Warning signals are under positive frequency- dependent selection in nature

Mathieu Chouteau, Mónica Arias, Mathieu Joron

Warning signals are under positive frequency-dependent selection in nature

Mathieu Chouteau[a,b,1], Mónica Arias[b], and Mathieu Joron[a]

[a]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS - Université de Montpellier - EPHE - Université Paul Valéry, 34293 Montpellier 5, France; and [b]Institut de Systématique, Evolution, Biodiversité, UMR 7205 CNRS MNHN UPMC EPHE, Muséum National d'Histoire Naturelle, 75005 Paris, France

Edited by Hanna Kokko, University of Zurich, Zurich, Switzerland, and accepted by the Editorial Board January 19, 2016 (received for review September 28, 2015)

Positive frequency-dependent selection (FDS) is a selection regime where the fitness of a phenotype increases with its frequency, and it is thought to underlie important adaptive strategies resting on signaling and communication. However, whether and how positive FDS truly operates in nature remains unknown, which hampers our understanding of signal diversity. Here, we test for positive FDS operating on the warning color patterns of chemically defended butterflies forming multiple coexisting mimicry assemblages in the Amazon. Using malleable prey models placed in localities showing differences in the relative frequencies of warningly colored prey, we demonstrate that the efficiency of a warning signal increases steadily with its local frequency in the natural community, up to a threshold where protection stabilizes. The shape of this relationship is consistent with the direct effect of the local abundance of each warning signal on the corresponding avoidance knowledge of the local predator community. This relationship, which differs from purifying selection acting on each mimetic pattern, indicates that predator knowledge, integrated over the entire community, is saturated only for the most common warning signals. In contrast, among the well-established warning signals present in local prey assemblages, most are incompletely known to local predators and enjoy incomplete protection. This incomplete predator knowledge should generate strong benefits to life history traits that enhance warning efficiency by increasing the effective frequency of prey visible to predators. Strategies such as gregariousness or niche convergence between comimics may therefore readily evolve through their effects on predator knowledge and warning efficiency.

Müllerian mimicry | aposematism | warning signal | predation | butterflies

Frequency-dependent selection (FDS) occurs when the fitness of a phenotype depends on its frequency in the population (1). Under negative FDS, the fitness of a phenotype will decrease as its frequency increases. This supposedly common mechanism (2) maintains adaptive polymorphism and has been documented in a variety of natural systems, ranging from foraging behavior (2, 3) and pollination syndromes (4) to predator–prey (5–7) and parasite–host coevolution (8). On the other hand, positive FDS is a selection regime where the fitness of a phenotype increases with its frequency and does not readily maintain polymorphism. However, positive FDS is thought to be a mechanism central to the evolution of a large diversity of adaptive strategies. Notably, traits used in signaling and communication, where efficiency depends on local frequency, such as languages and social signals (9), flower coloration for pollinator attraction (10), and warning signals of prey unpalatability (11), should be subjected to positive FDS. However, although the principles of positive FDS may be well understood from a theoretical point of view, the extent to which it is operating in nature remains largely unknown.

Warning coloration advertising prey defenses is a textbook example of a trait under positive FDS, and is thought to be responsible for the remarkable convergence among defended prey species known as Müllerian mimicry (11, 12). The theory developed by Müller (11) posits that naive predators learn to avoid warningly colored prey after a given number of attacks, conferring improved survival to individuals bearing resemblance to a locally abundant signal. The efficiency of Müllerian mimicry is well-supported by empirical evidence. Experiments in the laboratory (13, 14) and in natural contexts (15–18) have demonstrated that naive predators have the cognitive capabilities to associate conspicuous signals with toxicity, and to avoid them later. Transplant experiments with live prey or artificial prey models performed in the field have shown increased survival for prey matching the locally abundant warning signal and lower fitness for prey with exotic or novel signals, explaining the stable geographic mosaic of locally uniform warning signals observed (16, 17, 19–21).

However, selection favoring mimicry in those studies takes on the pattern of purifying selection for local signals (15, 17, 20), so the role of positive FDS in the evolution of mimicry is unclear. Indeed, distinguishing the effect of positive FDS from local purifying selection is particularly challenging because the polymorphism necessary to distinguish between them is intrinsically unstable when the most common forms are favored (15–17, 20, 22). Both selection regimes lead to population monomorphism and interspecific phenotypic convergence, but they differ importantly in the role of local frequency in defining fitness optima. Purifying selection is expected to favor the highest quality warning signal independent of frequency, whereas FDS stems from the dependence of predator knowledge on the local frequency of warning signals across species in the prey community. Therefore, understanding how selection for mimicry truly operates and shapes prey communities with multiple coexisting mimicry rings requires confirming positive FDS and distinguishing the effect of frequency

Significance

Warning coloration and defensive mimicry are textbook examples of antipredator strategies under frequency-dependent selection (FDS), where predators avoid defended prey that are common but attack rarer ones. Whether the efficiency of mimicry truly derives from positive FDS remains unverified, yet assessing FDS is essential for understanding the diversity of warning signals observed. Using wing patterns found in natural communities of toxic butterflies, we demonstrate that warning signals are under strong, nonlinear, positive FDS. Predator knowledge is saturated and protection is optimal only for a few superabundant prey types. Most warning signals fall below optimal protection, limited by their local abundance. This potential for improved protection may explain the extraordinary efficiency of selection to drive convergence in morphology and life history in mimetic prey.

Author contributions: M.C. and M.J. designed research; M.C. and M.A. performed research; M.C. analyzed data; and M.C. and M.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. H.K. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: mathieu.chouteau@cefe.cnrs.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1519216113/-/DCSupplemental.

itself from the intrinsic qualities of the different warning signals involved.

Results

To test for positive FDS, we took advantage of the natural phenotypic diversity found within butterfly communities in the Peruvian Amazon basin, where the warning colors displayed by butterfly species are structured into multiple coexisting mimicry assemblages. The Müllerian mimicry communities formed by butterflies with so-called "tiger" wing patterns [e.g., *Heliconius numata*, *Melinaea* butterflies (23, 24)] provide multicomponent polymorphism with different warning signals coexisting at different frequencies in different localities. Each locality harbors a distinct combination of high-, intermediate-, and low-frequency signals (25) (Fig. 1), and thus enables us to investigate the effect of signal frequency on predation.

We quantified attack rates associated with five major tiger-patterned phenotypes in adjacent localities of northern Peru (Fig. 1 and Fig. S1). We positioned artificial models built of malleable wax and photographic paper closely matching the local phenotypes along forest transects for 72 h (Fig. S2 and Table S1). Attacks were recorded as imprints left by predators (mostly avian predators, representing 69.9% of total attacks; Fig. S2), providing a measure of predation intensity on each wing pattern, and therefore a snapshot of predator avoidance knowledge for each locality (17, 19). Prey models bearing the different color patterns found in the locality, as well as exotic and control prey types, were placed in each locality, totaling 120 models per color pattern. To record predation across the natural morph frequency spectrum, we used models matching the distinct local wing patterns (two or more morphs found at different frequencies among local butterflies). To record predation on novel conspicuous prey unknown to local predators but known as an efficient warning signal elsewhere, we used models for one or two exotic phenotypes naturally absent from the locality but abundant in an adjacent locality (except in

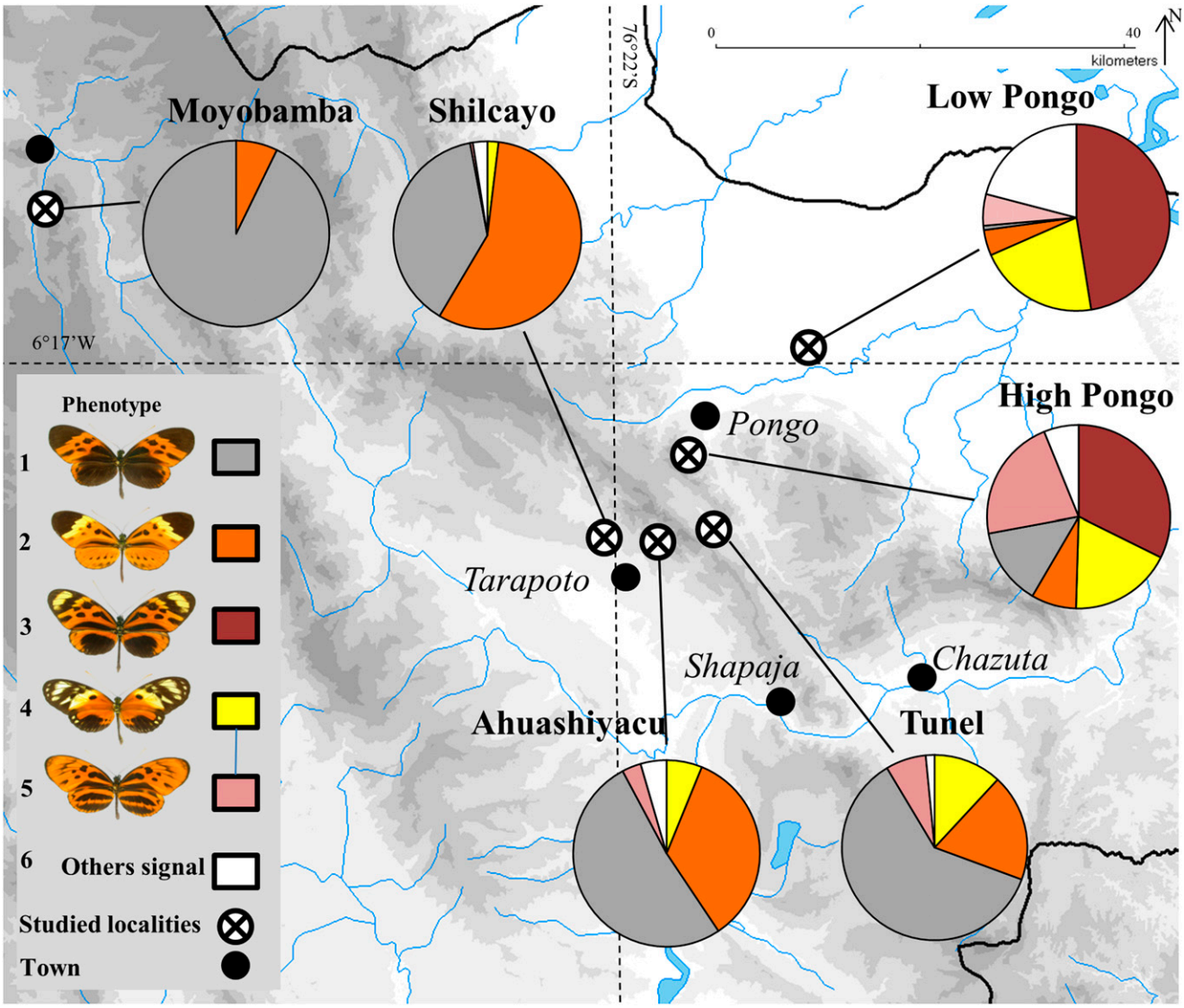

Fig. 1. Relative frequency of warning signals displayed by tiger-patterned butterflies in the study localities in northern Peru. Temporally stable relative frequencies were obtained from exhaustive field collecting (n = 1,873 butterflies) performed between 1997 and 2013 for the warning signals displayed by comimics (Fig. S1), such as (1) *Melinaea marsaeus mothone*, *Heliconius numata bicoloratus*, and *Chetone hydra*; (2) *Melinaea menophilus* ssp. nov.1, *Heliconius numata tarapotensis*, *Heliconius pardalinus sergestus*, and *Chetone histrio*; (3) *Melinaea marsaeus rileyi* and *Heliconius numata aurora*; (4) *Melinaea ludovica ludovica* and *Heliconius numata silvana*; (5) *Melinaea marsaeus phasiana* and *Heliconius numata arcuella*; and (6) other large sized tiger pattern signals such as those found in *Melinaea menophilus hicetas*, *Melinaea satevis cydon*, *Heliconius numata elegans*, or *Heliconius numata timaeus*.

High Pongo, where all warning signals co-occur). Finally, as a general predation control, we used models resembling a common palatable brown butterfly (i.e., *Pierella hyceta*). Predation rate on the palatable butterfly models revealed some variation in predation intensity between localities ($\chi^2_5 \geq 12.07$, $P = 0.033$), with high predation intensity in Moyobamba, Ahuashiyacu, and Low Pongo and lower predation intensity in the remaining localities.

In all localities, predation attempts decreased steeply but steadily as the local frequency of the warning signal increased from exotic to dominant ($R^2 \leq 0.876$, $P \leq 0.031$ in the three localities with high predation intensity and $R^2 \leq 0.652$, $P \leq 0.201$ in the three others). The most frequently encountered warning signal enjoyed the highest protection from predators ($P \leq 0.05$ in the three localities with high predation intensity and $P \leq 0.10$ in the three others), whereas exotic signals suffered up to nine times more attacks (average of 4.7 ± 2.6-fold more attacks). All local signals found at intermediate frequencies suffered intermediate predation attempts (Fig. 2*A* and Table S2). Because a given morph was strongly protected when locally abundant, but increasingly attacked when frequency decreased only a few kilometers away,

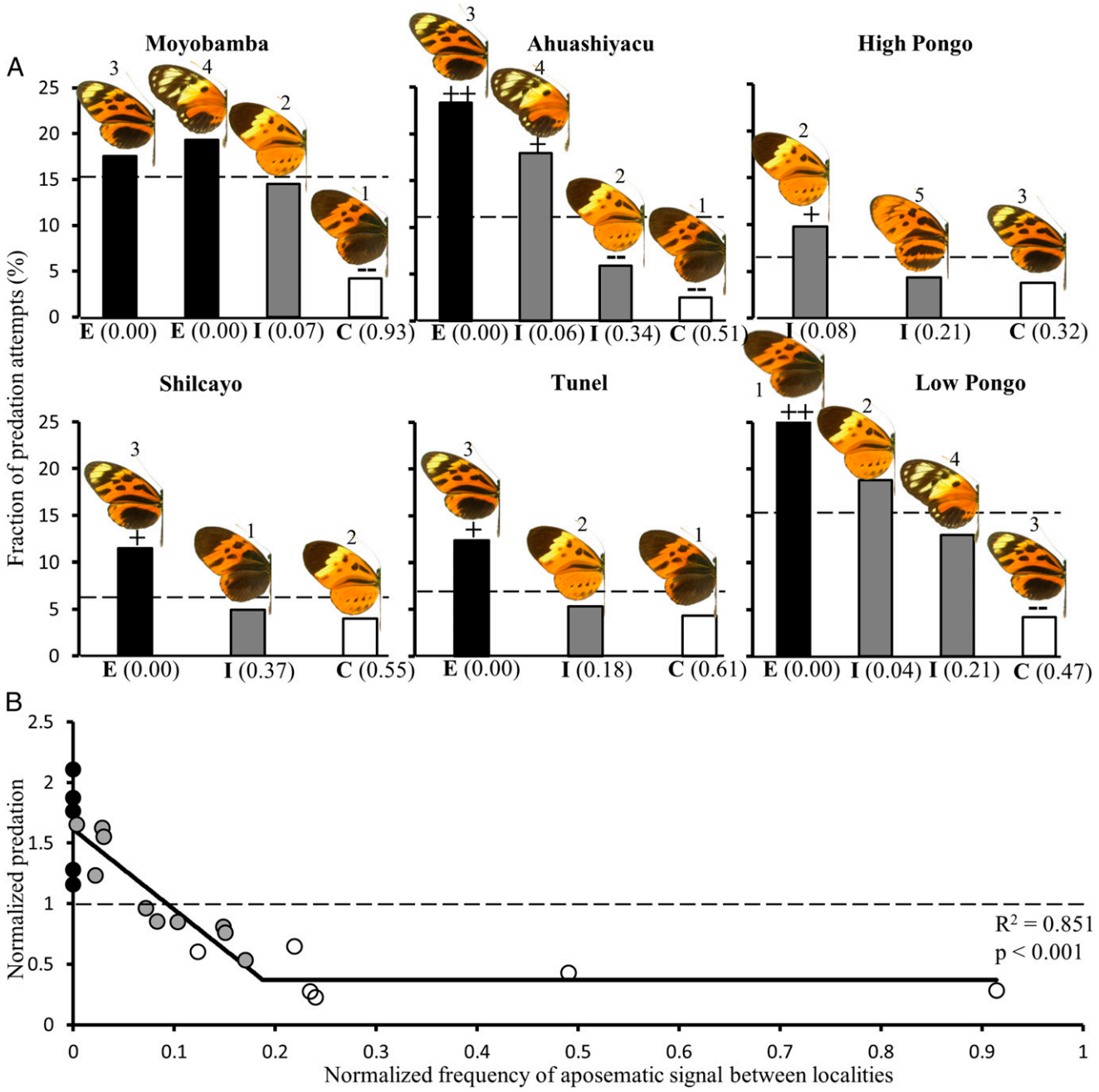

Fig. 2. Predator avoidance knowledge as a function of warning signal frequency. (*A*) For each of the six localities, phenotypes (numbered as in Fig. 1) are ordered from left to right according to their relative frequencies within the locality (E = exotic, I = intermediate abundance, C = most common, with the frequency indicated in brackets). Within a given locality, FTs identify which model phenotype suffered significantly more (++$P < 0.05$; +$P < 0.10$) or fewer (−−$P < 0.05$; −$P < 0.10$) predation attempts. (*B*) Entire study system with normalization of phenotype frequency and predation intensities between localities. The color of data points identifies the exotic (black), intermediate (gray), and most common (white) warning signals within each locality. The dotted line indicates the predation intensity on the "palatable" butterfly model.

our results indicate that local morph frequency strongly determined predation independent of the intrinsic warning qualities of each phenotype involved. For instance, the wing pattern shared by *Heliconius numata tarapotensis* and *Melinaea menophilus* ssp. nov. (phenotype 2 in Fig. 1) is highly protected in Shilcayo and Ahuashyacu, where it is abundant, but strongly attacked only 10 km away in High Pongo and Low Pongo, where it is rare. However, variations in species assembly among localities [according to microhabitat, vegetation structure, host plant, etc. (26, 27)] may generate variations in toxicity levels for certain mimicry groups, which might, in theory, increase predator avoidance on the most abundant groups if they tend to be the most toxic ones. However, because a pattern displayed by the same set of species in similar proportions can suffer drastically different levels of predation (e.g., *H. n. tarapotensis* and *M. menophilus* ssp. nov. encountered in a 4:3 ratio in both Shilcayo and High Pongo), we conclude that variations in toxicity are unlikely to confound the effect of variations in frequency observed in our results.

Because analyzing localities independently does not enable the assessment of frequency dependence spanning the entire frequency spectrum, we pooled the results from all six localities. To do so, we normalized both the frequencies and the predation intensity so as to correct for between-locality variations in the abundance of prey available for a given predation intensity (*Materials and Methods*). In contrast to the constant attack intensity expected under Müller's theory, a strongly nonlinear relationship was found between signal frequency and predation intensity across the normalized frequency spectrum (Akaike information criterion weight of 0.00 for the flat model vs. 0.71 for a linear decrease plateau model). Predation pressure may be divided into two main regimes of frequency dependence ($R^2 = 0.851$, $P < 0.001$ according to the best fit model; Fig. 2*B*). For low (exotic) to intermediate signal frequency in the environment, predation steadily decreased from high intensity to minimal intensity. Then, for morph frequencies above a certain threshold of natural occurrence, frequency dependence appears to abate and predation intensity remains nonzero with increasing signal frequency. This baseline for predation intensity, associated with locally abundant phenotypes, was similar for all localities (0.372 ± 0.171 normalized predation attempts) and reached a threshold of normalized frequency of 0.188 (best fit model; Fig. 2*B*).

Discussion

Our results, which show that the local frequency of a warning signal is a major determinant of its efficiency, demonstrate that positive FDS truly operates on local color patterns and support the theory proposed by Fritz Müller (11) over 140 y ago. However, FDS acting on warning signals is traditionally considered a dilution effect through number dependence: A learning community of predators will attack a certain number of prey per unit time [Müller's n_k (11, 12, 28)] until they associate a signal with noxiousness. Although a simplification of the real world (12), this model rests on the assumption that different warning signals suffer a similar number of predation events independent of their abundance, translating into a decreasing per capita mortality as more prey adopt the warning signal (28). Our data show that this prediction only holds for warning signals above an elevated frequency threshold (i.e., for very common prey). In contrast, below this threshold, predation risk is intimately related to the warning signal's local frequency. Although high predation on exotic variants is expected because local predators are naive to these signals (17), the range of intermediate predation intensities reflecting intermediate predator knowledge for signals occurring at intermediate frequencies is not predicted by individual predator learning. Neither current theories, which assume that predators should attack at a steady rate until knowledge is acquired, nor empirical studies, which show fast associative avoidance learning of warningly colored prey by naive individual predators (11, 18, 29–32), can account for the gradual decrease in predation intensity we observed. However, our study provides a measurement for the behavior and learning state of an entire community of predators. Common signals are often encountered and known to all predators, whereas rarer ones will be known to only a fraction of the predators, which, in turn, depends on local signal frequency. Because variations in predation rates in our study reflect the density of local predators not educated for a given warning signal, predation on rare or intermediately frequent signals reveals incomplete learning at the scale of the predator community (33) (Fig. S3). For these ranges of signal frequencies, an increase in signal frequency translates into improved predator knowledge. The incomplete predator knowledge therefore leaves some scope for improvement, which may explain the evolution of traits that maximize encounter rates of warning signals by predators, such as mimetic convergence (28), spatial aggregation (12), and the convergence of an ecological niche between comimics (26).

Above a certain frequency threshold (≥0.188 normalized abundance here; Fig. 2*B*), attacks stabilize at a frequency-independent level, suggesting the learning saturation of the local predator community. At this range of frequencies, warning signal efficiency is optimal and Müllerian number-dependent dilution acts (34). However, relatively few of the local signals studied here surpass this threshold. The saturation of predator knowledge for the most abundant signals provides support for the hypothesis that selection for convergence should be relatively inefficient between the most abundant warning signals (34), and may explain the coexistence of multiple major mimicry rings within each locality, attracting species from rarer mimicry rings through mimetic advergence (35).

FDS stems from the interaction of predator–prey encounters and learning; therefore, it is expected to act on any aposematic or Müllerian mimicry system for which predator avoidance is a learned response. However, whether the high prevalence of well-established, but suboptimally protected, warning signals found here in butterflies is also common in other types of warningly colored organisms is unknown, and may depend on the ecology of predator–prey encounters. Here, the prevalence of signals bringing incomplete protection suggests that strong selection for warning efficiency is not limited to transitional zones, new colonization events, or the emergence of new signals, but is operating continuously on most local warning signals. A prey community is therefore not defined primarily by locally abundant patterns vs. rare exotic ones but, rather, is composed of a rich diversity of differentially protected warning signals, subject to differential selection based on their frequencies, which vary spatially, and operating on traits that affect local exposure to predators.

Given the clear frequency-related benefits owing to positive FDS, mimicry benefits may change rapidly with geography, suggesting the possibility of rapid mimicry switches and a more dynamic situation than envisioned before. Simultaneously, local selection on warning signals may be expected to facilitate the evolution and gradual refinement of traits that maximize local warning signal visibility and density, and enhance predator knowledge (26, 27). Characterizing positive FDS in this system is an important step toward understanding the evolution of signaling in a natural and diverse ecosystem. Similar effects on the learning dynamics of a community of receivers in response to signalers may also apply in the evolution of languages or other communication traits.

Materials and Methods

Warning Signal Abundance. Relative frequency of the different mimetic phenotypes in each region (Moyobamba: 6°04′34″S, 7°57′27″W, altitude of 1,130 m; Shilcayo: 6°27′30″S, 76°21′00″W, altitude of 460 m; Ahuashyacu: 6°27′07″S, 76°18′44″W, altitude of 1,020 m; Tunel: 6°27′11″S, 76°17′11″W, altitude of 1,090 m; High Pongo: 6°21′19″S, 76°18′25″W, altitude of 595 m; Low Pongo: 6°17′53″S, 76°14′38″W, altitude of 200 m) was assessed from several intensive collecting expeditions between 1997 and 2013. These well-sampled localities are known to display temporally stable warning signal polymorphism (Fig. S1). In each locality, the relative frequency of each warning phenotype (*Heliconius numata*, *Heliconius pardalinus*, *Melinaea* sp.,

and *Chetone* sp. comimics) was calculated yearly for collections occurring between 1997 and 2013, and averaged between years for data analysis.

To compare warning signal abundance between localities, the average number of individuals bearing the signals of interest ($Abun_{Loc}$) collected over 8–10 sunny days, always by the same two collectors, throughout 2013 was calculated for each locality. The frequency of each warning signal per locality ($Freq_{Loc}$) was then normalized to between-locality frequency ($Freq_{Norm}$) using the previously assessed abundance ($Abun_{Loc}$). For instance, the normalized frequency of signal A found in locality 1 is defined as $Freq_{Norm\ A1} = Freq_{Loc\ A1} \times Abun_{Loc\ 1}/Abun_{Loc\ Max}$, where the $Abun_{Loc\ Max}$ is the highest abundance coefficient obtained (i.e., Moyobamba). The abundance obtained from collecting efforts for each locality ($Abun_{Loc}$) is as follows: Moyobamba = 14.1, Shilcayo = 5.6, Ahuashiyacu = 6.9, Tunel = 11.4, High Pongo = 5.4, and Low Pongo = 7.

Predation Tests. To assess predation rates, artificial models were built to resemble the local morphs of *H. numata* (mimetic with species of the genus *Melinaea, Chetone* sp., and other *Heliconius* sp.) and a model of a palatable brown butterfly, *Pierella hyceta* (Fig. S2). These artificial models were made by printing standardized high-resolution photographs of the ventral and dorsal butterfly wings on two-sided matte photographic paper (Epson C13S041569 paper and Epson L110 printer). Colors of printed wings were compared with colors of actual *H. numata* wings by measuring the reflectance spectra of yellow, black, and orange patches using a spectrophotometer (AvaSpec-3648; Avantes) and a deuterium-halogen light source (DH-2000; Avantes) connected to a 1.5-mm diameter sensor (FCR-7UV200-2-1.5x100; Avantes) inserted in a miniature black chamber. Reflectance spectra were taken at a 90° incidence relative to a 99% reflectance standard (300–700 nm; Spectralon) and to dark current. Spectra were recorded with Avasoft 7.0 software (Avantes) using an average of five measures with an integration time of 23 ms. Color spectra were then analyzed applying the method described by Vorobyev and Osorio (36) performed in Avicol v.6 (37). We contrasted black, orange, and yellow (when naturally present in the wings) on printed vs. actual wings, under two main avian vision systems. We used blue tit (*Parus caeruleus)* for UV vision, with cone proportion and sensitivity as described by Hart et al. (38), and peafowl (*Pavo cristatus*) (39) for violet (V) vision, as described by Hart (40). Photoreceptor activity is characterized by the signal-to-noise ratio, which is generally computed from its inverse, the Weber fraction (36), and which was set to 0.05 for all models in this study. Small gap light conditions from French Guiana (41) were included for all calculations. The difference in contrast was below the just noticeable threshold for both the UV-sensitive and V-sensitive avian visual systems, confirming the high color accuracy of our printed wings (Table S1). Model wings were assembled to a small galvanized wire, which provided both reinforcement for the body and a 20-cm-long anchor to facilitate positioning of the models in the habitats. Bodies were made of black-colored paraffin wax heated and molded, using a custom-made silicon mold, over the thorax portions of the wings and wire. Before positioning the models in the habitats, a 4% (vol/vol) permethrin solution was sprayed on the models to reduce destruction by invertebrates. In each locality, between 480 and 600 models (i.e., 120 models for each phenotype, four to five phenotypes per habitat) were placed following the same phenotype sequence over a 6- to 8-km transect (Table S2). Models were positioned at least at 10-m intervals to provide a uniform distribution of artificial warning signals and anchored on twigs, tree trunks, lianas, or leaf petioles, such that the models appeared to be perched on a tiny twig. Doing so enabled the models to be readily detectable to avian predators, as sun-basking butterflies would be, while preventing ground-dwelling nonpredatory animals (roaches, ants, and rodents) from damaging the models.

In all localities, models were left in the habitat for 72 h before collection. Predation marks on the malleable wax were categorized as "avian" when they showed a distinctive U- or V-shaped mark or large indentation or puncture marks (17, 19) and as "unknown vertebrate predator" when showing no such distinctive marks on the damaged area, or when the attacked parts of the model had been removed or torn by the predator (17). Models not recovered, or destroyed by ants or crickets, were scored as missing and excluded from the analysis (n = 229 models).

In an effort to control for variations in predation intensity between localities so as to analyze all localities combined, predation attempts on the different warningly colored models of a given locality were divided by the attack rate on the palatable butterfly model in that same locality.

Data Analysis. We used a χ^2 test of independence to assess variation in predation pressure between the different artificial models for each locality. When significant, the Freeman–Tukey deviate (FT) was compared with an alpha = 0.10 and 0.05 criterion, and corrected for multiple comparisons using the Bonferroni correction, to identify which artificial model phenotype suffered significantly more or fewer predation attempts, as shown by the sign of the FT, than expected under the null hypothesis of equal attack probability.

To describe the relationship between predation and abundance across all localities using normalized data (Fig. 2*B*), we compared four candidate models: a flat model (constant predation) as predicted by Müller (11), a linear model (linear decrease in predation), an exponential decay model (asymptotic decrease of predation), and a linear-lower plateau model (linear decrease in predation followed by a plateau at a given abundance threshold). A parameter search for both exponential decay and linear-lower plateau models was done using the solver function in Excel (Microsoft Corp.) using a least-squares optimality criterion and using the generalized reduced gradient nonlinear algorithm with default parameters. To evaluate which of the four models was a better fit to the data, we calculated Akaike information criterion corrected for small sample sizes (AICc) for each of the four models and compared them using Akaike model comparison differences (ΔAICc) and Akaike weights (42) using the residual sum of squares divided by the sample size as the likelihood criterion (Table S3). The best fit was found for the linear-lower plateau model with the equation $y = ax + b$ if $x < x_0$ and $y = y_0$ if $x \geq x_0$, where $x_0 = 0.188$ is the breakpoint for the relative abundance, $y_0 = 0.372$ is the basal predation rate, $a = -6.619$, and $b = 1.613$.

ACKNOWLEDGMENTS. We thank R. Mori-Pezo for his invaluable help with field work; D. Marina, V. Llaurens, and Y. Le Poul for the design of artificial butterflies; and J. Mallet, M. McClure, M. Elias, V. Llaurens, T. Aubier, A. Evin, and A. Whibley for their comments on manuscript drafts. We also thank the Peruvian government for providing the necessary research permits (236-2012-AG-DGFFS-DGEFFS and 201-2013-MINAGRI-DGFFS/DGEFFS). This research was supported by fellowships from the Natural Sciences and Engineering Research Council and from the European Union's Horizon 2020 research and innovation program under Marie Sklodowska-Curie Grant Agreement N 655857 (to M.C.), by a studentship provided by the Labex Biological and Cultural Diversities (LabEx ANR-10-LABX-0003-BCDiv) (to M.A.), and by grants from the European Research Council (Grant ERC-StG-243179-MimEvol) and the Agence Nationale de la Recherche (Grant ANR-12-JSV7-0005-01-Hybevol) (to M.J.).

1. Ayala FJ, Campbell CA (1978) Frequency-dependent selection. *Annu Rev Ecol Syst* 5:115–138.
2. Fitzpatrick MJ, Feder E, Rowe L, Sokolowski MB (2007) Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* 447(7141): 210–212.
3. Benkman CW (1996) Are the ratios of bill crossing morphs in crossbills the result of frequency-dependent selection? *Evol Ecol* 10(1):119–126.
4. Subramaniam B, Rausher MD (2000) Balancing selection on a floral polymorphism. *Evolution* 54(2):691–695.
5. Fitzpatrick BM, Shook K, Izally R (2009) Frequency-dependent selection by wild birds promotes polymorphism in model salamanders. *BMC Ecol* 9:12.
6. Hori M (1993) Frequency-dependent natural selection in the handedness of scale-eating cichlid fish. *Science* 260(5105):216–219.
7. Olendorf R, et al. (2006) Frequency-dependent survival in natural guppy populations. *Nature* 441(7093):633–636.
8. Conway DJ (1997) Natural selection on polymorphic malaria antigens and the search for a vaccine. *Parasitol Today* 13(1):26–29.
9. Cartwright RA (2011) Bards, poets, and cliques: frequency-dependent selection and the evolution of language genes. *Bull Math Biol* 73(9):2201–2212.
10. Epperson BK, Clegg MT (1987) Frequency-dependent variation for outcrossing rate among flower-color morphs of *Ipomoea purpurea*. *Evolution* 41(6):1302–1311.
11. Müller F (1878) Über die vortheile der mimicry bei schmetterlingen. *Zool Anz* 1:54–55. German.
12. Sherratt TN (2008) The evolution of Müllerian mimicry. *Naturwissenschaften* 95(8): 681–695.
13. Miller AM, Pawlik JR (2013) Do coral reef fish learn to avoid unpalatable prey using visual cues? *Anim Behav* 85(2):339–347.
14. Rowland HM, Ihalainen E, Lindström L, Mappes J, Speed MP (2007) Co-mimics have a mutualistic relationship despite unequal defences. *Nature* 448(7149):64–67.
15. Kapan DD (2001) Three-butterfly system provides a field test of müllerian mimicry. *Nature* 409(6818):338–340.
16. Mallet J, Barton NH (1989) Strong natural-selection in a warning-color hybrid zone. *Evolution* 43(2):421–431.
17. Chouteau M, Angers B (2011) The role of predators in maintaining the geographic organization of aposematic signals. *Am Nat* 178(6):810–817.
18. Langham GM (2004) Specialized avian predators repeatedly attack novel color morphs of Heliconius butterflies. *Evolution* 58(12):2783–2787.
19. Noonan BP, Comeault AA (2009) The role of predator selection on polymorphic aposematic poison frogs. *Biol Lett* 5(1):51–54.
20. Borer M, Van Noort T, Rahier M, Naisbit RE (2010) Positive frequency-dependent selection on warning color in Alpine leaf beetles. *Evolution* 64(12):3629–3633.
21. Chouteau M, Angers B (2012) Wright's shifting balance theory and the diversification of aposematic signals. *PLoS One* 7(3):e34028.
22. Comeault AA, Noonan BP (2011) Spatial variation in the fitness of divergent aposematic phenotypes of the poison frog, Dendrobates tinctorius. *J Evol Biol* 24(6):1374–1379.

23. Llaurens V, Joron M, Théry M (2014) 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(3):531–540.
24. Joron M (2005) Polymorphic mimicry, microhabitat use, and sex-specific behaviour. *J Evol Biol* 18(3):547–556.
25. Joron M, Wynne IR, Lamas G, Mallet J (1999) Variable selection and the coexistence of multiple mimetic forms of the butterfly Heliconius numata. *Evol Ecol* 13(7-8):721–754.
26. Elias M, Gompert Z, Jiggins C, Willmott K (2008) Mutualistic interactions drive ecological niche convergence in a diverse butterfly community. *PLoS Biol* 6(12):2642–2649.
27. Beccaloni GW (1997) Vertical stratification of ithomiine butterfly (Nymphalidae: Ithomiinae) mimicry complexes: The relationship between adult flight height and larval host-plant height. *Biol J Linn Soc Lond* 62(3):313–341.
28. Mallet J (1999) Causes and consequences of a lack of coevolution in mullerian mimicry. *Evol Ecol* 13(7-8):777–806.
29. Brower LP, Ryerson WN, Coppinger LL, Glazier SC (1968) Ecological chemistry and the palatability spectrum. *Science* 161(3848):1349–1350.
30. Skelhorn J, Rowe C (2007) Predators' toxin burdens influence their strategic decisions to eat toxic prey. *Curr Biol* 17(17):1479–1483.
31. Speed MP (2001) Can receiver psychology explain the evolution of aposematism? *Anim Behav* 61(1):205–216.
32. Speed MP (1999) Batesian, quasi-Batesian or Mullerian mimicry? Theory and data in mimicry Research. *Evol Ecol* 13(7-8):755–776.
33. Mallet J (2001) Mimicry: An interface between psychology and evolution. *Proc Natl Acad Sci USA* 98(16):8928–8930.
34. Mallet J, Joron M (1999) Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annu Rev Ecol Syst* 30(7):201–233.
35. Chouteau M, Summers K, Morales V, Angers B (2011) Advergence in Müllerian mimicry: The case of the poison dart frogs of Northern Peru revisited. *Biol Lett* 7(5):796–800.
36. Vorobyev M, Osorio D (1998) Receptor noise as a determinant of colour thresholds. *Proc R Soc Lond B Biol Sci* 265(1394):351–358.
37. Gomez D (2006) AVICOL, a program to analyse spectrometric data. Available at https://sites.google.com/site/avicolprogram/. Accessed December 11, 2013.
38. Hart NS, Partridge JC, Cuthill IC, Bennett ATD (2000) Visual pigments, oil droplets, ocular media and cone photoreceptor distribution in two species of passerine bird: The blue tit (Parus caeruleus L.) and the blackbird (Turdus merula L.). *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 186(4):375–387.
39. Hart NS (2002) Vision in the peafowl (Aves: Pavo cristatus). *J Exp Biol* 205(Pt 24): 3925–3935.
40. Hart NS (2001) Variations in cone photoreceptor abundance and the visual ecology of birds. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 187(9):685–697.
41. Thery M, Pincebourde S, Feer F (2008) Dusk light environment optimizes visual perception of conspecifics in a crepuscular horned beetle. *Behav Ecol* 19(3):627–634.
42. Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York).

Supporting Information

Chouteau et al. 10.1073/pnas.1519216113

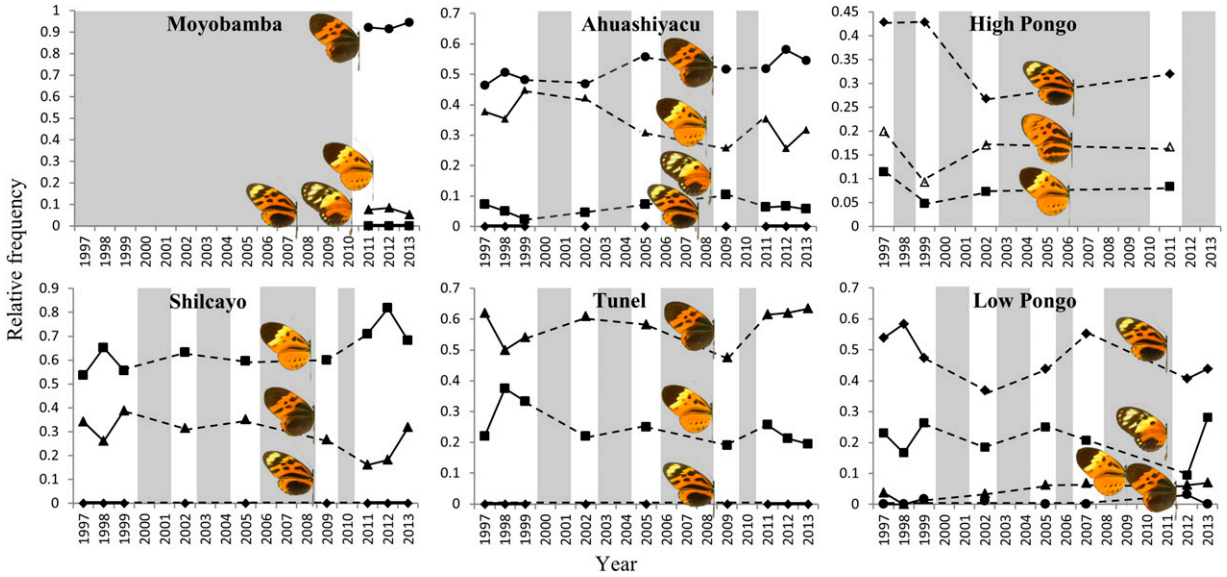

Fig. S1. Yearly relative frequency of each studied aposematic phenotype (i.e., *Heliconius numata*, *Heliconius pardalinus*, *Melineae sp.*, and *Chetone sp.* comimics) for each locality. Gray areas represent years for which no abundance data are available.

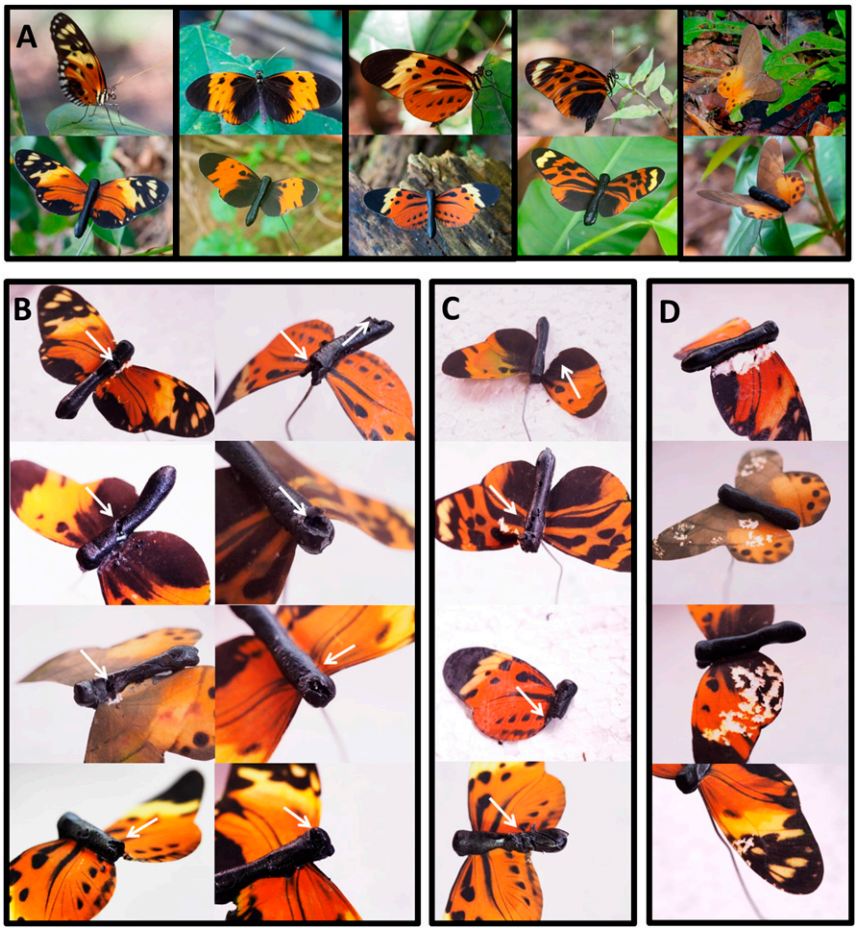

Fig. S2. (*A*) Four aposematic phenotypes of *Heliconius numata (H. n.)* (*Left* to *Right*: *H. n. silvana*, *H. n. bicoloratus*, *H. n. tarapotensis,* and *H. n. aurora*) and a nonconspicuous palatable butterfly, *Pierella hyceta* (*Top*), and their respective artificial models (*Bottom*) used to measure predation rates. Distinctive marks on the artificial models (white arrows) resulting from attacks by avian predators (*B*), unknown predators (*C*), and invertebrates (*D*) are shown.

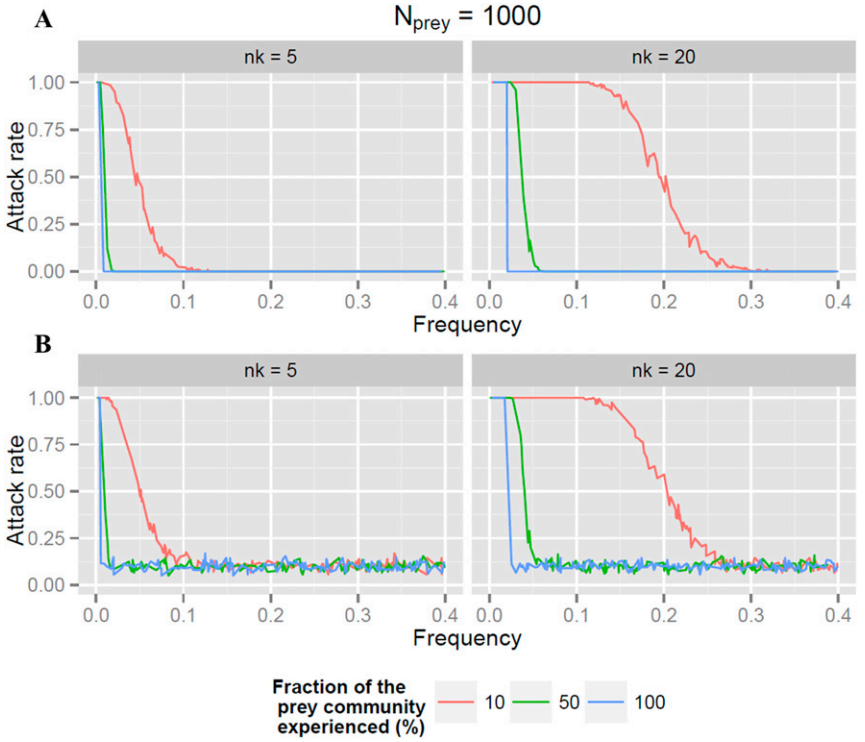

Fig. S3. Simulations of attack rates on a given prey type when predator education is limited by the predator encounter rate with local prey. To estimate the attack rate of predators on a given prey type (*i*) at a given time, we simulated the learning state of a set of N_{pred} predators at a given time, based on the assumption that each predator has gained experience from previous encounters with only a fraction f_{prey} of the entire prey community (N_{prey}). We randomly drew N_{prey}·f_{prey} prey from a community where prey type *i* is represented with a frequency p_i. For each draw (representing one individual predator), the number of prey n_i bearing a warning signal *i* and seen by the predator was compared with n_k, the number of trials required for predators to stop attacking a signal. If $n_i \geq n_k$, the predator was considered educated for prey type *i* [i.e., it had seen (and attacked) enough of this prey type to develop aversive learning, and would not attack it at the next opportunity]. Otherwise, the predator was considered uneducated (i.e., the predator had not had sufficient experience with this prey type, and would attack it at the next opportunity). In our experiment, the attack probability on a given prey model corresponds to the fraction of predators uneducated for the model's signal at the time of the experiment. Attack rate is therefore given by the iteration of this draw across all local predators. Graphs show the attack rate, at the time of an encounter with a prey (experiment), as a function of p_i for different values of f_{prey} and n_k, in two situations: when all predators have some prior knowledge (*A*) and when a fraction of the predator community (10%) is fully naive (*B*) at the time of the experiment. The simulated number of predators was of 200 (N_{pred} = 200). When predators are aware of the entire prey community (f_{prey} = 100%, blue line), the attack rate of the community of predators switches abruptly from 1 to 0 (or to a level determined by the fraction of naive predators) for all warning signal frequencies p_i·N_{prey} > *nk*. Otherwise, predator education is limited by their encounter rate with local prey (f_{prey} < 100%, green and red lines), and the attack rate of the predator community decreases gradually with increasing warning signal frequency, reflecting the increasing proportion of educated predators in the community.

Table S1. Mean and SD of the chromatic and achromatic differences for real butterfly wings and for wings of the printed artificial models under small gap light

Avian predator vision system	UV type		V type	
	DS	DQ	DS	DQ
Black	0.346 ± 0.014	0.188 ± 0.070	0.354 ± 0.100	0.270 ± 0.119
Orange	0.446 ± 0.207	0.062 ± 0.064	0.712 ± 0.178	0.088 ± 0.084
Yellow	1.357 ± 0.147	0.437 ± 0.297	1.033 ± 0.673	0.890 ± 0.673

Mean and SD of the chromatic [Delta S (DS)] and achromatic [Delta Q (DQ)] differences in orange, black, and yellow for real butterfly wings and for wings of the printed artificial models under small gap light are shown. These differences were computed for the two main visual systems known in birds (UV type: blue tit, V type: peafowl). We selected print settings that provided measures similar to the measures obtained from real wings (DS and DQ less or not much higher than 1).

Table S2. Summary data and statistics for each studied locality

Phenotype[†]	 *aurora*	 *silvana*	 *tarapotensis*	 *bicoloratus*	 *arcuela*	 *inconspicuous*
Moyobamba						
Relative frequency[‡]	0.00 ± 0.00	0.00 ± 0.00	0.07 ± 0.01	0.93 ± 0.01	0 ± 0	N.A.
Attacked/unattacked[§]	20/94	23/96	17/100	5/113	—	18/101
FT statistic[¶]	1.25	1.36	0.10	**3.61****	—	0.26
Shilcayo						
Relative frequency[‡]	0.00 ± 0.00	0.02 ± 0.02	0.55 ± 0.17	0.37 ± 0.17	0.04 ± 0.08	N.A.
Attacked/unattacked[§]	11/96	—	4/101	5/101	—	6/98
FT statistic[¶]	**1.55***	—	−0.95	−0.52	—	−0.06
Ahuashiyacu						
Relative frequency[‡]	0.00 ± 0.00	0.06 ± 0.02	0.34 ± 0.06	0.51 ± 0.04	0.03 ± 0.02	N.A.
Attacked/unattacked[§]	28/92	21/96	7/112	3/117	—	13/105
FT statistic[¶]	**2.98****	**1.67***	**2.18****	**3.96****	—	−0.28
Tunel						
Relative frequency[‡]	0.00 ± 0.00	0.11 ± 0.06	0.18 ± 0.12	0.61 ± 0.09	0.07 ± 0.04	N.A.
Attacked/unattacked[§]	10/81	—	5/94	4/92	—	7/98
FT statistic[¶]	**1.71***	—	−0.28	−0.65	—	−0.72
High Pongo						
Relative frequency[‡]	0.32 ± 0.11	0.18 ± 0.09	0.08 ± 0.02	0.13 ± 0.08	0.22 ± 0.18	N.A.
Attacked/unattacked[§]	4/105	—	9/92	—	5/113	6/94
FT statistic[¶]	−1.09	—	**1.60***	—	−0.77	0.21
Low Pongo						
Relative frequency[‡]	0.47 ± 0.07	0.21 ± 0.05	0.04 ± 0.02	0.01 ± 0.01	0.05 ± 0.04	N.A.
Attacked/unattacked[§]	5/114	15/101	22/95	30/89	—	18/100
FT statistic[¶]	**3.90****	−0.61	0.97	**2.46****	—	0.05

N.A., relative abundance has not been measured.

[†]For simplicity, only the name of the *Heliconis numata* morph is used to describe the aposematic phenotypes.

[‡]The relative frequency ($Freq_{Loc}$) of an aposematic signal in a given locality is presented as a yearly mean and its associated SD.

[§]For the 120 artificial models used, "attacked" represent models that were recovered with marks left due to predation attempts from avian and unknown assailants; "unattacked" is the number of models without any marks, when both missing models and models destroyed by invertebrates are removed from the equation. A one-em dash (—) was used when predation was not measured for a given aposematic phenotype.

[¶]Test of FTs in individual cells of contingency tables with the number of predation attacks for each site. Absolute values larger than the criterion with Bonferroni correction for simultaneous tests are shown in bold. These values identify the cells for which the number of observations significantly (*$P < 0.10$; **$P < 0.05$) differs (is either higher or lower as shown by the sign) from the corresponding expected frequencies.

Table S3. Results for the model fit for each of four candidate models describing the relationship between predation and abundance in the study system

Model	Equation	AICc	ΔAICc	AIC weight
Flat	$y = a$ ($a = 0.97$)	−20.42	34.76	0.00
Linear	$y = a + bx$ ($a = 1.28$, $b = 1.77$)	−32.52	22.66	0.00
Exponential decay	$y = a + 1/bx$ ($a = -0.10$, $b = 1.61$)	−53.34	1.84	0.29
Linear decrease plateau	$y = a + bx$ if $x < x_0$ and $y = a + bx_0$ if $x \geq x_0$ ($a = 1.61$, $b = 6.62$, $x_0 = 0.188$)	**−55.18**	**0.00**	**0.71**

The best supported model (as indicated by AICc) is shown in bold. AICc, Akaike information criterion corrected for small sample sizes; ΔAICc, Akaike models comparison difference; AIC weight, Akaike weight.