2019

Weak warning signals can persist in the absence of gene flow

J. P. Lawrence
University of Mississippi, jplarry@gmail.com

Bibiana Rojas
University of Jyväskylä

Antoine Fouquet
Laboratoire Evolution et Diversité Biologique

Johanna Mappes
University of Jyväskylä

Annelise Blanchette
John Carroll University

See next page for additional authors

Recommended Citation

Lawrence, J.P., B. Rojas, A. Fouquet, J. Mappes, A. Blanchette, R.A. Saporito, R.J. Bosque, E.A. Courtois, B.P. Noonan. Weak warning signals can persist in the absence of gene flow. Proceedings of the National Academy of Sciences. Published Sept. 3, 2019. https://doi.org/10.1073/pnas.1901872116

Authors
J. P. Lawrence, Bibiana Rojas, Antoine Fouquet, Johanna Mappes, Annelise Blanchette, Ralph A. Saporito, Renan Janke Bosque, Elodie A. Courtois, and Brice Noonan

Weak warning signals can persist in the absence of gene flow

J. P. Lawrence[a,1,2,3], Bibiana Rojas[b,1,3], Antoine Fouquet[c], Johanna Mappes[b], Annelise Blanchette[d], Ralph A. Saporito[d], Renan Janke Bosque[a], Elodie A. Courtois[e], and Brice P. Noonan[a]

[a]Department of Biology, University of Mississippi, University, MS 38677; [b]Department of Biological and Environmental Science, University of Jyväskylä, 40014 Jyväskylä, Finland; [c]Laboratoire Evolution et Diversité Biologique, Centre National de la Recherche Scientifique, UMR5174, 31062 Toulouse cedex 9, France; [d]Department of Biology, John Carroll University, University Heights, OH 44118; and [e]Laboratoire Ecologie, Evolution, Interactions des Systèmes Amazoniens, Centre de Recherche de Montabo, Universite de Guyane, BP 70620, 97334 Cayenne cedex, France

Edited by David B. Wake, University of California, Berkeley, CA, and approved August 2, 2019 (received for review January 31, 2019)

Aposematic organisms couple conspicuous warning signals with a secondary defense to deter predators from attacking. Novel signals of aposematic prey are expected to be selected against due to positive frequency-dependent selection. How, then, can novel phenotypes persist after they arise, and why do so many aposematic species exhibit intrapopulation signal variability? Using a polytypic poison frog (*Dendrobates tinctorius*), we explored the forces of selection on variable aposematic signals using 2 phenotypically distinct (white, yellow) populations. Contrary to expectations, local phenotype was not always better protected compared to novel phenotypes in either population; in the white population, the novel phenotype evoked greater avoidance in natural predators. Despite having a lower quantity of alkaloids, the skin extracts from yellow frogs provoked higher aversive reactions by birds than white frogs in the laboratory, although both populations differed from controls. Similarly, predators learned to avoid the yellow signal faster than the white signal, and generalized their learned avoidance of yellow but not white. We propose that signals that are easily learned and broadly generalized can protect rare, novel signals, and weak warning signals (i.e., signals with poor efficacy and/or poor defense) can persist when gene flow among populations, as in this case, is limited. This provides a mechanism for the persistence of intrapopulation aposematic variation, a likely precursor to polytypism and driver of speciation.

aposematism | frequency-dependent selection | polymorphism | unpalatability | secondary defenses

Aposematism is a widespread defensive strategy whereby organisms use conspicuous signals (i.e., color, sound, odor) to warn predators of secondary defenses (1), which are often chemical in nature (i.e., alkaloids, cardiac glycosides, venom) (2). In principle, aposematic signaling should be subject to strong, positive frequency-dependent selection (+FDS). This is because, in order to function efficiently, predators must recognize and avoid the aposematic signal, and thus novel signals will be selected against due to their rarity. Indeed, theoretical work (3–5), as well as laboratory (6, 7) and field experiments (8–10) support this hypothesis. However, there are numerous examples in nature where aposematic species display intrapopulation (11–14) or interpopulation (14–16) phenotypic variation (polymorphism or polytypism, respectively). While polytypic species may adhere to expectations of strong +FDS if predators' ranges do not include multiple phenotypically distinct populations, polymorphism would appear to violate the expectations of +FDS (3, 17, 18).

While the origins of polytypism are subject to rigorous debate, one likely avenue for a species to evolve polytypism is to first begin with polymorphism (19, 20). Founding propagules from polymorphic source populations may, through founder effect (21) or differential selection (22), allow for interpopulational differentiation, possibly even speciation. Therefore, understanding how polymorphism can persist is pivotal to understanding how and why polytypic species occur (23). Wright's shifting balance theory (24) has been invoked as a possible explanation for the existence of polytypism in aposematic signals (4, 21, 25) as the phenotypic differences among populations may represent different phenotypic optima for communicating unprofitability of prey to their potential predators based on local environmental conditions. Under Wright's shifting balance, polytypy may arise due to localized patterns of selection for local optima. It is worth noting that polytypic populations may not only reflect populations in which each has reached a locally distinct phenotypic optimum, but also instances in which populations have not yet reached an adaptive peak (4, 25), or have become "stranded" on suboptimal peaks of the fitness landscape (4, 22).

In order to understand the origin of polytypy, we might first consider how novel aposematic phenotypes persist within a population by examining the underlying mechanisms that promote or constrain diversity in aposematic signals. Species that are both polymorphic and polytypic are ideal subjects to explore the factors that promote or constrain phenotypic diversity within and among populations (Table 1). With these, we can examine how natural populations of predators react to novel phenotypes and

Significance

With our comprehensive set of field (model survival), laboratory (controlled learning, palatability, toxin analysis), and molecular data, we provide evidence that polymorphism can persist in an aposematic population, despite expectations of positive frequency-dependent selection. We show that this can happen if prey species carrying a strong signal can exploit predator learning to elicit broad avoidance of many signals, even if predators only have experience with a single signal. This could allow novel signals to be protected within a population of aposematic prey. Thus, under the expectations of broad generalization coupled with limited gene flow, weak aposematic signals can persist, contributing to the overall diversity of signals found within aposematic species.

Author contributions: J.P.L., B.R., A.F., J.M., R.A.S., and B.P.N. designed research; J.P.L., B.R., A.F., A.B., R.A.S., E.A.C., and B.P.N. performed research; J.P.L., B.R., A.B., R.A.S., and R.B.J. analyzed data; and J.P.L., B.R., and B.P.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data in this paper have been deposited in the Jyväskylä University data repository (DOI: 10.17011/jyx/dataset/65263). Molecular data are available on the University of Mississippi eGROVE (https://egrove.olemiss.edu/biology_facpubs/4/).

[1]J.P.L. and B.R. contributed equally to this work.

[2]Present address: Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697.

[3]To whom correspondence may be addressed. Email: jplarry@gmail.com or bibiana.rojas@jyu.fi.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1901872116/-/DCSupplemental.

how they react to, and learn, aposematic signals. By examining 2 populations of the dyeing poison frog (*Dendrobates tinctorius*) from northeastern French Guiana that differ in color, but not pattern (Fig. 1*A* and *SI Appendix*, Fig. S1), we sought to explore patterns of natural selection by avian predators and infer the conditions necessary to sustain the evolution of novel signals. In this study, we tested the responses of wild avian predators to novel phenotypes in situ, and used ex situ experiments with model predators to explore learning and response to visual and chemical signals. Coupled with analysis of toxin profiles and population genetic patterns, these findings present a multifaceted perspective of a complex evolutionary phenomenon and reveal a mechanism for the diversification of aposematic signals and the species that bear them (Table 1).

Results

Gene Flow Between Populations. Analysis of sequence data obtained from our 3-enzyme restriction-site–associated DNA (3RAD) libraries yielded 1,505 loci that included at least 1 SNP. For both populations, estimated migration rates (26) were found to be <0.0001 (Fig. 1*A*), effectively indicating complete genetic isolation. Similarly, analysis of population structure (27) best supported a model of 2 genetic populations with individuals from the yellow population and the white population grouped separately with very little genetic admixture (Fig. 1*A*).

Avoidance of Novel Phenotypes. According to previous work, birds are the most likely selective agent driving phenotypic diversity in these and other poison frogs (e.g., refs. 28–30). Thus, our results focus only on avian attacks, although it is worth noting that our models were also attacked by mammals and arthropods (*SI Appendix*, Fig. S2).

In the white population, we recovered 364 yellow-striped, 362 white-striped, and 367 solid white models for a total of 1,093 of the 1,136 models deployed. Of the 1,093 models, 9 yellow-striped, 23 white-striped, and 16 solid white models were unambiguously attacked by avian predators, for an overall attack rate of 4.39%. (*SI Appendix*, Fig. S2). In the yellow population, we recovered 439 yellow-striped, 439 white-striped, and 440 solid yellow models for a total of 1,318 of 1,378 models deployed. Bird predators attacked 20 yellow-striped, 23 white-striped, and 23 solid yellow models (*SI Appendix*, Fig. S2), for an overall attack rate of 5.01%.

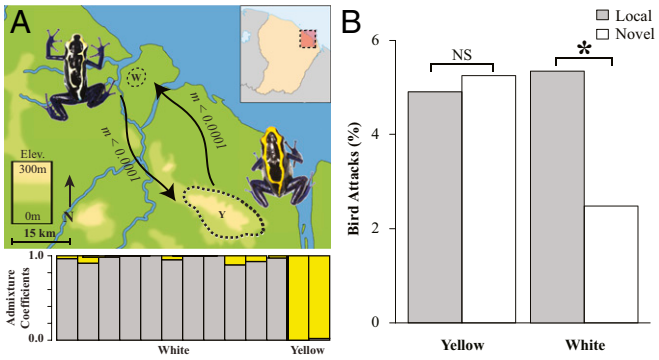

Fig. 1. Distribution, population structure, and predation in the studied populations. (*A*) Map in northeastern French Guiana of the 2 populations displaying migration rate (m) between the 2 populations, white (W) and yellow (Y), and population admixture with each bar representing an individual, and (*B*) distribution of attacks on clay models within the 2 populations. Local/novel indicates local or novel colors to the respective population. *$P < 0.05$; NS, nonsignificant results.

We found a significant interaction between population and color, such that the novel color (yellow) in Matoury was proportionately less attacked than the local one (white) (estimate ± SE = −1.131 ± 0.509, $z = -2.224$, $P = 0.026$), but not between population and pattern (estimate ± SE = −0.542 ± 0.458, $z = -1.183$, $P = 0.237$) (Fig. 1*B*). However, the overall attack rate was not significantly different for the 2 populations (estimate ± SE = 0.3516 ± 0.3145, $z = 1.118$, $P = 0.2635$), and birds exhibited no differential avoidance of color (estimate ± SE = 0.147 ± 0.314, $z = 0.469$, $P = 0.639$) or pattern (estimate ± SE = 0.144 ± 0.313, $z = 0.461$, $P = 0.645$).

Avoidance Learning and Generalization. With model predators, we found a significant interaction between color and chloroquine concentration affecting the latency to attack (Cox regression: estimate ± SE = 0.653 ± 0.289; $z = 2.26$; $P = 0.024$). Similarly, the interaction between color and chloroquine concentration had a significant effect on the number of trials in which the chicks attacked the mealworm (estimate ± SE = 0.564 ± 0.239; $z = 2.36$; $P = 0.018$). Based on this, mealworms with a high chloroquine concentration (highly unpalatable) presented on the

Table 1. Summary table of the questions, predictions, and supporting literature used for this study

Questions	Study	Prediction of effect on phenotypic diversity	Theory support
Is there gene flow between populations?	Genetic (ddRAD) test of population connectivity	Constrain: The homogenizing effects of genetic exchange hinder phenotypic divergence	(91, 92)
How do predators respond to known and novel signals in the 2 populations?	Plasticine clay models in the field	Constrain: Positive frequency dependent selection will act against novel phenotypes.	(10, 93)
	Plasticine clay models in the field	Promote: Neophobia/dietary conservatism will make predators weary about novel forms, allowing them to thrive	(41–44)
How do predators learn to avoid different phenotypes, and when learned, do they extend experience to novel signals?	Learning and generalization assays using naïve model predators	Constrain: Predator learning will favor known phenotypes and select against novel forms	(81, 94, 95)
Do alkaloid profiles vary between populations?	Alkaloid characterization between populations	Promote: Differential alkaloid profiles will correspond to differential unpalatability and promote signal honesty	(33, 96)
Does predator response vary in relation to different alkaloid profiles?	Behavioral response to known amounts of alkaloids by model predators	Promote: Differential predator response to unpalatability will promote signal honesty	(34, 40)

These questions seek to better understand the overall question of how phenotypic diversity can persist in aposematic species.

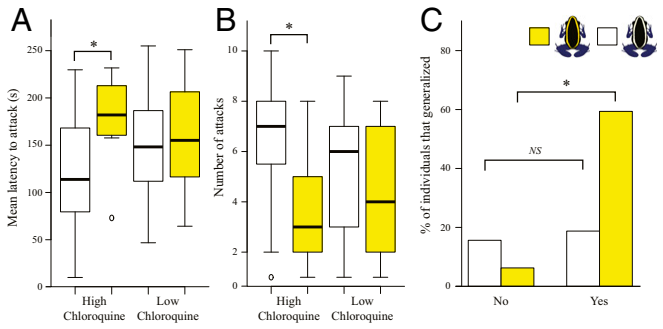

Fig. 2. Results of the learning (*A* and *B*) and generalization (*C*) experiments for white and yellow models. Chickens (*G. gallus domesticus*) were exposed to either the yellow or white treatment which were each split into a high (10%) and low (5%) chloroquine treatment. The results are characterized by (*A*) mean latency to attacking distasteful prey and (*B*) number of trials in which a bird attacked a mealworm. (*C*) Bars in white represent the proportion of birds that learned avoidance of the white signal and were exposed to a novel yellow signal while the yellow bars are the opposite (learned yellow, exposed to novel white). These bars represent both low and high chloroquine treatments combined. Boxes denote the median and the 25th and 75th percentiles of data distribution. Vertical lines indicate data range, and circles represent outliers in data distribution. Significant differences between treatments are denoted by $*P < 0.05$ and nonsignificant (NS).

white signal were more likely to be attacked, elicited a shorter latency to attack (Fig. 2*A*), and were subject to a higher number of attacks across trials (Fig. 2*B*) in comparison to highly unpalatable mealworms offered in association with a yellow signal. However, whether or not the chicks learned the offered signal depended only on the signal color (estimate ± SE = 2.398 ± 0.870; $z = 2.755$; $P = 0.006$), and not on the chloroquine concentration (estimate ± SE = 0.878 ± 0.780; $z = 1.125$; $P = 0.260$). Thus, chicks that were offered mealworms in association with a yellow signal were more likely to learn to avoid them, regardless of distastefulness.

Chloroquine concentration had no effect on generalization of the alternative signal (estimate ± SE = −0.911 ± 0.999; $z = -0.911$; $P = 0.362$), but signal color did (estimate ± SE = 1.950 ± 0.974; $z = 2.001$; $P = 0.045$), such that chicks that had learned to avoid yellow were more likely to also avoid white than the converse (Fig. 2*C*).

Population Variability of Alkaloids. We found that the white population has a higher amount of alkaloids than the yellow population (white mean ± SE: 471 ± 80 μg per frog skin; yellow mean ± SE = 249 ± 33 μg per frog skin; generalized linear model [GLM]: estimate ± SE = 244.45 ± 105.33, $z = 2.321$; $P = 0.036$), but the amount of alkaloids does not depend on the frog's size (estimate ± SE = 21.47 ± 20.85; $z = 1.030$; $P = 0.321$). Alkaloid composition was also significantly different between the 2 populations (global $R = 0.783$; $P \leq 0.001$) (Fig. 3). The white population possessed 49 different alkaloids (with multiple isomers for some) representing 11 different structural classes. These included 8 alkaloids unique to the white population that had not previously been isolated from poison frogs. The yellow population possessed 46 different alkaloids (with multiple isomers for some) representing 12 different structural classes. This also included 9 alkaloids unique to the yellow population that had not previously been isolated from poison frogs (*SI Appendix*, Table S1). One yellow individual had 8 times more alkaloids than the population average (yellow outlier = 2,040 μg per frog skin), likely the result of measurement error, and this outlier was excluded from analysis.

Unpalatability. We conducted 2 skin content assays to assess unpalatability: 1 that examined natural variation of skin contents among individual frogs (unpalatability assay A) and a second that controlled for skin content composition (via dry mass) to examine how composition affected response (unpalatability assay B). Skin extracts were added to oats and given to birds to gauge behavioral response (*Methods*). In unpalatability assay A, we found both populations to elicit significantly higher beak wiping (yellow: estimate ± SE = 1.627 ± 0.485; $z = 3.35$; $P < 0.001$; white: estimate ± SE = 1.318 ± 0.461; $z = 2.856$; $P < 0.01$) (Fig. 4*A*) than the ethanol-soaked oats, but no differences between the 2 populations. Oats coated with the yellow population extracts were eaten significantly less than control oats (estimate ± SE = −0.6807 ± 0.2827; z = −2.408; $P = 0.016$) (Fig. 4*C*), while those coated with extracts from the white population did not differ from controls. These results seem to point at both populations being unpalatable to some extent, but a tendency for the frogs from the Kaw Mountains (yellow) to be less palatable than those from Matoury (white).

Unpalatability assay B revealed that birds exposed to skin extracts from the yellow population showed aversive behavior (beak wiping) at a significantly higher rate than those exposed to the ethanol control (estimate ± SE = 1.250 ± 0.520, $z = 2.400$, $P = 0.016$) (Fig. 4*B*) and to the extracts of white frogs (estimate ± SE = 1.030 ± 0.380, $z = 2.669$, $P = 0.008$) (Fig. 4*B*), suggesting that yellow frogs are more unpalatable. These differences were not influenced by the quantity of alkaloids found in frogs' skin (estimate ± SE = 0.189 ± 0.226, $z = 0.838$, $P = 0.402$). Beak wiping in response to the extracts of white frogs did not differ from the ethanol-only control oats (estimate = 0.153 ± 0.520, $z = 0.293$, $P = 0.769$), and neither population differed from the controls in the proportion of oats eaten by the birds (Fig. 4*D*). Together, these 2 assays suggest that both populations have enough alkaloids to negatively affect avian predators (unpalatability assay A), but the yellow population appears to have more aversive compounds (unpalatability assay B) despite its lower quantity of alkaloids.

To facilitate synthesis of the results of our 5 assays, we summarize them in Table 2 and reference the assay number in the discussion below.

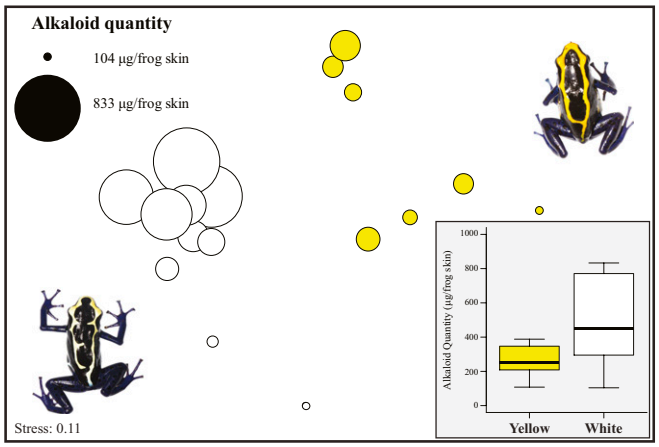

Fig. 3. nMDS plot of variation in alkaloid composition between the 2 populations and distribution of quantity variation (*Inset*). Each circle represents an individual frog, and the distance between symbols is proportional to the difference in alkaloid composition. The diameter of each circle is scaled to represent the quantity of alkaloids present in that frog (microgram per frog skin). Alkaloid composition is significantly different between the 2 populations (global $R = 0.783$; $P \leq 0.001$). *Inset* box plot depicts the range of quantities of alkaloids in the yellow ($n = 7$) and white ($n = 10$) populations.

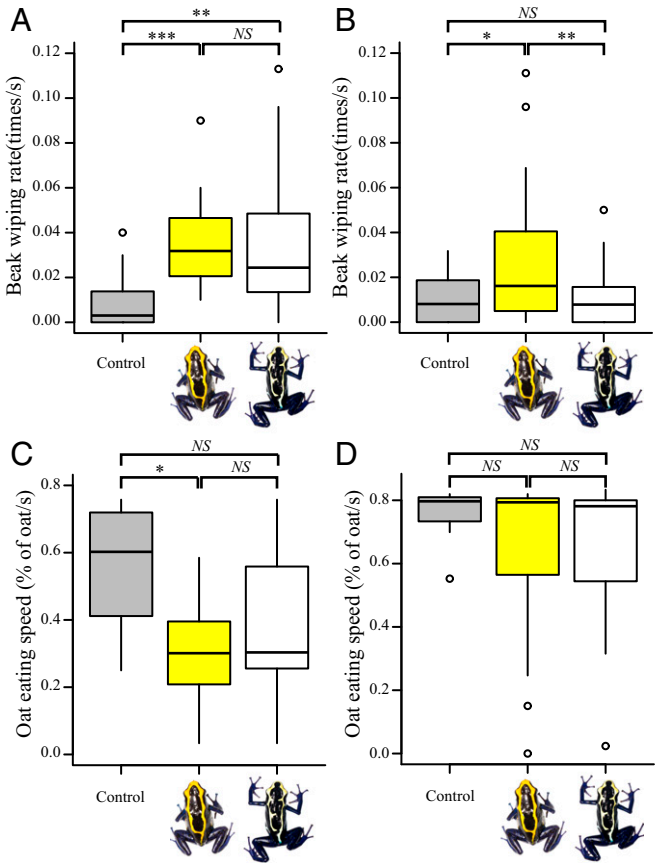

Fig. 4. Results of unpalatability experiments using blue tits (*C. caeruleus*) when comparing natural variation of skin contents (unpalatability assay A: *A* and *C*) and dry mass controlled proportions (unpalatability assay B: *B* and *D*). Beak wiping is significantly higher in response to extracts from both populations compared to controls in (*A*) unpalatability assay A but (*B*) only yellow differed from the control and white when dry mass was controlled (unpalatability assay B). (*C*) Natural variation resulted in proportionately fewer oats being eaten but, interestingly, this pattern disappeared when (*D*) controlling for dry mass. Boxes denote the median and the 25th and 75th percentiles of data distribution. Vertical lines indicate data range, and circles represent outliers in data distribution. Differences denoted by ***$P < 0.001$; **$0.001 < P < 0.01$; *$0.01 < P < 0.05$; and NS, nonsignificant.

Discussion

Aposematic signals under strong +FDS are expected to quickly reach an adaptive peak (4). Poor or ineffective signals should, in turn, quickly be selected against as the population is driven to an optimal signal given the ecological and evolutionary conditions present. We provide evidence that, while having a selective disadvantage in the field (assay 2) and the laboratory (assay 3), a poor aposematic signal can persist in the absence of gene flow (assay 1). This research supports the idea of shifting balance (24) as a possible explanation for the existence of polytypic populations. The yellow and white populations examined here appear to represent different adaptive optima based on local conditions. While the Matoury (white) population possesses a less effective aposematic signal, this may represent a suboptimal "adaptive peak" rather than a "valley." Given the asymmetry between the 2 populations in terms of protection afforded to novel signals (assays 2 and 3), our research highlights a plausible scenario whereby a novel white signal may have evolved within a yellow population and subsequently was isolated through stochastic processes (i.e., genetic drift, founder effects, and so forth). Following this, the white population, which was limited by gene flow, climbed a suboptimal adaptive peak among neighboring populations, resulting in fixation of the white phenotype. We emphasize, however, that the origin of the white phenotype is currently unknown but does warrant further research.

In examining the Kaw and Matoury populations, we have demonstrated that a weak aposematic signal can persist and, further, how novel signals could persist in a population despite expectations of strong +FDS operating against individuals that display such signals. The 2 populations examined here have virtually no gene flow between them (assay 1), suggesting that more adaptive alleles from the yellow population are not inundating the population and overwhelming the weak white signal (assay 3). We found that predators have greater difficulty learning avoidance of the white signal (assay 3), and those that do are less likely to generalize novel signals (assay 3). Furthermore, we showed that yellow signals elicit a stronger aversive response in model predators than white signals. This apparent weakness of the white signal is not overcome by the presence of greater quantities or diversity of alkaloid defenses (assays 4 and 5). While we found differences in alkaloid quantity between the 2 populations (assay 4), with higher amounts in the white population, we found no effect of alkaloid quantity influencing avian behavior (assay 5). Alkaloid quantities ranged from 104 μg per frog skin to 833 μg per frog skin (Fig. 3), yet these did not have an effect on predator behavior. The observed differences in our unpalatability assays (Fig. 4) are thus likely the result of different compositions between populations, which we demonstrate are distinct from one another (Fig. 3). Coupled with the interaction observed in the chloroquine learning experiments (assay 2), we can conclude that increased unpalatability (assay 5) can augment the effectiveness of a strong warning signal, and that there is likely a subset of the alkaloids present in these frogs that is driving predator response. We demonstrate that a weak signal cannot be overcome simply by a stronger defense, and conversely, a strong signal can further be enhanced by an increased defense (assay 3).

Prior research has focused on toxicity of alkaloids via injections (31, 32), even suggesting that unpalatability cannot be empirically assessed for toxins (33). Researchers are beginning to work on unpalatability assays of anuran toxins in invertebrates (34–36), but our unpalatability assay provides a practical assessment of distastefulness, as opposed to toxicity (33), of anuran alkaloids in avian predators in a controlled manner, without having to use live frogs (33, 37). While informative, toxicity assays do not give an evolutionary context to defensive compounds as predators experience these alkaloids through taste, not via toxicity effects in the bloodstream (38). Experiments with live frogs, on the other hand, do not account for interindividual or interpopulational differences in toxin profiles (as sampling complete alkaloid profiles is not possible with live frogs) or alkaloid amount, which we have found not to relate directly to unpalatability. Our findings suggest that relationships between toxicity and coloration (i.e., refs. 32 and 39) may actually be more complex than previously proposed or nonexistent all together. This finding aligns with a recent study on nudibranchs, where it has become clear that toxicity and taste must be disentangled for a better understanding of chemical defenses in aposematic species (40). To this end, we also demonstrate that it is possible to quantitatively assess variation in unpalatability among skin secretions independent of alkaloid content. Future studies would benefit from greater within-population sampling to better characterize the extent of intrapopulation variation in the amount of toxic compounds present.

With our findings we also show how experienced model avian predators respond to novel signals after learning either yellow or white frog patterns (assay 3). Predators were more likely to avoid attacking a novel signal if they had previously learned to avoid yellow signals. Importantly, this is corroborated by our field experiments with wild predators and clay models in the yellow

Table 2. Summary of results from the different assays conducted for this study

Study	Questions	Species tested	Results and conclusions
Assay 1: Population genetics	Is there gene flow between yellow and white?	White and yellow *D. tinctorius*	Fig. 1*A*; Virtually no migration ($m < 0.0001$) between populations
Assay 2: Clay models	How do predators respond to known and novel signals in the yellow and white populations?	Native avian predator community	Fig. 1*B*; White stripes attacked more in white stripe population; no differences among models in yellow population
Assay 3: Learning experiments	How do predators learn to avoid yellow and white, and when learned, do they extend experience to novel signals?	Naïve chickens	Fig. 2; Naïve predators learned to avoid yellow faster, hesitate longer, and extend experience to novel signals when compared to white
Assay 4: Alkaloid profiles	Do alkaloid profiles vary between populations?	White and yellow *D. tinctorius*	Fig. 3 and *SI Appendix*, Table S1; Populations are distinct in their alkaloid composition (number, type, quantity). Wide variation in alkaloid profiles observed among populations.
Assay 5: Unpalatability	Does predator response vary in relation to different alkaloid profiles?	Blue tits	Fig. 4; Populations contain alkaloids more aversive than controls, but yellow appears to have more effective secondary defenses Skin secretions significantly different from controls, but variably different among populations, depending on concentration/proportions

population (assay 2) where, counter to expectations of anti-apostatic selection (6), we observed no difference in attack rates between known and novel phenotypes. This indicates that the experienced predators in these areas are likely to avoid attacking novel signals, thereby allowing novel signals to persist in the population (if/when they evolve), likely due to phenomena such as neophobia or dietary conservatism (41–44). Interestingly, in Matoury we observed that the local (white) form is attacked significantly more than the novel, yellow-striped form, which, based on our laboratory experiments, can be explained by the difficulty predators have in learning avoidance of the white form and, when they do learn to avoid white, they were less likely to extend that avoidance to novel forms (assay 3). Avian predators' avoidance of novel yellow-striped models in the Matoury area could further be explained by our laboratory experiments where birds had greater latency to attack yellow models, suggesting that neophobia may play a role in protecting novel yellow signals.

The most obvious question from the above results is: Why does the white population persist despite possessing a weak aposematic signal? If there was an exchange of alleles between the white and yellow populations, we would expect strong selection against white coloration and the population should quickly become fixed for yellow stripes. However, we observe virtually no gene flow between the 2 populations, and as a result, the individuals in the white population present a suboptimal aposematic signal (45, 46), which may be traversing from 1 adaptive peak to another, climbing a "yellow peak," or be atop a suboptimal peak, unable to cross a fitness valley to a more fit yellow phenotype. It is also important to note that our laboratory experiments do demonstrate that avian predators can learn to avoid the suboptimal (white) aposematic signal and that the skin secretions of the Matoury population are distasteful, so while it may not be as efficient a signal as yellow, it does offer some protection. In fact, it has been previously suggested that weak signals could evolve and be maintained if predators are variable in their response toward defended prey (18, 47, 48), and if accompanied by a behavioral adaptation (49). Individuals in the white population, for example, appeared to be more secretive and less bold than individuals in the yellow population. This behavior may aid in the persistence of this population despite their suboptimal signal.

Our research provides important insight into the origin and persistence of aposematic signals. Notably, we provide a reasonable mechanism by which a novel signal, even a weak one, persists via broad generalization of existing/known signals, despite the expectation of +FDS. While we did not observe white individuals in the yellow population, our results support the idea that if novel signals were to arise in the yellow population, they would be protected by the strong yellow signal. This possibility is supported by the existence of phenotypic polymorphism in some yellow populations of *D. tinctorius*, where a variety of patterns indicates that phenotype is not constrained (13), and that predators may indeed generalize among them (30). Interestingly, such color pattern variation in populations with white in their signals has never been reported. Consequently, this system may support the idea of an adaptive landscape. It is important to note that while this is but 1 scenario highlighting how a multipeaked adaptive landscape could be responsible for poor signals, there may be others that warrant further research.

Much of our knowledge on the evolution of polymorphism/polytypism in aposematic taxa has largely focused on mimicry complexes. In these instances, Batesian (33, 50–52) and Müllerian (5, 37, 53–56) mimics can diversify as selection matches them to models and both can drive one another to new phenotypes. These systems, understandably, allow researchers to investigate mechanisms promoting phenotypic diversity. While this is an important mechanism in driving phenotypic variation in defended prey species, our findings differ in that we provide a mechanism for the origin and persistence of intrapopulation polymorphism in the absence of a model species. As many aposematic species are not involved in mimicry complexes, this finding, in particular, is important for our understanding of how aposematic signals evolve and diversify.

Given theoretical constraints imposed by selection, the existence of polymorphic aposematic signals is a difficult phenomenon to explain (but see ref. 57 for a review where this paradigm is challenged). Herein, we have documented 2 mechanisms that can promote phenotypic diversity within and among populations. Polymorphism in aposematic signals may not be selectively disadvantageous due to broad generalization by predators that have learned to avoid a sufficiently strong signal, and weak signals may occur at high frequencies (up to fixation) when gene flow among populations is limited. Novel aposematic signals can indeed persist, and some signals, despite being weak, may endure and even become fixed in a population. Together, these findings contribute to our understanding of the forces generating the diverse array of aposematic signals seen across the animal kingdom.

Methods

Field Sampling. The 2 focal populations were Matoury, French Guiana (4.897°, −52.350°), where frogs have white stripes, and the Kaw Mountains (4.542°, −52.158°), French Guiana, where frogs have yellow stripes, hereafter referred to as white and yellow, respectively. These sites were chosen because frogs have the same pattern (stripes on a black body with blue legs) but differ in the color of their stripes (white and yellow) (Fig. 1*A*). We collected frogs from white ($n = 10$) and yellow ($n = 8$) populations in 2013 and 2014. Notably, we initially collected 2 frogs from the Kaw Mountains in 2013, which were used for molecular analysis, but to increase sample size for alkaloid variability, we collected 6 additional frogs in 2014. Frogs were killed by deep-freezing to avoid possible contamination from killing agents (58). Livers were preserved in 95% ethanol, and whole skins were preserved in 100% methanol for subsequent alkaloid analysis. Skins were stored in 4-mL glass vials with PTFE caps.

Gene Flow between Populations. To assess gene flow between populations, we examined thousands of variable sites (SNPs) across the *D. tinctorius* genome that could then be used to estimate population demography. To do this, we employed the 3RAD method of refs. 59 and 60. Samples were digested in a reaction that contained 100 ng of DNA, 20 units of each of 3 restriction enzymes (XbaI, EcoRI-HF, and NheI-HF; New England Biolabs), 1× NEB Cutsmart Buffer, 1 μL of both a forward and reverse double-stranded adapter at 5 μM and water to 15 μL total volume. Enzymes were chosen based on an in silico digest of the genome of *Xenopus laevis*, which suggested these enzymes would produce ∼3,500 unique loci within the desired size range of 300 to 350 bp. Each adapter contained a 6- to 9-bp barcode and each sample was assigned a unique forward and reverse barcode combination. Digestions were carried out at 37 °C for 1 h and followed immediately by the addition of ligation mix. As adapters were already present in digestion reactions, the ligation mix consisted of 1.5 μL of 10 mM rATP (New England Biolabs), 0.5 μL 10× ligase buffer (New England Biolabs), 100 units T4 DNA Ligase (New England Biolabs), and water to 5 μL. Ligation reactions were incubated for 2 cycles of 22 °C for 20 min and 37 °C for 10 min followed by 20 min at 80 °C. Ligated samples were then cleaned by mixing with 1.2 volumes of Sera-mag Speedbeads (Fisher Scientific; prepared as in ref. 61), 2 washes of 70% ethanol, and eluted from beads with 20 μL of IDTE pH 8 (Integrated DNA Technologies). Samples were then amplified in 20-μL reactions with 10 μL of postbead template, 1× HIFI Buffer (Kapa Biosystems), 0.75 μL dNTPs (Kapa Biosystems), 0.5 μL HIFI DNA polymerase (Kapa Biosystems), 0.5 μM iTru5 and iTru7 primers containing sample-specific 8-bp barcodes, and H_2O to 25 μL total volume. Library amplification began with 3 min at 95 °C followed by 16 cycles of 98 °C for 20 s, 55 °C for 15 s, and 72 °C for 30 s, followed by a 5-min extension of 72 °C.

Success of individual PCRs was determined by gel electrophoresis. When digestion, ligation, and amplification were successful, 10 μL from each PCR were pooled into sets of 24. Each pool was concentrated down to 30 μL using a Qiagen PCR purification column and then run in a single lane of a 1.5% Pippin Prep gel cassette (Sage Science) selecting for fragments sized 400 to 450 bp. Selected sized fragments were removed from Pippin elution wells and used directly in qPCR library quantification (Kapa Biosystems) and Illumina Sequencing on a NextSEq. 500 to obtain single-ended, dual-indexed reads of 75 bp. Libraries were sequenced in the National Center for Natural Products Research, University of Mississippi.

Raw reads were downloaded from Basespace using basespacerundownloader (Illumina) and demultiplexed using BCL2FASTQ (Illumina) allowing up to 2-bp errors in barcodes as our barcodes had 3-bp degeneracy built in. After demultiplexing, each sample was trimmed from the beginning of each read to remove remaining internal (adapter) barcode and restriction site overhang. All samples were then trimmed to 58 bp in length to prevent downstream difficulties in assignment of homology. Loci were identified within individuals and aligned among individuals using *pyRAD* (62).

Population Genetic Analysis. We used the SNP data with the Bayesian coalescent program G-PhoCS (Generalized Phylogenetic Coalescent Sampler) (26) to estimate rates of migration between the 2 populations. Six independent, simultaneous runs were conducted, each with a 500,000 Markov chain Monte Carlo iterations, which were then compiled for a total of 3,000,000 iterations. We used a 10% burn-in independently for each run, which resulted in 2,700,000 iterations from which summary statistics were calculated. We defined a constant mutation rate in order to conduct these analyses as a constant mutation rate would be a null assumption when lacking data, suggesting fluctuating mutation rates over evolutionary history (26). Bidirectional migration rates between the 2 populations were then calculated in G-PhoCS. We examined population structure between the 2 populations using *snmf* function in the *LEA* R package (27). We tested from 1 up to 10 populations and ran 1,000,000 iterations.

Avoidance of Novel Phenotypes. Plasticine clay model experiments are commonly used to assess predator responses to novel aposematic signals (e.g., refs. 29 and 63–65). These allow observation of predation attempts and provide insight into selection pressures on various phenotypes. We constructed 45-mm-long (snout to vent) replica frogs using Van Aken polymer modeling clay (29, 66), which does not harden and retains evidence of predation attempts. We poured melted clay into silicone molds made from plastic model replicas (67). 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 (28). Both stripes and eyes were created with clay and affixed to models (63).

Models were placed along transects on which 3 model types were randomized with 1 model every 5 m (29, 68). The 3 model types for each location were local color and pattern, local color and novel pattern, local pattern and novel color, ensuring that at least 1 component of the aposematic signal was familiar to predators. Transects were separated by at least 100 m. We placed 7 transects from 495 m long to 1.5 km long in the white population and 13 transects from 495 m long to 1.5 km long in yellow population. Models (1,378 and 1,136 in the yellow and white populations, respectively) were left for 72 h to allow for predation attempts (29, 63). 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 (69, 70). The nonavian 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 (71, 72). Similarly, snakes, which are some of the few known predators of poison frogs (73, 74), can also see color, but they are highly motivated by movement and chemical cues, making it highly unlikely that they would attack clay models.

Clay Model Data Analysis. If models had multiple bite marks, they were scored as a single predation attempt as we were not able to determine whether 1 predator attacked multiple times or multiple predators attacked once (29). Missing models were excluded from the analysis (43 of 1,136 and 57 of 1,378 for white and yellow, respectively). We used a general linear mixed model (GLMM) 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.

Avoidance Learning and Generalization. As avian predators are the likely drivers of aposematic signal evolution in *D. tinctorius* (29, 30, 63), we used a model avian predator, the domestic chicken (*Gallus gallus domesticus*), to assess how naïve predators learn and extend experience (generalize) with aposematic signals. Chickens are well-known for their capabilities of recognizing and learning different colors quickly (75), and thus widely used in this type of experiment (33, 37, 76–79). We trained naïve 1-wk-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 (*SI Appendix*, Fig. S3). 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 (80, 81). Over successive trials, we examined whether chicks would learn to avoid distasteful mealworms associated with an aposematic signal, with learning defined as 3 consecutive refusals of a distasteful mealworm (82). 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 to 3 h. Chicks were equally divided (n = 15 per treatment; 60 chicks total) into 4 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 or distastefulness (chloroquine concentration). The trials were done in 30-cm × 60-cm wooden compartments under full spectrum lights. Chicks were placed individually into a compartment and allowed to habituate for 2 h. Chicks were food-deprived during this acclimation period to ensure motivation to feed during the trials. Once in the compartment and after the 2-h 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 3 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) (*SI Appendix*, Fig. S3). 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 3 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 3 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 5 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 3 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 afterward. See *SI Appendix*, Fig. S3 for a schematic on how the learning and generalization experiments were conducted.

Learning Experiment Data Analysis. We analyzed the data from the learning experiments using a GLMM in 2 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 2, 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 10, and whether or not the chicks learned to avoid the signal they were presented (as defined by 3 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 2. 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 (83), with the RStudio interface and using the packages *lme4* (84) and *coxme* (85).

Population Variability of Alkaloids. Whole frog skins were collected from 18 specimens (8 yellow, 10 white) and stored in 4 mL of methanol. Alkaloids were extracted and analyzed from skins using the procedure outlined in ref. 34, which is described here briefly. One milliliter of each methanol extract was transferred into a 10-mL conical vial and 10 μg of nicotine as an internal standard [(-)-nicotine ≥ 99%, Sigma-Aldrich] and 50 μL of 1N hydrochloric acid (to acidify the solution) were added. Each sample was then mixed and evaporated with nitrogen gas to a volume of 100 μL. Subsequently, each sample was diluted with 200 μL of deionized water. The samples were then extracted with 4 300-μL portions of hexane. The resulting hexane (organic) layer was disposed of and the remaining aqueous layer was basified with saturated sodium bicarbonate. Basicity was tested with pH paper. Once the pH was greater than 7, each sample was extracted with 3 300-μL portions of ethyl acetate. Anhydrous sodium sulfate was added to the newly extracted mixture to remove any excess water. The remaining samples were carefully evaporated to dryness with nitrogen gas. Alkaloid fractions were resuspended in 100 μL of methanol and stored at −20 °C.

Gas chromatography-mass spectrometry (GC-MS) was performed for each individual frog on a Varian Saturn 2100T ion-trap MS instrument, which was coupled to a Varian 3900 GC with a 30-m × 0.25-mm inner diameter Varian Factor Four VF-5-ms fused silica column. GC separation of alkaloids was achieved using a temperature program from 100 to 280 °C at a rate of 10 °C per minute with helium as the carrier gas (1 mL/min). Each alkaloid fraction was analyzed in triplicate with electron impact MS and once with chemical ionization (CI) MS with methanol as the CI reagent.

Individual alkaloids of *D. tinctorius* were identified based on comparison of MS properties and GC retention times with those of previously reported alkaloids in dendrobatid frogs (86). Alkaloid quantities for each individual frog were calculated by comparing the average observed peak area of individual alkaloids to the average peak area of the nicotine standard from the triplicate electron ionization-MS analyses using a Varian MS Workstation v.6.9 SPI. Only alkaloids that were present in quantities of ≥0.5 μg were included in the analyses (34). We compared average alkaloid quantity between populations with a GLM with Gaussian distribution using the frog's body size as a covariate. Differences in alkaloid composition among individual frogs from each population were graphically visualized using nonmetric multidimensional scaling (nMDS), and statistical differences were examined with a 1-way analysis of similarity (ANOSIM). nMDS and ANOSIM statistics were based on Bray–Curtis similarity matrices, and were performed in PRIMER-E (v5).

Predator Response to Chemical Defenses. Prior research into dendrobatid alkaloids has largely quantified toxicity through either LD50 (31) or subcutaneous injections (32, 33) (but see refs. 34 and 35 for assays on alkaloid unpalatability using arthropods as predators), which, while informative, are limited in their evolutionary significance, as predators experience alkaloids through ingestion, not injection into the bloodstream or musculature (38). In

other studies (33, 37), naïve chickens have been offered live frogs, which does not allow investigators to account for interindividual variation in toxin profiles and alkaloid amounts. This is because the amount of alkaloids in living frogs cannot be known and, thus, the most aversive response may be assumed to result from the highest amount of alkaloids. Consequently, we sought to examine unpalatability of skin extracts, as this is what drives predator decision making, in a controlled manner. We used an additional 1 mL of the methanol extract from the same skins of the 18 assayed frogs (10 white, 8 yellow) and prepared toxins for 2 unpalatability assays using wild-caught blue tits (*Cyanistes caeruleus*), which have been shown to be able to respond to subtle variations in animal chemical defenses (87, 88). These methanol extracts were not fractionated and were simply whole extracts of skin contents.

We assessed skin contents in 2 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 nonalkaloid content impact predator response. We note, however, that predator responses could be skewed if nonalkaloid content varied among individuals (see below). The 2 different assays were identical except for the skin secretion toxin preparation. The purpose of doing 2 assays was to attempt to 1) understand how toxins influence predator response and 2) ensure that nontoxin components of skin secretions do not disproportionately affect results and mask potential aversion to the alkaloids. For each assay, we tested 1 frog sample with 1 bird (white, $n = 10$, yellow, $n = 8$), with 6 control birds for the first assay and 7 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 0.5 mL 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, and so forth, 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 (approximately 1:1 toxin mass to ethanol). Toxins extracts were then transferred to oats (15 μL per 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 nonalkaloid content varied among frogs, especially in a manner independent of alkaloid mass, this may have unintentionally altered our concentrations among samples, eliciting nonbiologically 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 Burdfield-Steel et al. (88). Aversive behavior (i.e., beak wiping) and percentage of oat eaten were recorded.

For each assay, each of 2 oats were soaked with 15 μL of extract of 1 frog skin and left for 24 h at room temperature to ensure that all ethanol had evaporated. Two other oats were soaked—and subsequently allowed to dry for 24 h—with 15 μL 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 (1 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 (89), 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 GLMM using the package *lme4* (84). 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 (83) using the RStudio interface (90), unless stated otherwise.

Ethics Statement. Chick and blue tit experiments were conducted under University of Mississippi Institutional Animal Care and Use Committee protocols 14-025 and 14-026 and the Central Finland Centre for Economic Development, Transport, and Environment and license from the National Animal Experiment Board (ESAVI/9114/04.10.07/2014), and the Central Finland Regional Environment Centre (VARELY/294/2015), respectively.

Data Availability. Data supporting the findings reported in this study are available from the data repository of the University of Jyväskylä (DOI: 10.17011/jyx/dataset/65263).

ACKNOWLEDGMENTS. We thank Tim Muhich, John Constan, Antoine Baglan, and Mathilde Segers for their invaluable help in creating and placing model frogs in French Guiana; Ombeline Vrignaud for providing valuable information and assistance and securing permission for work within the Grand Matoury Natural Reserve; John Endler, Emily Burdfield-Steel, Janne Valkonen, and Carita Lindstedt for providing helpful feedback on an earlier version; Helinä Nisu at the Konnevesi Research Station for help with the blue tit experiment; and Sara Calhim and J. Valkonen for invaluable statistical advice. This work was supported in part by Investissement d'Avenir Grant CEBA ANR-10-LABX-25-01 of the Agence Nationale de la Recherche; the American Society of Ichthyologists and Herpetologists' Gaige Award (to J.P.L.); and a Society for the Study of Amphibians and Reptiles' grant in Herpetology (to J.P.L.). B.R. and J.M. were funded by the Finnish Centre of Excellence in Biological Interactions (Project 28466, to J.M.). B.R. also acknowledges funding from the Academy of Finland (Academy Research Fellowship, project No. 21000042021).

1. E. B. Poulton, *The Colours of Animals: Their Meaning and Use, Especially Considered in the Case of Insects* (Kegan Paul, Trench, Trubner, 1890).
2. G. D. Ruxton *et al.*, *Avoiding Attack: The Evolutionary Ecology of Crypsis, Warning Signals and Mimicry* (Oxford University Press, 2004).
3. J. A. Endler, Frequency-dependent predation, crypsis and aposematic coloration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **319**, 505–523 (1988).
4. J. Mallet, M. Joron, Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annu. Rev. Ecol. Syst.* **30**, 201–233 (1999).
5. F. Müller, *Ituna* and *Thyridia*: A remarkable case of mimicry in butterflies. *Trans. Entomol. Soc. Lond.* **1879**, 20–29 (1879).
6. L. Lindström, R. V. Alatalo, A. Lyytinen, J. Mappes, Strong antiapostatic selection against novel rare aposematic prey. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9181–9184 (2001).
7. H. M. Rowland, J. Mappes, G. D. Ruxton, M. P. Speed, Mimicry between unequally defended prey can be parasitic: Evidence for quasi-Batesian mimicry. *Ecol. Lett.* **13**, 1494–1502 (2010).
8. M. Borer, T. Van Noort, M. Rahier, R. E. Naisbit, Positive frequency-dependent selection on warning color in Alpine leaf beetles. *Evolution* **64**, 3629–3633 (2010).
9. M. Chouteau, M. Arias, M. Joron, Warning signals are under positive frequency-dependent selection in nature. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2164–2169 (2016).
10. J. Mallet, N. H. Barton, Strong natural selection in a warning-color hybrid zone. *Evolution* **43**, 421–431 (1989).
11. P. O'Donald, M. E. N. Majerus, Polymorphism of melanic ladybirds maintained by frequency-dependent sexual selection. *Biol. J. Linn. Soc. Lond.* **23**, 101–111 (1984).
12. H. Ueno, Y. Sato, K. Tsuchida, Colour-associated mating success in a polymorphic Ladybird Beetle, *Harmonia axyridis*. *Funct. Ecol.* **12**, 757–761 (1998).
13. B. Rojas, J. A. Endler, Sexual dimorphism and intra-populational colour pattern variation in the aposematic frog *Dendrobates tinctorius*. *Evol. Ecol.* **27**, 739–753 (2013).
14. R. H. Hegna, J. A. Galarza, J. Mappes, Global phylogeography and geographical variation in warning coloration of the wood tiger moth (*Parasemia plantaginis*). *J. Biogeogr.* **42**, 1469–1481 (2015).
15. C. W. Myers, J. W. Daly, Dart-poison frogs. *Sci. Am.* **248**, 120–133 (1983).
16. A. B. Roland *et al.*, Radiation of the polymorphic Little Devil poison frog (*Oophaga sylvatica*) in Ecuador. *Ecol. Evol.* **7**, 9750–9762 (2017).
17. B. Rojas, Behavioural, ecological, and evolutionary aspects of diversity in frog colour patterns. *Biol. Rev. Camb. Philos. Soc.* **92**, 1059–1080 (2017).
18. J. A. Endler, J. Mappes, Predator mixes and the conspicuousness of aposematic signals. *Am. Nat.* **163**, 532–547 (2004).
19. A. F. Hugall, D. Stuart-Fox, Accelerated speciation in colour-polymorphic birds. *Nature* **485**, 631–634 (2012).
20. C. A. McLean, D. Stuart-Fox, Geographic variation in animal colour polymorphisms and its role in speciation. *Biol. Rev. Camb. Philos. Soc.* **89**, 860–873 (2014).

21. M. Chouteau, B. Angers, Wright's shifting balance theory and the diversification of aposematic signals. *PLoS One* **7**, e34028 (2012).
22. M. Arias *et al*., Crossing fitness valleys: Empirical estimation of a fitness landscape associated with polymorphic mimicry. *Proc. Biol. Sci.* **283**, 1829 (2016).
23. S. M. Gray, J. S. McKinnon, Linking color polymorphism maintenance and speciation. *Trends Ecol. Evol.* **22**, 71–79 (2007).
24. S. Wright, The shifting balance theory and macroevolution. *Annu. Rev. Genet.* **16**, 1–19 (1982).
25. J. Mallet, Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. *Ecol. Entomol.* **35**, 90–104 (2010).
26. I. Gronau, M. J. Hubisz, B. Gulko, C. G. Danko, A. Siepel, Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034 (2011).
27. E. Frichot, O. François, LEA: An R package for landscape and ecological association studies. *Methods Ecol. Evol.* **6**, 925–929 (2015).
28. R. A. Saporito, R. Zuercher, M. Roberts, K. G. Gerow, M. A. Donnelly, Experimental evidence for aposematism in the dendrobatid poison frog *Oophaga pumilio*. *Copeia* **2007**, 1006–1011 (2007).
29. B. P. Noonan, A. A. Comeault, The role of predator selection on polymorphic aposematic poison frogs. *Biol. Lett.* **5**, 51–54 (2009).
30. B. Rojas, P. Rautiala, J. Mappes, Differential detectability of polymorphic warning signals under varying light environments. *Behav. Processes* **109**, 164–172 (2014).
31. J. W. Daly, C. W. Myers, Toxicity of Panamanian poison frogs (*Dendrobates*): Some biological and chemical aspects. *Science* **156**, 970–973 (1967).
32. M. E. Maan, M. E. Cummings, Poison frog colors are honest signals of toxicity, particularly for bird predators. *Am. Nat.* **179**, E1–E14 (2012).
33. C. R. Darst, M. E. Cummings, Predator learning favours mimicry of a less-toxic model in poison frogs. *Nature* **440**, 208–211 (2006).
34. E. M. Murray, S. K. Bolton, T. Berg, R. A. Saporito, Arthropod predation in a dendrobatid poison frog: Does frog life stage matter? *Zoology (Jena)* **119**, 169–174 (2016).
35. S. K. Bolton, K. Dickerson, R. A. Saporito, Variable alkaloid defenses in the dendrobatid poison frog *Oophaga pumilio* are perceived as differences in palatability to arthropods. *J. Chem. Ecol.* **43**, 273–289 (2017).
36. L. M. Schulte, R. A. Saporito, I. Davison, K. Summers, The palatability of Neotropical poison frogs in predator-prey systems: Do alkaloids make the difference? *Biotropica* **49**, 23–26 (2017).
37. A. M. M. Stuckert, P. J. Venegas, K. Summers, Experimental evidence for predator learning and Müllerian mimicry in Peruvian poison frogs (*Ranitomeya*, Dendrobatidae). *Evol. Ecol.* **28**, 413–426 (2014).
38. P. J. Weldon, Poison frogs, defensive alkaloids, and sleepless mice: Critique of a toxicity bioassay. *Chemoecology* **27**, 123–126 (2017).
39. I. J. Wang, Inversely related aposematic traits: Reduced conspicuousness evolves with increased toxicity in a polymorphic poison-dart frog. *Evolution* **65**, 1637–1649 (2011).
40. A. E. Winters *et al*., Toxicity and taste: Unequal chemical defences in a mimicry ring. *Proc. Biol. Sci.* **285**, 1880 (2018).
41. N. M. Marples, J. Mappes, Can the dietary conservatism of predators compensate for positive frequency dependent selection against rare, conspicuous prey? *Evol. Ecol.* **25**, 737–749 (2011).
42. R. J. Thomas, L. A. Bartlett, N. M. Marples, D. J. Kelly, I. C. Cuthill, Prey selection by wild birds can allow novel and conspicuous colour morphs to spread in prey populations. *Oikos* **106**, 285–294 (2004).
43. N. M. Marples, T. J. Roper, D. G. C. Harper, Responses of wild birds to novel prey: Evidence of dietary conservatism. *Oikos* **83**, 161–165 (1998).
44. N. M. Marples, D. J. Kelly, Neophobia and dietary conservatism: Two distinct processes? *Evol. Ecol.* **13**, 641–653 (1999).
45. G. D. Ruxton, M. P. Speed, M. Broom, The importance of initial protection of conspicuous mutants for the coevolution of defense and aposematic signaling of the defense: A modeling study. *Evolution* **61**, 2165–2174 (2007).
46. C. A. McLean, A. Moussalli, D. Stuart-Fox, Local adaptation and divergence in colour signal conspicuousness between monomorphic and polymorphic lineages in a lizard. *J. Evol. Biol.* **27**, 2654–2664 (2014).
47. J. K. Valkonen *et al*., Variation in predator species abundance can cause variable selection pressure on warning signaling prey. *Ecol. Evol.* **2**, 1971–1976 (2012).
48. O. Nokelainen, J. Valkonen, C. Lindstedt, J. Mappes, Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. *J. Anim. Ecol.* **83**, 598–605 (2014).
49. B. Willink, E. Brenes-Mora, F. Bolaños, H. Pröhl, Not everything is black and white: Color and behavioral variation reveal a continuum between cryptic and aposematic strategies in a polymorphic poison frog. *Evolution* **67**, 2783–2794 (2013).
50. H. W. Bates, XXXII. Contributions to an insect Fauna of the Amazon Valley. Lepidoptera: Heliconidæ. *Trans. Linn. Soc. Lond.* **23**, 495–566 (1862).
51. K. Kunte, The diversity and evolution of batesian mimicry in *Papilio* swallowtail butterflies. *Evolution* **63**, 2707–2716 (2009).
52. M. Katoh, H. Tatsuta, K. Tsuji, Rapid evolution of a Batesian mimicry trait in a butterfly responding to arrival of a new model. *Sci. Rep.* **7**, 6369 (2017).
53. W. W. Benson, Natural selection for Miillerian mimicry in Heliconius erato in Costa Rica. *Science* **176**, 936–939 (1972).
54. P. M. Brakefield, Polymorphic Müllerian mimicry and interactions with thermal melanism in ladybirds and a soldier beetle: A hypothesis. *Biol. J. Linn. Soc. Lond.* **26**, 243–267 (1985).
55. D. D. Kapan, Three-butterfly system provides a field test of Müllerian mimicry. *Nature* **409**, 338–340 (2001).
56. P. E. Marek, J. E. Bond, A Müllerian mimicry ring in Appalachian millipedes. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9755–9760 (2009).
57. E. S. Briolat *et al*., Diversity in warning coloration: Selective paradox or the norm? *Biol. Rev. Camb. Philos. Soc.* **94**, 388–414, (2019).
58. R. A. Saporito, T. Grant, Comment on Amézquita et al. (2017) "Conspicuousness, color resemblance, and toxicity in geographically diverging mimicry: The pan-Amazonian frog *Allobates femoralis*". *Evolution* **72**, 1009–1014 (2018).
59. C. F. Graham *et al*., Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Mol. Ecol. Resour.* **15**, 1304–1315 (2015).
60. S. L. Hoffberg *et al*., RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Mol. Ecol. Resour.* **16**, 1264–1278 (2016).
61. N. Rohland, D. Reich, Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22**, 939–946 (2012).
62. D. A. R. Eaton, PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* **30**, 1844–1849 (2014).
63. A. A. Comeault, B. P. Noonan, Spatial variation in the fitness of divergent aposematic phenotypes of the poison frog, *Dendrobates tinctorius*. *J. Evol. Biol.* **24**, 1374–1379 (2011).
64. B. Willink, A. García-Rodríguez, F. Bolaños, H. Pröhl, The interplay between multiple predators and prey colour divergence. *Biol. J. Linn. Soc. Lond.* **113**, 580–589 (2014).
65. D. D. Dell'aglio, M. Stevens, C. D. Jiggins, Avoidance of an aposematically coloured butterfly by wild birds in a tropical forest. *Ecol. Entomol.* **41**, 627–632 (2016).
66. E. D. Brodie, 3rd, Differential avoidance of coral snake banded patterns by free-ranging avian predators in Costa Rica. *Evolution* **47**, 227–235 (1993).
67. J. Yeager, C. Wooten, K. Summers, A new technique for the production of large numbers of clay models for field studies of predation. *Herpetol. Rev.* **42**, 357–359 (2011).
68. R. H. Hegna, R. A. Saporito, M. A. Donnelly, Not all colors are equal: Predation and color polytypism in the aposematic poison frog *Oophaga pumilio*. *Evol. Ecol.* **27**, 831–845 (2013).
69. V. Llaurens, M. Joron, M. Théry, Cryptic differences in colour among Müllerian mimics: How can the visual capacities of predators and prey shape the evolution of wing colours? *J. Evol. Biol.* **27**, 531–540 (2014).
70. M. J. Henze, O. Lind, J. Mappes, B. Rojas, A. Kelber, An aposematic colour-polymorphic moth seen through the eyes of conspecifics and predators—Sensitivity and colour discrimination in a tiger moth. *Funct. Ecol.* **32**, 1797–1809 (2018).
71. J. K. Valkonen, J. Mappes, Comments on Guimarães & Sawaya. Pretending to be venomous: Is a snake's head shape a trustworthy signal to a predator? *J. Trop. Ecol.* **28**, 123–124 (2012).
72. D. C. Rößler, H. Pröhl, S. Lötters, The future of clay model studies. *BMC Zoology* **3**, 6 (2018).
73. M. Ringler, E. Ursprung, W. Höld, Predation on *Allobates femoralis* (Boulenger 1884; Anura: Aromobatidae) by the colubrid snake *Xenopholis scalaris* (Wucherer, 1861). *Herpetol. Notes* **3**, 301–304 (2010).
74. D. R. Lenger, J. K. Berkey, M. B. Dugas, Predation on the toxic *Oophaga pumilio* (Anura: Dendrobatidae) by *Rhadinaea decorate* (Squamata: Collubridae). *Herpetol. Notes* **7**, 83–84 (2014).
75. D. Osorio, M. Vorobyev, C. D. Jones, Colour vision of domestic chicks. *J. Exp. Biol.* **202**, 2951–2959 (1999).
76. G. Gamberale-Stille, B. Kazemi, A. Balogh, O. Leimar, Biased generalization of salient traits drives the evolution of warning signals. *Proc. R Soc. Lond. B* **285**, 20180283 (2018).
77. G. Gamberale-Stille, B. S. Tullberg, Experienced chicks show biased avoidance of stronger signals: An experiment with natural colour variation in live aposematic prey. *Evol. Ecol.* **13**, 579–589 (1999).
78. J. P. Lawrence, B. P. Noonan, Avian learning favors colorful, not bright, signals. *PLoS One* **13**, e0194279 (2018).
79. M. Aronsson, G. Gamberale-Stille, Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Anim. Behav.* **75**, 417–423 (2008).
80. L. Lindström, R. V. Alatalo, J. Mappes, M. Riipi, L. Vertainen, Can aposematic signals evolve by gradual change? *Nature* **397**, 249 (1999).
81. K. Rönkä, C. De Pasqual, J. Mappes, S. Gordon, B. Rojas, Colour alone matters: No predator generalization among morphs of an aposematic moth. *Anim. Behav.* **135**, 153–163 (2018).
82. A. Exnerová *et al*., Different reactions to aposematic prey in 2 geographically distant populations of great tits. *Behav. Ecol.* **26**, 1361–1370 (2015).
83. R Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2017), Version 3.6.1.
84. D. Bates *et al*., Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
85. T. M. Therneau coxme: Mixed Effects Cox Models (*R Package Version*:2–2, 2015). https://cran.r-project.org/web/packages/coxme/vignettes/coxme.pdf. Accessed 21 August 2019.
86. J. W. Daly, T. F. Spande, H. M. Garraffo, Alkaloids from amphibian skin: A tabulation of over eight-hundred compounds. *J. Nat. Prod.* **68**, 1556–1575 (2005).
87. B. Rojas *et al*., How to fight multiple enemies: Target-specific chemical defences in an aposematic moth. *Proc. R Soc. Lond. B* **284**, 20171424 (2017).
88. E. Burdfield-Steel, M. Brain, B. Rojas, J. Mappes, The price of safety: Food deprivation in early life influences the efficacy of chemical defence in an aposematic moth. *Oikos* **17**, 293 (2018).
89. T. J. Roper, N. M. Marples, Odour and colour as cues for taste-avoidance learning in domestic chicks. *Anim. Behav.* **53**, 1241–1250 (1997).
90. RStudio Team, RStudio: Integrated Development for R, Version 1.1.463 (2015). https://www.rstudio.com. Accessed 21 August 2019.
91. E. A. Hoffman, F. W. Schueler, A. G. Jones, M. S. Blouin, An analysis of selection on a colour polymorphism in the northern leopard frog. *Mol. Ecol.* **15**, 2627–2641 (2006).
92. R. A. Sánchez-Guillén, B. Hansson, M. Wellenreuther, E. I. Svensson, A. Cordero-Rivera, The influence of stochastic and selective forces in the population divergence of female colour polymorphism in damselflies of the genus *Ischnura*. *Heredity* **107**, 513–522 (2011).
93. C. R. Linnen *et al*., Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312–1316 (2013).
94. A. D. Ham, E. Ihalainen, L. Lindström, J. Mappes, Does colour matter? The importance of colour in avoidance learning, memorability and generalisation. *Behav. Ecol. Sociobiol.* **60**, 482–491 (2006).
95. G. Gamberale, B. S. Tullberg, Evidence for a peak-shift in predator generalization among aposematic prey. *Proc. Biol. Sci.* **263**, 1329–1334 (1996).
96. T. Stankowich, T. Caro, M. Cox, Bold coloration and the evolution of aposematism in terrestrial carnivores. *Evolution* **65**, 3090–3099 (2011).