A Müllerian mimicry ring in Appalachian millipedes

Paul E. Marek[1] and Jason E. Bond

Department of Biology, Howell Science Complex, East Carolina University, Greenville, NC 27858

Edited by May R. Berenbaum, University of Illinois at Urbana-Champaign, Urbana, IL, and approved April 10, 2009 (received for review October 17, 2008)

Few biological phenomena provide such an elegant and straightforward example of evolution by natural selection as color mimicry among unrelated organisms. By mimicking the appearance of a heavily defended aposematic species, members of a second species gain protection from predators and, potentially, enhanced fitness. Mimicking a preexisting warning advertisement is economical because a potentially costly novel one can be avoided; simultaneously, the addition of more aposematic individuals enhances the overall warning effect. The better-known mimetic systems comprise tropical taxa, but here, we show a remarkable example of color mimicry in 7 species of blind, cyanide-generating millipedes endemic to the Appalachian Mountains of temperate North America. Because these millipedes lack eyes, there is no sexual selection or intraspecific signaling for coloration, providing an ideal system for mimicry studies. We document a Müllerian symbiosis where unrelated species vary in color and pattern over geographical space but appear identical where they co-occur. By using spectral color data, estimations of evolutionary history, and detailed field observations of species abundance, we test 4 predictions of Müllerian mimicry theory and begin to unravel the story of an elaborate mimetic diversification in the forests of Appalachia.

aposematic | reflectance | Apheloriini | Diplopoda | *Brachoria*

Müllerian mimicry, a shared coloration system in which all participant species are defended (1), is the more commonly documented mimicry phenomenon. Batesian mimicry, in contrast, is a form where one species is undefended and the other is defended (2). Müllerian theory predicts mutualism between species and positive frequency dependence whereby "honest" warning signals (backed up by antipredation defense) are maximized to reinforce a shared signal (3, 4). Alternatively, Batesian mimicry predicts negative frequency dependence whereby an "honest" signal is copied by a "dishonest" signal. Batesian and Müllerian mimicries in nature are exemplified by tropical Amazonian butterflies, which compose diverse and fantastically complex systems (1–4).

Apheloriine millipedes, endemic to the forests of temperate North America, are aposematic (i.e., their appearance signals to predators that they are unprofitable prey). Each individual can secrete 18-fold the amount of hydrogen cyanide necessary to kill pigeon-sized birds (5, 6). Cyanogenesis occurs in 2 internal glands—one secreting the stable precursor mandelonitrile, and the other secreting the enzyme hydroxynitrile lyase (7). Their admixture creates the cyanide, and the millipede opens a muscle-actuated valve to squirt it through lateral gland openings called ozopores. Apheloriine millipede aposematism involves a conspicuous display of color patterns that vary in hue, including yellow, red, orange, and pink (Fig. 1).

For more than 30 years, scientists have postulated that color mimicry occurs within communities of Apheloriini in the Appalachian Mountains of the eastern United States (8, 9). Extensive field observations by P.E.M. suggest that certain apheloriine millipedes constitute complex mimicry rings, where variable aposematic hues and patterns are shared among sympatric species. However, these hypotheses have never been formally tested within the context of a comparative evolutionary framework. Because all apheloriines produce cyanide, a Müllerian mimetic system is considered to be operable in this millipede tribe in the Appalachian Mountains.

We report the discovery and evaluate the evolutionary dynamics of an intricate Müllerian mimicry complex among 7 apheloriine species endemic to the valleys and ridges of the Appalachian Mountains. Species in this group are differentiated by the structure of male genitalia. Color and pattern vary considerably within species, and co-occurring species often appear identical (Fig. 1). Some sympatric populations have up to 5 comimic species within an area less than 50 m^2 (video of mimicry community from site 1, Stone Mountain, Virginia; Movie S1). The evolutionary phenomenon of geographical covariation in color and pattern between apheloriine species parallels exemplary cases in butterfly species from the Amazon. This mimicry ring phenomenon in millipedes agrees with the theory of mimetic advergence that states a mimic species converges over evolutionary time with a preestablished, widespread, and variably colored model species (4).

Apheloriine millipedes are ideal models for studying aposematism and mimicry for several reasons. First, they are abundant and have high endemic species diversity; up to 43 individuals comprising 5 unique species can be found in a single 50-m^2 area. Second, apheloriines display extremely variable colors and patterns both within and between species. Third, there is no known diet–color correlation: all co-occurring species feed on decaying leaves of the same deciduous tree species (tulip-poplar, maple, and oak). Fourth, all apheloriines—indeed, the entire order Polydesmida—lack eyes. Thus, there is no sexual selection for coloration or any other intraspecific color signaling. Instead, coloration appears to be linked solely and directly to warning colors and mimicry. This is perhaps the most exciting aspect of using apheloriine millipedes as models to study aposematism. Their coloration provides a controlled and isolated view of warning signaling and mimetic resemblance. To study this mimetic symbiosis in millipedes, we investigated 7 species endemic to the Appalachian Mountains as a model system to test the following predictions of Müllerian mimicry: (*i*) species share the same aposematic signal, (*ii*) a shared signal is not attributable to close evolutionary relatedness, (*iii*) rarer species tend to mimic the most abundant and/or most highly defended model, and (*iv*) the mimicry varies geographically as a function of the model's geographical variation in color pattern.

Results

The study species are endemic to the mountainous confluence of Virginia, Kentucky, and Tennessee. We surveyed 2 species of *Apheloria* (*A.* clade A and *A.* clade B) and 5 species of *Brachoria* (*B. cedra*, *B. dentata*, *B. insolita*, *B. mendota*, and *B.* species "n"). Together, these comprise all of the region's known apheloriines.

Author contributions: P.E.M. and J.E.B. designed research; P.E.M. performed research; P.E.M. analyzed data; and P.E.M. and J.E.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed at the present address: Center for Insect Science, Department of Entomology, Forbes Building, University of Arizona, Tucson, AZ 85721. E-mail: pmarek@email.arizona.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0810408106/DCSupplemental.

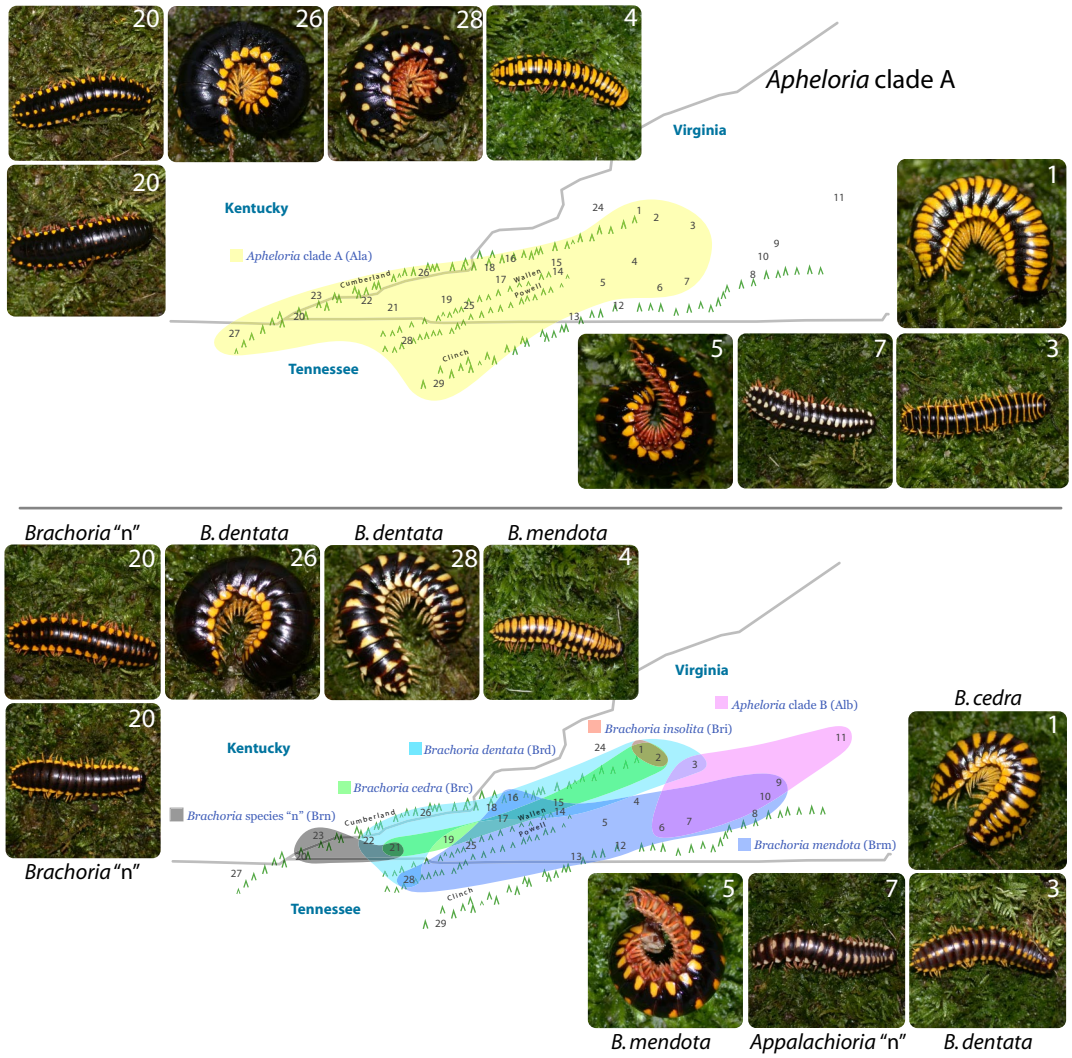

Fig. 1. Müllerian mimicry in apheloriine millipedes of the U.S. Appalachian Mountains. The geographically variable aposematic signal of several mimic species (*Lower*) is a function of the highly variable model species *Apheloria* clade A (*Upper*). The bottom right images in *Upper* and *Lower* display a single site of 7 where *A.* clade A and *B. dentata* do not share a mimetic resemblance. Species ranges are colored by taxa. Numbers indicate sites. Major mountain ridges are identified by name. Millipede images from sites 1, 5, 26, and 28 are magnified 1.65×.

Populations were sampled throughout the entirety of their known species ranges. We collected at 29 sites spanning Cumberland and Stone Mountain in the west, and through Powell, Wallen, and Clinch Mountains in the east (Fig. 1). We sampled each site randomly to collect in proportion to naturally occurring species, mitochondrial DNA haplotypes (16S rDNA), and color morph frequencies. We collected 398 millipedes at 29 sites (see color images of millipedes in Figs. S3–S8). Phylogenetic analysis of sampled individuals resulted in species categories largely consistent with traditional boundaries. Fig. S1 shows the 7 codistributed species population phylogenies and their location in the apheloriine phylogeny. Species population phylogenies were well-supported and corresponded closely with geography (i.e., closely related clades were nearer spatially, and those distantly related were further spatially or separated by significant geographical barriers). We explicitly tested for an association between mtDNA haplotypes and geography by using a phylogeographical approach with Nested Clade Analysis (10). We found a significant association between genetic and geographical variation, with inferences favoring restricted gene flow with isolation by distance concomitant with distributions along linear mountain ridges (Wallen, Powell, and Clinch Mountains; *SI Text*).

To test prediction one—that species share an aposematic signal—we measured millipede color patterns by using a fiber optic spectrometer. We then compared spectral color measurements between species to determine the presence or absence of mimicry at a site. Between 7 and 9 surface reflectances were measured on living millipedes. The number of measurements varied proportionally to the number of color patches (spots) on the millipede to encapsulate both color and pattern. Each color patch reflectance was measured twice at a distance of 3 mm from the fiber optic probe and at a 45° angle (to minimize glare from the millipede's cuticle). Reflectance measurements, $R(\lambda)$, were taken from 5,710 color patches. Ambient habitat light, or irradiance $[I(\lambda)]$, was measured in 2 old-growth Appalachian forests. We analyzed color radiance, $Q(\lambda)$, spectra, calculated as the product of reflectance and irradiance (the way color appears under forest lighting conditions) by using a nonparametric compositional analysis that allows comparison of entire color patterns (11). The compositional analysis provides the test statistic, δ, to evaluate the null hypothesis of no color differences, and yields K, which is a measure of color pattern difference (in

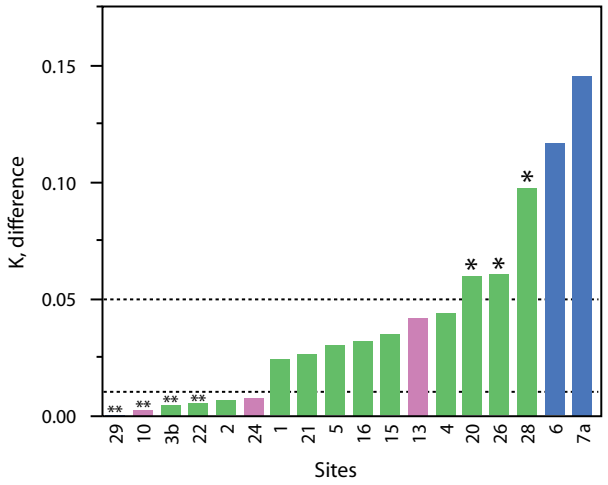

Fig. 2. Results of the nonparametric compositional analysis of color. Differences between species color patterns at a site as measured by the effect size *K*. Green bars indicate sites with mimicry; blue bars, without mimicry; and pink bars, with a single species. $K = 0.01$ (lower dashed line) is a threshold for consistent color pattern difference (11). $K = 0.05$ (upper dashed line) is the threshold for mimicry. A single asterisk indicates the site's color measurements violated homogeneity prerequisite ($P < 0.01$), and the site was reanalyzed without outliers (*SI Text*). Double asterisks indicate failure to reject the null hypothesis of no color differences ($P > 0.05$).

our case, a measure of color pattern difference between species at a site). *K* increases with color pattern differences between groups; the value of *K* increases with mean differences and other distributional differences, like variance, skewness, and shape (11, 12). Based on a study to examine the performance of the method with color pattern differences found in nature, $K = 0.01$ is a threshold for consistent color pattern difference (11). We predict that *K* values at sites with mimicry will be less than 0.05. This greater value is appropriate as a statistical threshold for mimicry because mimetic resemblance is imperfect or "impressionistic" (4), resulting from convergence of distinct ancestral color patterns from unrelated species. Hence, species involved in mimicry should retain a trace of color differentiation despite mimetic convergence. Finally, at sites with 1 species present, we predict *K* values should be less than 0.01, whereas at sites with >1 species present, without mimicry the color pattern differences should be greater than 0.05. These are estimates of visual discriminability—i.e., *K* differences are commensurate with perceived differences—and we set an estimated threshold for mimicry above 0.01 (arbitrarily at 0.05) because it is yet unknown what the precise limits of detection are for the generalized color segmentation method, which is not specific to any visual system. The null hypothesis of no color difference was rejected in all cases except 4, 2 of which were between individuals of a single species (sites 10 and 29) and the other 2 between 2 mimetic species (sites 3b and 22). The differences between color patterns, measured by the effect size *K*, ranged from 0.0000182 (Fig. 2, site 29, with 1 species and 1 color morph) to 0.145 (site 7a, with 4 species and 2 distinct color morphs). At sites with >1 species and a single color morph (i.e., with hypothesized mimetic resemblance; Table S1), *K* values ranged from 0.005128 to 0.097. At sites with >1 species and multiple color morphs (i.e., without mimicry), *K* values ranged from 0.116 to 0.145. At sites with hypothesized mimetic resemblance, *K* values were smaller than at sites without mimicry and with multiple color morphs. For sites with a single species and color morph and without mimicry, *K* values were generally smaller than at sites with mimetic resemblance, and they were much smaller than at sites with multiple color morphs. Color pattern *K* differences were notably higher at sites with a single *Brachoria* species, which always comprised multiple color morphs (sites 13 and 24, null hypothesis of no color difference rejected), over those with a single *Apheloria* species, which always comprised a single color morph (sites 10 and 29, unable to reject null hypothesis of no color difference). Thus, at most sites (9 of 14 with >1 species), we recovered consistent statistical evidence that hypothetical mimetic species share an aposematic signal, with color pattern differences much smaller ($K < 0.05$) than the others.

Table 1. Results of the test that a shared warning signal is attributed to close evolutionary relatedness

Species comparison	Likelihood difference	*P*
A. clade A and *A.* clade B	14.3318	0.7605
A. clade A and *B. cedra*	147.9892	**<0.0001**
A. clade A and *B. dentata*	179.7570	**<0.0001**
A. clade A and *B. insolita*	173.9368	**<0.0001**
A. clade A and *B. mendota*	149.7075	**<0.0001**
A. clade A and *B.* species "n"	172.4197	**<0.0001**
B. cedra and *B. dentata*	63.6940	**0.0260**
B. cedra and *B. insolita*	48.1714	0.0892
B. cedra and *B. mendota*	6.6009	0.9229
B. cedra and *B.* species "n"	48.3602	0.1006
B. dentata and *B. insolita*	17.0138	0.6651
B. dentata and *B. mendota*	66.3181	**0.0185**
B. dentata and *B.* species "n"	15.7115	0.7098

The left column shows the putative mimic species compared. The middle column contains the maximum likelihood difference between the constrained phylogeny (species branches artificially constrained to where the putative mimics are closely related sister species) versus the unconstrained Apheloriini phylogeny. If the likelihood difference between trees was significant (indicated by a *P* value <0.05, boldface, right column), we rejected the null hypothesis: that the constrained tree does not differ from the unconstrained tree and that shared coloration is likely a result of a recent common ancestry.

To test prediction two—that a shared warning signal cannot be attributed to close evolutionary relatedness—we used the phylogeny of Apheloriini (13) as an evolutionary context to determine whether mimetic species are closely related. The null hypothesis—putative mimic species are closely related evolutionarily (i.e., shared coloration is likely a result of a recent common ancestry)—is rejected in all comparisons between species and *Apheloria* clade A (except between *A.* clade A and *A.* clade B). However, in likelihood comparisons between species of *Brachoria*, except between *B. cedra* and *B. dentata* and between *B. dentata* and *B. mendota*, the null hypothesis of close relatedness cannot be rejected (Table 1), indicating that the coloration patterns likely resulted from recent ancestry rather than mimicry.

To test prediction three, we used the data from random population sampling in the field to determine whether one species is more common. Müllerian theory predicts that rarer species will mimic the most abundant and/or most highly defended model (1, 4). We measured census population size (N_c), determined from random population sampling in the field, as the sum of all individuals collected. The species *A.* clade A had by far the largest pool of individuals, with $N_c = 156$ specimens (*Apheloria* clade B, $N_c = 40$ individuals; *Brachoria* species, N_c between 9 and 37 individuals). N_c is correlated with the coalescent-based estimate of historical population size derived from the parameter theta: $\theta = N_e\mu$ (N_e indicates effective population size, and μ indicates mutation rate; Fig. S2). The most ubiquitous species, *A.* clade A, was also the most commonly encountered millipede involved in mimetic resemblance; it had the highest θ, and at 10 of 17 sites it was the most abundant. Species of *Brachoria* were less frequent and had θ values 3.5-fold lower, on average, than species of *Apheloria* (2-sample *t* test, $P = 0.0197$).

EVOLUTION

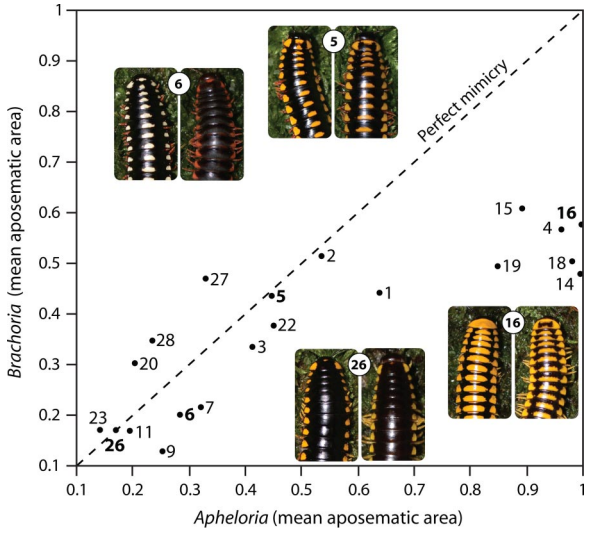

Fig. 3. Congruence between model and mimic warning patterns by site. Mean aposematic area (area covered by aposematic pattern divided by total area of collum: the flat dorsal plate of the millipede's first segment) between models and mimics by site. The dashed line indicates theoretical one-to-one correspondence expected with perfect pattern mimicry. Four representative sites (5, 6, 16, 26) are shown with the model (*Apheloria*) in the left box and the mimic (*Brachoria*) in the right box.

Based on the assumption that a species mimics the most common species, we designate the ubiquitous *Apheloria* as the model and the rare *Brachoria* as the mimic.

To test prediction four—the mimic's signal varies across its distribution as a function of the model—we tested for a correlation in pattern between model and mimic species. Both color and pattern were similar locally; however, pattern varied across the geographical space of the mimicry ring more than color, which was generally a yellow-based coloration (Fig. 1). By using a photographic analysis, we measured the amount of warning coloration in models versus mimics and tested them for correlation. In a scenario of perfect mimicry, the fraction of the body covered with aposematic color (or the size of the color patches) is the same in the model and the mimic. In a scatter plot of model versus mimic patch coverage, the mimic's coverage was variable across the geographical space of the study sites. This variation in the mimic strongly corresponds to that in the models according to a positive linear relationship (Fig. 3, Spearman's $\rho = 0.8797$, $P < 0.0001$). Thus, the correspondence in pattern area indicates that the mimic's signal covaries as a function of a geographically variable model. However, the relationship is not the one-to-one correspondence expected with perfect pattern mimicry (Fig. 3, dashed line) because patch area is slightly less in *Brachoria* (mean, 0.4172) than *Apheloria* (mean, 0.4864), but not significantly less (2-sample Mann–Whitney test, $P = 0.1593$).

Our analyses indicate that all 7 species have congruent phylogeographical patterns. Where they co-occur on Appalachian ridges and highlands, they display strong color and pattern similarity. *Apheloria* clade A, a widespread species with 6 color morphs that vary considerably across its phylogeny, appears to mimic each of the other species at no fewer than 2 localities (where it is often the most abundant). *B. cedra*, with 2 color morphs, and *B. dentata*, with 5 color morphs, mimic *A.* clade A and the remaining *Brachoria* species in at least one site. The other 4 species (*A.* clade B, *B. insolita*, *B. mendota*, and *B.* species "n") share a mimetic resemblance at fewer localities. *Apheloria* clade A and *B. dentata*, the 2 most widespread species with the greatest number of color morphs, display strong resemblance in color and pattern at 6 of their 7 shared sites.

Discussion

The 7 apheloriine species studied constitute a ring complex of Müllerian mimics in the U.S. Appalachian Mountains, comparable with mimetic radiations from the tropics. Geographical concordance in color patterns of the unrelated species provides strong evidence for mimicry, which may otherwise be a result of random co-occurrence or ancestral polymorphism. At most sites tested for color differences (9 of 14 with >1 species), mimetic resemblance occurred between species, predominately between the widespread model *Apheloria* and the narrowly endemic *Brachoria* species. Results of the phylogeographic and coalescent analyses indicate that those species with widespread distributions and larger θs evolved first, potentially shifting predator avoidance strategies from crypsis to aposematism, or from one aposematic signal to another. Based on the bright, multihue aposematic colors, species like *A.* clade A may have been responding to strong predation force from predators, like birds, with a foraging modality tuned to color and pattern signals. Millipede cuticle contains large quantities of calcium (14) and would thus be an ideal source for avian egg formation. The aposematic signal variation in *A.* clade A is probably older than that of its codistributed species and may reflect a selective regime involving components other than mimicry. Thereafter, coloration in *Brachoria* (species with smaller θ and N_e values as well as geographical distributions) most likely evolved through mimetic advergence on aposematic signals of established, widespread species (4, 15, 16).

There are no empirical examples of mutualistic parity in Müllerian mimicry; i.e., where 2 defended species converge equilaterally toward a shared color pattern. Instead, in the majority of examined cases, 1 of 2 defended species evolved unilaterally toward a preestablished color pattern via mimetic advergence (4, 16). This is plausible because theoretical predictions of mimicry between defended species suggest that the less abundant species is favored (1), and one species must have acquired the mimicked aposematic signal first. It is only after advergence that mutualism and coevolution toward an ideal color pattern may transpire ("two-step" theory of Müllerian mimicry (17, 18). In the multispecies mimicry ring of apheloriine millipedes, several species (up to 5 at site 1) co-occur, displaying a shared warning signal at a given site, and species abundance varies from a maximum in the common model (*A.* clade A) to a minimum in less common species (*Brachoria* species). These unequal abundances, θs, and distributions support mimetic advergence as the evolutionary mechanism generating the Müllerian ring pattern in these millipedes.

Several evolutionary processes, including (*i*) biotic drift, (*ii*) quasi-Batesian mimicry, and (*iii*) genetic drift (3, 19–21) have each been suggested as sources for the evolution of novel warning colors within Müllerian ring systems. Genetic drift provides the most likely explanation for color pattern diversity in apheloriine millipedes. On Stone Mountain in Virginia, sites 1 and 2 exhibited very high millipede densities and a remarkable variety of color patterns. We hypothesize that these sites provide evidence for the shifting balance model of color pattern diversity. They are consistent with observations from previous studies that invoke shifting balance to explain color variation in Müllerian mimics (3, 22). In a shifting balance scenario (23), mutation and genetic drift generate phenotypic variation upon which natural selection acts. Shifting balance would appear to be a nonsensible explanation for color pattern variation because an evolutionary "sieve" predicts an immediate elimination of new color morphs (3, 18). However, if strong purifying selection is temporarily relaxed, then color morphs can more thoroughly explore an adaptive landscape where an alternative aposematic pattern might be equally successful, if not more so, than a preexisting one. At sites with high densities of multiple species

and individuals (e.g., sites 1, 2, 20, 22, and 24), correspondingly high color and pattern variation occurred. In contrast, at sites with low densities, low color and pattern variation occurred. Several studies have suggested that strong selection, which would normally act to eliminate new color morphs, is relaxed at high densities (20, 22). Under these circumstances, a strongly positive, density-driven signal effect (density-dependent generalized avoidance, ref. 24; or another type of Allee effect refs. 25 and 26) may grant novel color morphs a higher likelihood of survival, enabling exploration of an adaptive landscape otherwise made inaccessible by purifying selection. At sites with low densities (e.g., site 4 with 9 individuals; images of sites are in Figs. S3–S8), we observed 1 morph of *A.* clade A and 1 morph of *B. mendota*—the model morph corresponded in pattern with the mimic morph. At sites with medium-high densities (e.g., site 20 with 20 individuals), we observed 2 model morphs of *A.* clade A and 3 mimic morphs of *B.* species "n"—the 2 model morphs corresponded in pattern with 2 of the 3 mimic morphs. Finally, at sites with high densities (e.g., site 2 with 30 individuals), we noted 3 model morphs in *A.* clade A, 2 in *B. insolita*, and 1 in both *B. cedra* and *Brachoria hoffmani*. Overall, there seems to be a positive correlation between density and color morph diversity, suggesting the same mode of color pattern generation in millipedes that occurs in *Heliconius* butterflies of Ecuador (22).

Every species co-occurred with and mimicked *A.* clade A in all but 2 locations, with *B. dentata* doing so in 6 of the 7 sites where they co-occurred. *Apheloria* clade A represents the single most probable model based on its (*i*) high frequency of mimetic interactions across the study region; (*ii*) high relative abundance (N_c); (*iv*) large θ and widespread distribution; and (*iv*) considerable geographical color variation. *Apheloria* clade A possesses several other qualities that are predicted for models in 1-sided advergent Müllerian mimicry (4): they are generally larger in body size, with larger warning spots (see prediction four), and are more gregarious in habit. At sites south of Powell Mountain, such as along Clinch Mountain (sites 6, 7, 8, 9, 12, and 13), red taillight color morphs in *B. mendota* predominated, whereas at sites north of Powell Mountain (sites 4, 5, 15, and 16), yellow morphs—caution, headlight, striped, and yellowjacket—were more common. Interestingly, in southern taillight sites south of Powell Mountain, *B. mendota* was not involved in mimetic resemblance (sites 6 and 7; $K > 0.10$), but in northern yellow morph sites, where *A.* clade A occurred, *B. mendota* displayed a strong mimetic resemblance with it and co-occurring *Brachoria* species (sites 4, 5, 15, and 16; $K < 0.05$). This pattern supports the role of *A.* clade A as the widespread model and potential nexus of a multispecies mimetic diversification. Where *B. mendota* overlapped geographically with *A.* clade A, so did their colors converge.

Species that are mimics in Müllerian rings are predicted to be "impressionistic" versions of the models (4). A model species is expected to be more conspicuous and larger, and to have a clearer pattern. Over evolutionary time, natural selection favors more accurate resemblance of a common and/or highly defended model. Our results from the pattern analysis are consistent with impressionistic mimicry. The mimics (*Brachoria*) have less aposematic area on average than the models (*Apheloria*). The observed correspondence in pattern area has a strong linear relationship, which is noticeably below the line indicated by a theoretical one-to-one correspondence for perfect pattern mimicry (Fig. 3). The largest pattern difference was at sites 4, 14, 15, 16, 18, 19, where models had a collum that was nearly filled with aposematic color, and therefore had a ratio approaching one (e.g., site 16, ratio 0.998 in models and 0.574 in mimics). The smallest difference occurred at sites 2, 5, and 26, where models had a collum with only a few small spots (e.g., site 5, ratio 0.448 in models and 0.434 in mimics; site 26, ratio 0.171 in models and 0.168 in mimics). These results suggest that impressionistic resemblance is stronger when the model's aposematic signal consists of a smaller area (sites 2, 5, and 26).

Conclusions

We document a remarkable example of a Müllerian mimicry ring in millipedes from the Appalachian Mountains. We satisfied 4 predictions of Müllerian theory through analyses of genetic, population, and spectral color data within a statistical and evolutionary context. By using a phylogeographic approach, we recovered a significant association of genetic groups with geographical boundaries. We note that many of these areas contain geographically and genetically independent populations of several sympatric and unrelated species. Through careful documentation of distribution, genetic isolation, and coloration and pattern, we have a biologically relevant approach for testing a correlation of pattern and color between sympatric species and their geographical distributions. We recovered consistent statistical evidence of mimetic resemblance by using spectral color data and recovered differences (aggregate color pattern differences, *K*) much smaller at sites with mimicry than without. Aposematic patterns vary drastically in apheloriine species distributions in the Appalachians; however, patterns are similar locally (coupled by mimetic resemblance), and as a result are tightly correlated across geographical scale. These patterns of color similarity are highly consistent with mimetic advergence between narrowly endemic mimic species in the genus *Brachoria* and a widespread and abundant model species in the genus *Apheloria*. Our apheloriine data provide the unique opportunity to test the predictions of Müllerian mimicry theory—one of the first mathematical models of natural selection, first described more than a century ago (1).

Materials and Methods

Phylogeographic Analyses. The genetic datasets for the 7 species allowed us to test for consistent geographical breaks and patterns across multiple codistributed species lineages. Because we collected populations by using random field sampling, haplotype frequencies are biologically relevant and appropriate for conducting nested clade analysis (NCA) (10). NCA tests for a significant association between haplotypes and geography, and it provides a framework to infer demographic processes. Of the 17 significant inferences drawn, 8 were restricted gene flow with isolation by distance (IBD), and 4 involved allopatric fragmentation. In many of the species, a large number of IBD inferences (5 of 8) were coincident with clade distributions superimposed along the linear mountain ridges (Wallen, Powell, and Clinch). Inferences favoring allopatric fragmentation were often recovered between distant clades and between those on opposite slopes of a large mountain (e.g., Cumberland Mountain).

Prediction One: Test of a Shared Aposematic Signal. We measured millipede color pattern reflectance $R(\lambda)$ by using a USB4000-VIS-NIR fiber optic spectrometer (blazed at 500 nm) attached to an LS-1 tungsten halogen light source with a QR400-7-VIS-NIR bifurcated probe (Ocean Optics). Habitat irradiance $I(\lambda)$ was measured by using a radiometrically calibrated spectrometer and a QP600-2-UV-VIS irradiance probe connected to a CC-3 cosine-corrected filter. Radiance $Q(\lambda)$ is the photon flux by wavelength arriving at a viewer's eye (e.g., a predator) reflected from a color patch $R(\lambda)$ in an ambient light environment $I(\lambda)$. Radiance was calculated as the product of reflectance and irradiance by using the assumption that the color surface of the millipede diffusely reflects ambient light over a solid angle of 2π steradians (180°) and that the receiver has a hemispherical field of view. We analyzed color radiance spectra $Q(\lambda)$ by using the segment classification method where spectra are divided into wavelength segments (between 400 and 700 nm) that have variable probabilities of exciting different photoreceptor classes in a vertebrate predator (27). This method examines the inherent physical properties of color spectra and does not process the data according to a species-specific perspective (e.g., according to a bird's-eye view), because apheloriine predators have not been precisely identified. Then, by using the segmented radiance spectra, we tested hypothesized cases of mimetic resemblance with a quantitative, repeatable method of comparing animal color patterns. We used LAD-MRPP, the least absolute difference regression and multirange permutation probability test (11), to test for aggregate differences between entire animal color patterns between mimic species at a site. It is based on a nonparametric multivariate

technique involving a combination of least absolute difference regression analysis and a multirange permutation procedure (11).

Prediction Two: Phylogenetic Tests of Mimicry. To test prediction two, that a shared warning signal cannot be attributed to close evolutionary relatedness, we used a Shimodaira–Hasegawa (28) test of alternative phylogenetic hypotheses—a strategy successfully implemented for investigating mimicry in poison dart frogs (29). We constrained tree branches for species that co-occur and share a warning signal and compared that constrained likelihood (T_0) to the unconstrained likelihood (T_1) of the Apheloriini phylogeny. If the likelihood difference between trees was significant, we rejected the null hypothesis that the constrained tree does not differ from the unconstrained tree and that hypothesized mimic species are closely related. We used a 1-tailed test to compare phylogenetic hypotheses.

Prediction Three: Test of Model and Mimic Abundances. Census population size, N_c, the sum of all individuals collected, was determined from random population sampling in the field. We used the program LAMARC, version 2.1.2 (30) to estimate the population size parameter θ. To test the relationship between θ and N_c, we used linear regression of census abundance as a predictor of θ by using the statistical package JMP, version 7 (SAS).

Prediction Four: Test of Correlation Between Model and Mimic Pattern. To evaluate prediction four, we used a photographic analysis of color pattern to measure the fraction of the collum (flattened dorsal plate of first segment) covered with aposematic color. Because millipede segments are serially repetitive and nearly identical in color and pattern (in the case of polydesmidan millipedes, repeating 20 times), the collum pattern encapsulates the whole body pattern. With the image processing package ImageJ (National Institutes of Health), we measured the area covered by the aposematic pattern relative to the total area of the collum. We measured the area from photographs taken so that the flat collum surface was parallel to the focal plane. Area is an effective estimator because millipede color pattern predominately varies according to size of the color patches. Patches remain in the same location, and when all patches are present, they are located in the left and right corners and in the middle of the anterior and posterior margins. This procedure was conducted for every specimen ($n = 312$) from sites with more than one species. We calculated the ratio of aposematic pattern area divided by total collum area for models and mimics from each site, plotted mean values to visualize the relationship, and computed a nonparametric Pearson correlation (Pearson's ρ) to measure the strength of the linear relationship with the statistical package JMP, version 7 (SAS).

ACKNOWLEDGMENTS. We thank Petra Sierwald, Trip Lamb, Carol Goodwillie, David Maddison, Jeff McKinnon, Amy Stockman, Kyle Summers, Charity Hall, Rob Marek, and anonymous reviewers for providing useful suggestions on the manuscript. Chad Spruill, Bob Marek, and Matt Walker provided great assistance collecting millipedes in the field and help in the laboratory. Bill Shear, Richard Hoffman, Rowland Shelley, and Sue McRae provided valuable discussion of concepts. This research was supported by National Science Foundation Doctoral Dissertation Improvement Grant DEB 0607996 (to P.E.M. and J.E.B.) and Partnerships for Enhancing Expertise in Taxonomy Grant DEB 0529715 (to P. Sierwald, J.E.B., and B. Shear).

1. Müller F (1879) *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies. *Trans Ent Soc Lond* 1879 xx–xxix.
2. Bates HW (1862) Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. *Trans Linn Soc Lond* 23:495–566.
3. Joron M, Mallet JLB (1998) Diversity in mimicry: Paradox or paradigm? *Trends Ecol Evol* 13:461–466.
4. Mallet J (1999) Causes and consequences of a lack of coevolution in Müllerian mimicry. *Evol Ecol* 13:777–806.
5. Eisner T, Eisner HE, Hurst JJ, Kafatos FC, Meinwald J (1963) Cyanogenic glandular apparatus of a millipede. *Science* 139:1218–1220.
6. Eisner T, Eisner M, Siegler M (2005) *Secret Weapons*: *Defenses of Insects, Spiders, Scorpions, and Other Many-Legged Creatures* (Belknap, Cambridge, MA), p 372.
7. Duffey SS, Towers GHN (1978) On the biochemical basis of HCN production in the millipede *Harpaphe haydeniana* (Xystodesmidae Polydesmida). *Can J Zool* 56:7–16.
8. Hoffman RL (1971) Millipeds of the genus *Brachoria* from southwestern Virginia (Polydesmida: Xystodesmidae). *Radford Rev* 25:83–99.
9. Whitehead DR, Shelley RM (1992) Mimicry among aposematic Appalachian xystodesmid millipedes (Polydesmida, Chelodesmidea). *Proc Entomol Soc Wash* 94:177–188.
10. Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum. Genetics* 140:767–782.
11. Endler JA, Mielke PW (2005) Comparing entire colour patterns as birds see them. *Biol J Linn Soc* 86:405–431.
12. Endler JA, Westcott DA, Madden JR, Robson T (2005) Animal visual systems and the evolution of color patterns: Sensory processing illuminates signal evolution. *Evolution* 59:1795–1818.
13. Marek PE, Bond JE (2007) A reassessment of apheloriine millipede phylogeny: Additional taxa, Bayesian inference, and direct optimization (Polydesmida: Xystodesmidae). *Zootaxa* 1610:27–39.
14. Hopkin SP, Read HJ (1992) *The Biology of Millipedes* (Oxford Univ Press, Oxford), p 233.
15. Eltringham H (1916) On specific and mimetic relationships in the genus *Heliconius. Trans Ent Soc Lond* 1916:101–148.
16. Flanagan NS, et al. (2004) Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proc Natl Acad Sci USA* 101:9704–9709.
17. Clarke CA, Sheppard PM (1960) The evolution of mimicry in the butterfly *Papilio dardanus* Brown. *Heredity* 14:163–173.
18. Turner JRG (1984) Mimicry: The palatability spectrum and its consequences. *The Biology of Butterflies*, eds Vane-Wright RI, Ackery PR (Academic, New York) pp 141–161.
19. Endler JA (1988) Frequency-dependent predation, crypsis and aposematic coloration. *Philos Trans R Soc Lond B Biol Sci* 319:505–523.
20. Mallet JLB, Joron M (1999) Evolution of diversity in warning color and mimicry: Polymorphisms, shifting balance, and speciation. *Annu Rev Ecol Syst* 30:201–233.
21. Speed MP (1993) Muellerian mimicry and the psychology of predation. *Anim Behav* 45:571–580.
22. Kapan DD (2001) Three-butterfly system provides a field test of Müllerian mimicry. *Nature* 409:338–340.
23. Wright S (1982) The shifting balance theory and macroevolution. *Annu Rev Genet* 16:1–19.
24. Darst CR, Cummings ME (2006) Predator learning favours mimicry of a less-toxic model in poison frogs. *Nature* 440:208–211.
25. Allee WC (1938) *The Social Life of Animals* (Heinemann, London), p 233.
26. Joron M, Iwasa Y (2005) The evolution of a Müllerian mimic in a spatially distributed community. *J Theor Biol* 237:87–103.
27. Endler JA (1990) On the measurement and classification of color in studies of animal color patterns. *Biol J Linn Soc* 41:315–352.
28. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116.
29. Symula R, Schulte R, Summers K (2001) Molecular phylogenetic evidence for a mimetic radiation in Peruvian poison frogs supports a Müllerian mimicry hypothesis. *Proc Biol Sci* 268:2415–2421.
30. Kuhner MK (2006) LAMARC 2.0: Maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics* 22:768–770.