel or local

birdatt = whether the model exhibits marks attributable to a bird attack (1) or not (0). NA was assigned to missing models.

Each row represents one model.

Assay 3: Learning experiment

We trained naïve one-week old chickens to eat mealworms (*Tenebrio molitor*) from a petri dish with an image of a frog beneath the dish. Frogs were printed illustrations that were brown (control) or had white or yellow stripes, on a black body and blue legs, similar to the clay models. White and yellow colors were taken from photographs of frogs in the white and yellow populations, respectively, using the color dropper tool in Photoshop. Unmodified mealworms were served in association with the control, brown frog, whereas mealworms associated with white or yellow striped frogs were tainted with a distasteful chloroquine solution. Using this design, we measured predator-learning rates for white-striped and yellow-striped frogs and then explored their response to similar, but novel aposematic signals. To make mealworms distasteful we soaked them in a chloroquine diphosphate (98%, Arcos Organics™) solution, which is an alkaloid and known to be distasteful, but not harmful, to birds. Over successive trials, we examined whether chicks would learn to avoid distasteful mealworms associated with an aposematic signal, with learning defined as three consecutive refusals of a distasteful mealworm. Once a chick learned to avoid a particular aposematic signal, it was presented with the other signal (e.g., trained to avoid yellow stripes, presented with white stripes). We further examined the effect of distastefulness level on learning by varying levels of distastefulness by training chicks with mealworms that had been soaked in either a 5% or 10% chloroquine solution for 1-3 hours. Chicks were equally divided (N = 15 per treatment; 60 chicks total) into four different treatments: 5% chloroquine with a yellow signal, 5% chloroquine with a white signal, 10% chloroquine with a yellow signal, and 10% chloroquine with a white signal. Thus, we were able to explore avoidance in the context of associated color and/or

distastefulness (chloroquine concentration). The trials were done in 30 cm x 60 cm wooden compartments under full spectrum lights. Chicks were placed individually into a compartment and allowed to habituate for 2 hours. Chicks were food-deprived during this acclimation period to ensure motivation to feed during the trials. Once in the compartment and after the 2h acclimation period, training consisted of teaching the chicks to eat a dried mealworm from a petri dish on top of an illustration of a brown frog on a tan background. The training phase was completed once the chick had eaten three consecutive times, after which we allowed the birds to rest for a period of 5 min.

The avoidance-learning trials consisted of the consecutive presentation of mealworms on a petri dish on top of an illustration of *D. tinctorius* on a tan background. (i.e., a frog with either white or yellow dorsal stripes). We recorded the latency (i.e., the time until the chick approached and picked up the mealworm) and noted any behavioral reaction to the distastefulness of the mealworm after tasted or eaten. Such behaviors most often involved beak wiping and head shaking. Each trial ran for 5 min., followed by 5 min. rest, after which a new petri dish with an unpalatable mealworm was offered. This procedure was repeated until the chick refused to eat the mealworm over three consecutive trials, at which point the chick had learned the signal. The test ended either when chicks “learned” the signal or proceeded through 10 trials, whichever came first. If chicks proceeded through 10 trials without three consecutive refusals, they were considered to not have learned the signal. When a chick did not eat the chloroquine-soaked mealworm on the *D. tinctorius* pattern, a palatable mealworm was offered on the neutral (brown frog) background, which the chick did not associate with an unpleasant experience. In that way, we made sure that the chicks had refrained from eating the treated mealworm because of an association with the warning signal, rather than satiation. In five cases, chicks stopped eating and subsequently refused the palatable mealworms and as a result, were excluded from analysis. In these cases, we tested additional chicks to ensure each treatment had 15 replicates. We registered the number of trials that it took for each chick to learn to avoid an aposematic (white or yellow) signal.

The next step was to run a generalization trial, which aimed to test whether, once a signal was learned, the aversion would be extended to other signals. Therefore, after the three trials in which the chick would refrain from eating the presented mealworm, chicks were given a 5 min rest, and then a new unpalatable mealworm was presented in association with the other (white/yellow) aposematic signal. The unpalatable mealworm had the same chloroquine concentration on which the chick was originally trained. Chicks that had learned to avoid yellow-striped frog images were presented with a mealworm atop a white-striped frog image, and vice versa. We recorded the chicks' response both as a binary variable (whether the mealworm was eaten or not) and the hesitation time. Generalization trials were only conducted on chicks that learned to avoid a signal in the first set of trials. Chicks were considered to have generalized avoidance to the novel signal if they avoided the first presentation of the novel signal. This situation best simulates choices wild predators that have learned to avoid a locally common aposematic signal would make when encountering a novel signal. All trials, both avoidance learning and generalization, were filmed in order to extract details on the chick's behavior afterwards.

We analyzed the data from the learning experiments using a generalized linear mixed model (GLMM) in two ways. First, we used a survival analysis (Cox regression) to examine the effect of color and chloroquine content (5% vs. 10%), and the interaction between the two, on the latency to attack (i.e., time to event) each mealworm. Chick ID was included in the model as a random factor to account for repeated presentations to the same individual. Then, for each chick, we counted the number of trials in which the mealworm was attacked, to a maximum of ten, and whether or not the chicks learned to avoid the signal they were presented (as defined by three consecutive refusals). In the first case, we used a Poisson error distribution, whereas in the second we used a binomial distribution. In both cases, our predicting variables were color, chloroquine concentration, and the interaction between the two. Finally, we used a last GLMM to test whether the chloroquine concentration and the color of the signal learned predicted whether or not the chicks would

generalize their learned aversion to the alternative signal. All analyses were done in R, with the RStudio interface and using the packages *lme4* and *coxme*.

Data file name: LawrenceEtAl2019PNAS_Assay3Data

The data file contains two tabs, one for the information of each of the trials (Learning Trials), and a second one (LearningAver+Generalization) with the averages for each chick + information of whether each chick generalized or not (see methods). The variable names and description apply for both tabs, when appropriate.

Variable description:

chickid = individual id of each of the birds used

signal = colour (yellow, white) of the stripes of the frog illustration used as a stimulus

chlor_treatm = amount of chloroquine (low/high) used with the mealworms offered to the birds

trial = as its name implies, trial number.

latency = time taken by each bird to “attack” a mealworm during each trial

learnt = whether or not the bird showed to have learnt the signal shown in that particular trial

avg_latency = average time taken by each bird to “attack” a mealworm across trials

trials_avoid = number of trails it took for the bird to avoid attacking the mealworms presented

number_taste = total number of times the bird attacked (and thus tasted) the mealworm presented

learnt_final = whether or not the bird ended up learning to avoid attacking the mealworms

generalize = whether or not, after having learned to avoid the mealworms presented on a signal, the bird refused to attack the mealworm presented on the other signal

In ‘Learning Trials’, each row represents one trial. In ‘LearningAver+Generalization’, each row represents an individual bird.

Assay 5: Unpalatability experiments

We assessed skin contents in two ways. First, we took equal proportions of methanol extracts (Unpalatability Assay A). This represents how predators respond to natural variation among individuals. Our second assay (Unpalatability Assay B) sought to examine how skin content composition affected predator response by controlling for dry weight of the skin contents. This aids in inferring how different alkaloid profiles coupled with non-alkaloid content impact predator response. We note, however, that predator responses could be skewed if non-alkaloid content varied among individuals (see below). The two different assays were identical except for the skin secretion toxin preparation. The purpose of doing two assays was to attempt to 1) understand how toxins influence predator response and 2) ensure that non-toxin components of skin secretions do not disproportionately affect results and mask potential aversion to the alkaloids. For each assay, we tested one frog sample with one bird (white, N = 10, yellow, N = 8), with six control birds for the first assay and seven control birds for the second assay.

In Unpalatability Assay A, we diluted each methanol extract equally, evaporating 1.0 mL of methanol to dryness under N₂, and then reconstituted with 0.5mL ethanol regardless of dry mass. We added 15 µl of the reconstituted sample to each oat and allowed the oats to dry. These oats would then be presented to birds. While this is a more realistic scenario of how birds would respond to natural variation among these frogs, we sought to control for dry mass in an effort to examine how composition affected response. Thus, we conducted a second assay that controlled for dry mass.

For Unpalatability Assay B, we evaporated 1 mL of methanol extracts to dryness and weighed to determine the approximate quantity of skin content present for each sample (notably, this included everything in the methanol such as mucus, cholesterol, fatty acids, carotenoids, etc., which, while not necessarily defensive in nature, their presence may impact the efficacy of distasteful alkaloids). Samples were then reconstituted in a volume of ethanol such that the concentration of toxin for each sample was the same based on the mass of the dried sample (ca. 1:1

toxin mass to ethanol). Toxins extracts were then transferred to oats (15 μ l/oat) and allowed to dry. This was in an effort to control for the quantity of toxins present in the extracts and not allow individual variation in alkaloids quantity skew results. However, if the mass of the non-alkaloid content varied among frogs, especially in a manner independent of alkaloid mass, this may have unintentionally altered our concentrations among samples eliciting non-biologically relevant behaviors among the birds.

Prior to experiments, Blue Tits were trained to eat untreated oats. After training, birds were given oats to which toxins had been added following the protocol described in Rojas et al. 2017 (Proc. Roy. Soc B 284:20171424). Aversive behavior (i.e., beak wiping) and percentage of oat eaten were recorded. For each assay, each of two oats were soaked with 15 μ l of extract of one frog skin and left for 24 h at room temperature to ensure that all ethanol had evaporated. Two other oats were soaked with 15 μ l each of pure ethanol and used at the beginning and end of the experiment with each bird. The first ethanol-only oat needed to be consumed entirely by the bird before the experiment could be initiated, to ensure motivation to eat; the second ethanol-only oat was offered in the final trial to ensure that the birds were not refusing to eat the oats coated with toxins due to satiation or lack of motivation to eat in general. Birds in the control treatment received oats soaked with pure ethanol for all trials in order to compare directly the response of birds to oats containing frog toxins versus oats with ethanol only.

Each oat (one at a time) was presented on a hatch that had a visual barrier, which allowed us to detect the exact moment at which the oat was seen, which set the actual beginning of the trials. We measured the number of times the bird wiped its beak, which is a known aversive behavior, and the percentage of the oat eaten. Birds were watched for a 2 min. period after they finished eating the oats, or for a maximum of 5 min. in those instances in which the oat was not fully eaten, to make sure that any delayed response to the oat taste would not be missed. Data were analyzed using Generalized Linear Mixed Models using the package *lme4*. We entered frog population as the predicting variable, while number of times the beak was wiped and percentage of oat eaten were

entered as the response variables. Because the response of each bird was measured twice, we included bird ID as a random factor. Given that the duration was not the same for all trials, we used the function “offset” in R to account only for the effect of population on our response variable once the effect of duration was controlled for. These and all other statistical analyses were done in R using the RStudio interface, unless stated otherwise.

Data file name: LawrenceEtAl2019PNAS_Assay5Data

The data file contains two tabs, one for Assay 5A, and another one for Assay 5B. The variable names and description apply for both tabs.

Variable description:

birdid = individual id of the bird that was exposed to the skin extracts of a particular frog

frogid = individual id of each frog whose skin extracts were tested

frog_svl = body size (snout-vent length) in mm of every frog individual

pop = population where each frog comes from

treatm = treatment; either 0 (pure ethanol) or 1 (frog skin extracts)

trial = as its name implies, trial number. Trials 1 and 4 were always done with pure ethanol (see methods) while trials 2 and 3 were always done with frog skin extracts.

prop oat = proportion of the oat eaten by the bird during each trial.

ncb = number of times that each bird wiped its beak (sign of disgust) during each trial.

trialdur = duration of each trial. Used as a baseline to calculate the rate of beak wiping and the speed at which the oat was eaten.

Each row represents one trial.