

The impact of life stage and pigment source on the evolution of novel warning signal traits

Carita Lindstedt,^{1,2,3}  Robin K. Bagley,^{4,5}  Sara Calhim,¹  Mackenzie Jones,⁴ and Catherine R. Linnen^{4,6} 

¹Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä 40014, Finland

²Department of Forest Science, P.O. Box 27, 00014, University of Helsinki, Finland

³E-mail: carita.a.lindstedt@jyu.fi

⁴Department of Biology, University of Kentucky, Lexington, Kentucky 40506

⁵Department of Evolution, Ecology, and Organismal Biology, The Ohio State University at Lima, Lima, Ohio 45804

⁶E-mail: Catherine.linnen@uky.edu

Received April 3, 2021

Accepted December 6, 2021

Our understanding of how novel warning color traits evolve in natural populations is largely based on studies of reproductive stages and organisms with endogenously produced pigmentation. In these systems, genetic drift is often required for novel alleles to overcome strong purifying selection stemming from frequency-dependent predation and positive assortative mating. Here, we integrate data from field surveys, predation experiments, population genomics, and phenotypic correlations to explain the origin and maintenance of geographic variation in a diet-based larval pigmentation trait in the redheaded pine sawfly (*Neodiprion lecontei*), a pine-feeding hymenopteran. Although our experiments confirm that *N. lecontei* larvae are indeed aposematic—and therefore likely to experience frequency-dependent predation—our genomic data do not support a historical demographic scenario that would have facilitated the spread of an initially deleterious allele via drift. Additionally, significantly elevated differentiation at a known color locus suggests that geographic variation in larval color is currently maintained by selection. Together, these data suggest that the novel white morph likely spread via selection. However, white body color does not enhance aposematic displays, nor is it correlated with enhanced chemical defense or immune function. Instead, the derived white-bodied morph is disproportionately abundant on a pine species with a reduced carotenoid content relative to other pine hosts, suggesting that bottom-up selection via host plants may have driven divergence among populations. Overall, our results suggest that life stage and pigment source can have a substantial impact on the evolution of novel warning signals, highlighting the need to investigate diverse aposematic taxa to develop a comprehensive understanding of color variation in nature.

KEY WORDS: Aposematism, carotenoids, chemical defense, ecological genetics, host adaptation, polytypic coloration.

Determining how novel traits originate and become abundant is a core goal of evolutionary biology. To this end, vibrant colors that animals use to communicate their unprofitability to predators have made aposematic species attractive model systems (Cuthill et al. 2017; Briolat et al. 2018). Aposematic colors, such as orange, yellow, red or white combined with black pattern elements, can make naïve predators reluctant to attack prey due to innate biases such as neophobia (Mappes and Alatalo 1997; Thomas et al. 2004). Furthermore, predators learn to associate warning color pattern with the unprofitability of the prey and avoid attacking prey individuals with a similar appearance in future encoun-

ters (Ruxton et al. 2004). As a result, the fitness of a particular aposematic color morph is expected to be dependent on its local abundance and evolve under positive frequency-dependent selection (Rowland et al. 2007; Briolat et al. 2018). This frequency-dependence is expected to reduce genetic and phenotypic variation in signal design because rare or novel color patterns can suffer increased predation risk (Kapan 2001; Arias et al. 2016; Chouteau et al. 2016). However, intraspecific variation in warning color is surprisingly common both within populations (polymorphism) and among populations (polytypism) of aposematic species (Briolat et al. 2018). Such variation offers valuable

opportunities for characterizing the evolutionary and ecological processes that generate novel warning color traits and signal divergence (Dasmahapatra et al. 2012).

Three evolutionary scenarios could explain the origin and spread of novel warning color alleles that produce color differences among aposematic populations. First, under some circumstances, novel color alleles may evolve neutrally or under weak selection from predators. For example, when predators focus only on a specific element or combination of elements in an aposematic signal, variation in other signal elements may have little to no impact on predation rates (Winters et al. 2017; Rönkä et al. 2018a). There can also be variation in predator cognition and foraging behavior due to neophobia or dietary conservatism, which can result in weaker selection against novel signal forms by predators (e.g., Thomas et al. 2004; Aubier and Sherratt 2015), especially if the prey is aggregated (Mappes and Alatalo 1997; Mappes et al. 1999). Under this scenario, among-population differentiation at color loci evolves under migration-drift balance.

Second, even when novel color alleles evolve under negative (purifying) selection, stochastic shifts in color allele frequencies that occur when a small number of individuals colonize a new area or survive a population bottleneck could enable novel color alleles to reach threshold frequencies at which they are common enough to be protected (Mallet and Singer 1987; Turner et al. 1996; Mallet and Joron 1999; Mallet 2010). For both of the first two scenarios (neutrality and purifying selection), novel color alleles increase in frequency via genetic drift. The main difference between these scenarios is in the degree of isolation and drift needed to explain differentiation at color loci: the stronger the selection against a novel color allele, the more isolation and drift required to enable the allele to spread.

A third explanation for polytypic warning color is that novel warning color alleles are favored in some populations. Because a multitude of selection pressures can act on color alleles, the net selection coefficient for a novel color allele can be positive, even when the color morph it produces experiences an elevated predation risk. For example, warning signal efficacy may trade off with enhanced thermoregulation (Lindstedt et al. 2009; Hegna et al. 2013) or improved defense against pathogens (Friman et al. 2009). Less efficient warning signal forms can also be favored under certain dietary conditions (Talloen et al. 2004; Lindstedt et al. 2010, 2020).

Ultimately, determining the relative contribution of drift and selection and the relative importance of different selection pressures to the evolution of warning color requires integrating genetic, demographic, and ecological analyses. Despite several promising study systems (e.g., Galarza et al. 2014; Hegna et al. 2015; Lawrence et al. 2019), this level of integration remains rare with one notable exception: Müllerian mimics in the genus *Heliconius* (e.g., Merrill et al. 2015; Nadeau et al. 2016). These iconic

butterflies have contributed substantially to our understanding of warning color evolution based on endogenously produced pigmentation. However, the extent to which lessons learned from *Heliconius* apply to other aposematic species remains unclear.

For example, key factors such as the life stage that expresses warning signals (larval vs. adult) (Willmott et al. 2011; Gaitonde et al. 2018) could have profound impacts on color evolution that are not typically considered in models for warning color evolution. Specifically, adult coloration of many aposematic taxa is often subject to positive frequency-dependent selection via not only predator learning, but also positive assortative mating by color (Summers et al. 1999; Jiggins et al. 2001; Reynolds and Fitzpatrick 2007; Gordon et al. 2015). Color-based mating preferences in the adults can impact warning coloration evolution in two ways. First, when reproductive adults mate assortatively by color and color varies among populations, gene flow will be reduced and genome-wide differentiation can accumulate more readily between populations via genetic drift and divergent natural selection. Second, positive assortative mating by warning color may make it even more difficult for shifts in coloration to occur because rare color morphs will not only experience increased predation risk, but also reduced mating success (Jiggins et al. 2001; Naisbit et al. 2001). As a result, this can further strengthen selection against novel warning signal forms (Jiggins et al. 2001; Reynolds and Fitzpatrick 2007). By contrast, because coloration tends to be decoupled across ontogeny (Gaitonde et al. 2018; Galarza et al. 2019; Medina et al. 2020; Herrig et al. 2021; but see Lindstedt et al. 2016), aposematic larvae are not usually subject to direct sexual selection. Furthermore, aposematic larvae can occur in aggregations (Sillén-Tullberg 1988; Terbot et al. 2017; Wang et al. 2021), which can weaken selection against novel signals (Mappes and Alatalo 1997; Riipi et al. 2001). Thus, all else equal, we hypothesize that the cost of being a rare aposematic color allele may be lessened in the larval stage, making the evolutionary shift to a novel color morph comparatively easier for larvae. Currently, we have very little data on how novel warning signals evolve and become abundant in immature stages in natural populations (Willmott et al. 2011), which prevents comparisons with the mechanisms found to be important in the adult stage.

Similarly, the source of color pigments—whether they are produced endogenously or acquired from the diet—can influence the relative importance of top-down and bottom-up selective agents acting on warning color variation (Grill and Moore 1998; Bezzerides et al. 2007; Blount et al. 2012). Bottom-up selection may have an especially strong impact in herbivorous species, for which variation in host-plant use across space can lead to differences in access to defensive compounds acquired from the host (Codella and Raffa 1995), the visual background against which coloration is displayed (Nosil et al. 2018), and access to

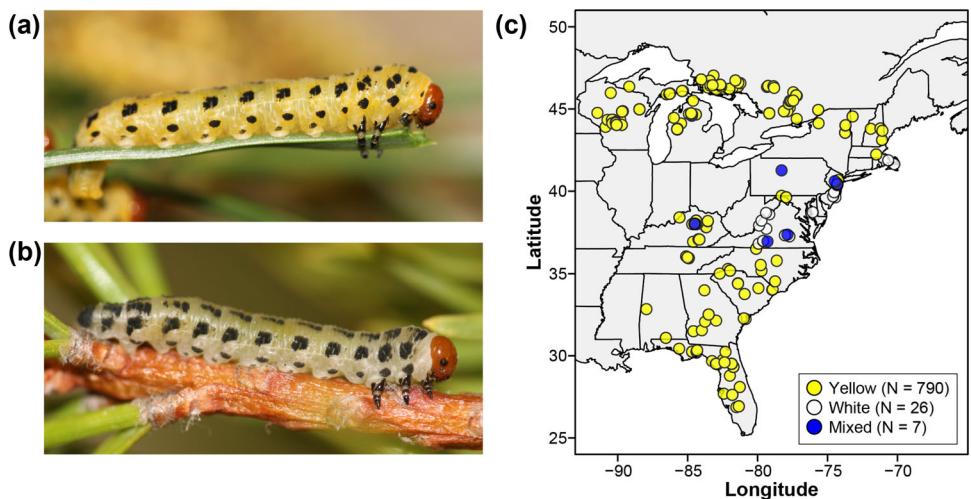


Figure 1. Appearance and distribution of white and yellow color morphs of *Neodiprion lecontei* larvae. Photographs depict representative yellow-bodied (a) and white-bodied (b) larvae (both photos by R. K. Bagley). (c) Approximate collecting locations and recorded color of 823 *N. lecontei* larval colonies collected between 2001–2004 and 2009–2016.

diet-derived pigments (Carroll et al. 1997). Although host shifts may also impact the availability of nutrients needed to produce pigments endogenously, warning signals based on endogenous pigments seem to be more robust and less variable under nutritional stress (Lindstedt et al. 2020). As a consequence, spatial variation in diet quality and nutrient availability is hypothesized to have a much stronger impact on diet-derived warning signal pigmentation than on endogenously produced pigments (Blount et al. 2012). Yet, few studies have evaluated the contribution of diet quality to geographic variation in warning color pigmentation derived from the diet.

As a starting point to test these hypotheses and to address key gaps in currently available study systems, we investigated diet-based warning color evolution in redheaded pine sawfly (*Neodiprion lecontei*) larvae. *Neodiprion lecontei* are specialist herbivores of pines and occur over a wide geographical and climatic range in eastern North America (Fig. 1). They are semisocial hymenopterans, meaning that larvae feed in large groups until the final instar, at which point they disperse from the feeding site to spin cocoons and pupate (Terbot et al. 2017). Based on their bright coloration, *N. lecontei* larvae are assumed to be aposematic. They defend against predators and parasitoids collectively using a synchronized display in which they raise their heads and regurgitate resinous droplets of sticky fluid sequestered from the host plant (Eisner et al. 1974; Codella and Raffa 1996; Costa 2006).

Throughout most of *N. lecontei*'s range, larvae have a bright, carotenoid-based yellow body color overlaid with several rows of melanic black spots (Fig. 1) (Linnen et al. 2018). Like most insects, *N. lecontei* cannot synthesize carotenoids, which must therefore be obtained from the host plant. In addition, carotenoids

are also thought to be involved in protecting the insect's own tissues from defensive toxins (scavenging free-radicals) and in supporting immune defense (Cornet et al. 2007; Babin et al. 2010; Blount et al. 2012; Dhinaut et al. 2017; Koch et al. 2018). Thus, diet quality impacts *N. lecontei* fitness in multiple ways, including the availability of defensive terpene compounds (Codella and Raffa 1995) as well as acting as a direct source of carotenoids for conspicuous warning coloration (Linnen et al. 2018) and immune function. By contrast, *N. lecontei* adults are not aposematic, and expression of pigmentation genes is decoupled between larval and adult stages of this species (Herrig et al. 2021). Therefore, larval color is not likely to be subject to sexual selection.

Although most *N. lecontei* populations have yellow larvae, field surveys have revealed the presence of white-bodied larvae in the eastern United States (Fig. 1). Previous demographic and genetic mapping analyses provide insight into the origin of this larval color variation. First, a range-wide demographic analysis indicates that the white-body phenotype is the derived state for *N. lecontei* larval pigmentation (Bagley et al. 2017). Second, a quantitative trait locus (QTL) mapping analysis of *N. lecontei* larval color suggests that the loss of yellow pigmentation in white-bodied populations is attributable to small number of large-effect loci that reduce or halt the transport of carotenoids from the gut to the integument (Linnen et al. 2018). Thus, although larvae obtain yellow pigments from their diet, the difference between white-bodied and yellow-bodied larvae is largely genetic, with a multiple-QTL model explaining ~86% of the phenotypic variation in recombinant F2 larvae derived from a cross between white-bodied and yellow-bodied populations (Linnen et al. 2018). Although

previous studies have shed light on the origin of white-bodied larvae, additional data are needed to explain how novel white-bodied alleles persisted and spread in some *N. lecontei* populations.

To explain spatial variation in diet-based larval warning color, we combined field surveys, predation experiments, population genomic analyses, and analysis of trait correlations. First, we compiled hundreds of field observations collected over 12 years to characterize the spatial distribution of the white-bodied phenotype. Second, we combined avian vision modelling and predation assays to test if *N. lecontei* larvae are aposematic and how the white allele affects the efficacy of their warning coloration. Third, we used population genomic analyses to make inferences about the historical context of white allele spread and the evolutionary mechanisms responsible for among-population variation in color. Fourth, we examined phenotypic correlations between larval color and other larval defense traits (immune function, chemical defense) to determine whether changes in larval coloration were likely to have correlated effects on other traits. Fifth, we combined field observations and chemical analysis of common pines to determine whether white-bodied larvae are found disproportionately on low-carotenoid hosts. We then integrated these data to make inferences regarding the most likely scenario under which the white allele spread (negative selection, neutrality, or positive selection), as well as the relative importance of top-down and bottom-up selection pressures. When contrasted with study systems involving aposematic adults that use endogenous warning pigments, our results suggest that life stage and pigment source may have profound impacts on warning color evolution.

Materials and Methods

GEOGRAPHIC DISTRIBUTION OF LARVAL COLOR IN *N. lecontei*

To describe the geographic distribution of the derived white-body morph in *N. lecontei* larvae, we examined field notes and collection logs from multiple collecting trips that took place over 12 years (2001–2004 and 2009–2016). During this time, we collected 823 distinct *N. lecontei* larval colonies, each containing anywhere from a few individuals to over 100 individuals (Table S1). We used colonies—not individual larvae—as our unit of replication because each colony typically comprises the offspring of a single mated female. Because we consistently noted all colonies that were not uniformly yellow (color was assessed by eye), we could confidently assign each collected colony to one of three color categories: white-bodied ($N = 26$), mixed-color ($N = 7$), or yellow-bodied ($N = 790$). To visualize the geographical data, we used R version 3.6 (R Core Team 2020) and the “maps” (version 3.3.0), “maptools” (version 0.9-5), “mapplots” (version 1.5.1), and “mapdata” (version 2.3.0) packages.

LARVAL COLOR AS AN ANTIPREDATOR DEFENSE

Conspicuousness of yellow- and white-bodied N. lecontei larvae against different host plants

To determine if the color of *N. lecontei* larvae could function as a warning signal to predators (as has been assumed [Costa 2006]), we first asked whether birds can distinguish (1) the white- and yellow-bodied larvae from each other and (2) the white- and yellow-bodied larvae from their natural background. We quantified the conspicuousness of larvae against the three most common pine hosts for Central *N. lecontei*: *Pinus virginiana* (VA pine), *Pinus echinata* (shortleaf pine), and *Pinus rigida* (pitch pine) by using mathematical models that simulate the vision of an insectivorous avian predator (blue tit, *Cyanistes caeruleus*) (Vorobyev and Osorio 1998; Vorobyev et al. 1998). This approach assumes that the blue tit visual model available is representative of avian predators *N. lecontei* may encounter in North America.

Larvae used for these analyses were derived from a laboratory colony that was established from a single mixed-color colony collected on an introduced host, *Pinus sylvestris*, in Piscataway, NJ ($40^{\circ} 32'58.4''N$, $74^{\circ} 25'50.9''W$) in August 2013. For host material, we collected five clippings from each of three individual trees. All clippings were collected in July 2018 from the University of Kentucky Arboretum in Lexington, KY. This does not cover the whole geographical variation in color among host plants, but will give a rough estimate of the potential variation in contrast between the larval color and host plant species.

To quantify color of *N. lecontei* larvae and their host plants, we recorded reflectance spectra from each larval and host sample with a USB2000 spectrophotometer equipped with a 200- μ m probe and a PX-2 pulsed xenon light source (Ocean Optics, Largo, FL). To minimize stray light, all reflectance measurements were taken in a dark, windowless room. Plant and larval spectra were recorded with SpectraSuite software (Ocean Optics) relative to a WS-1-SL reflectance standard (Ocean Optics). Each spectral measurement was taken with an integration time of 0.2 s, a boxcar width of 4, and 10 spectra averaged per measurement. For larvae, we recorded nine reflectance spectra across the dorsal, lateral, and ventral surfaces (three spectra per surface) of 20 individuals (10 white and 10 yellow) chosen at random from our laboratory population and immobilized with CO₂. The reflectance probe was held at a 90° angle to each larva and as close to the larvae as possible, while still enabling us to ensure that we were recording measures from nonmelanic parts of the larval body. For host plants, we recorded 15 spectra across the new foliage (current year's growth), old foliage (previous years' growth), and bark (five spectra per region) from each of 45 clippings (15 fresh clippings from each *Pinus* species). The reflectance probe was held at a 90° angle and as close as possible to the plant material. We used the program CLR: Color Analysis Programs version 1.05 (Montgomerie 2008) to bin and trim the raw spectra to 1-nm

intervals between 300 and 750 nm. We then averaged each bin for each sample to obtain a single summary spectrum for the dorsolateral (dorsal + lateral) and ventral surfaces of each larva and for the foliage and bark portions of each plant clipping. These averages were used in vision model analyses.

To predict whether a blue tit could discriminate between *N. lecontei* larvae and their natural pine backgrounds in terms of their color and luminance, we used a discrimination threshold model that assumes that noise in the receptors limits discrimination ability (Vorobyev and Osorio 1998; Vorobyev et al. 1998). The model uses information about the visual system, such as the sensitivity and relative abundance of different receptor types, and estimates of noise that arises in the photoreceptors. An average spectrum per each stimulus type was modeled for a blue tit's photon catch values for the single and double cones with a standard D65 irradiance spectrum. Color vision in birds stems from the four single cone types, whereas luminance-based tasks apparently stem from the double cones (Osorio and Vorobyev 2005). For the color model, we therefore used the four single cones, whereas the luminance model was based on the double cones (Siddiqi et al. 2004). We used a Weber fraction of 0.05 for the discrimination model for the most abundant cone type and the relative proportion of cone types in the blue tit retina (long wave = 1.00, medium wave = 0.99, short wave = 0.71, and ultraviolet [UV] sensitive = 0.37).

Using the same approach as in Nokelainen et al. (2012), we calculated "just noticeable difference" (JND) values for every combination of larval body region (dorsolateral and ventral) and host background (foliage and bark for *P. virginiana*, *P. echinata*, and *P. rigida*) for each larva. For reference, JND values for prey/background combinations that are <1 are indistinguishable, values between <1 and 3 are hard to distinguish unless under optimal conditions, and values >5 are easy to tell apart under most conditions (Vorobyev and Osorio 1998).

Predator avoidance learning assays with white and yellow *N. lecontei* larvae

If *N. lecontei* larvae are aposematic, their unpalatability and conspicuousness should yield avoidance learning in potential predators. To test this prediction, we conducted avoidance learning experiments in plywood cages (plywood $50 \times 50 \times 70$ cm [$w \times d \times h$]) at Konnevesi Research Station (described in, e.g., Lindstedt et al. 2008). Prey were offered through a hatch behind a visual barrier, which enabled us to record the exact time of prey detection because the bird had to go around the barrier or fly on top of it to see the prey.

Eleven great tits were tested with white *N. lecontei* larvae, 11 great tits with yellow *N. lecontei* larvae, and 10 great tits with light green *Diprion pini* larvae. *Diprion pini* were included as a

cryptic control: although they are similarly chemically defended, they lack the conspicuous coloration of *N. lecontei* larvae (Lindstedt et al. 2011b) (Fig. S1). *Neodiprion lecontei* larvae were obtained from the same mixed-color colony described above. After sorting by color (white or yellow), we froze live larvae at -80°C until needed. We used spectrophotometric measurements (collected as described above) to verify that human-sorted larvae reflected true underlying differences in color and that color differences between thawed larvae that had been frozen recapitulated differences between living larvae (Table S2). *Diprion pini* larvae were obtained from laboratory insect culture described in Lindstedt et al. (2018) and killed by freezing. We thawed all larvae for 30 min prior to the experiments with birds. Because variation in larval defensive behaviors could affect how larvae appear to birds (see *Results* and also Lindstedt et al. 2018), we used dead larvae to isolate the effects of color and taste on bird behavior independent of larval behavior. Although we do not know how freezing and thawing impact taste, we note that birds readily ate similarly treated mealworms (see below). Due to the oily structure of the pine sawfly larvae's defensive fluid, the chemical composition of the fluid stored in the defensive glands is stable over long time periods (Eisner et al. 1974).

After birds were acclimated to the cage, *N. lecontei* larvae were offered one at a time and dorsal side up on a white dish (i.e., all the prey items were easy to detect for birds) in three consecutive trials similar to Lindstedt et al. (2008). At the beginning of each experiment, a bird's motivation to feed was confirmed by offering them a mealworm that had been killed by freezing and thawed prior to the experiment. To measure learning rate, we used attack latency (the time from when the bird noticed the prey to when it touched/attacked the prey with its beak [Lindstedt et al. 2008, 2011a]). Because hunger can affect a predator's readiness to attack defended prey (Sandre et al. 2010), we also quantified the hunger level as the mass of thawed mealworms eaten after the experimental trials.

We used delta attack latency, that is., change in attack latency in between first and last trial (trial 3 – trial 1 hesitation times), as a measure of avoidance learning. If *N. lecontei* larvae are aposematic, we expected to see an increase in attack latency between the first and third encounter, indicative of an increased reluctance to attack similarly colored prey once birds learned to associate their bright color with an unpleasant taste. For green and chemically defended *D. pini* larvae, we expected either no change or a less pronounced change in attack latency, indicative of reduced avoidance learning in the absence of a conspicuous color cue. We used a linear model ("lm" function from the base R package), because each bird had only one data point. Model diagnostics were performed on the residuals. We used planned contrasts to test for differences in delta attack latency: (i) between *D. pini* and

(the average) *N. lecontei* and (ii) between white and yellow *N. lecontei*. To control for hunger level, the linear model included the amount of meal worms (mg) eaten after all trials as a covariate. We used the “summary” function to obtain the results table.

POPULATION GENOMICS OF LARVAL COLOR

Assuming that a novel white allele would have been deleterious in a predominantly yellow-bodied population, isolation and drift would have been required for the allele to persist and eventually reach high frequencies in some parts of the *N. lecontei* range. Such a scenario would have impacted not only color loci, but also loci distributed throughout the rest of the genome, giving rise to color-associated population structure. Thus, although there is no way to definitively test whether the white allele was deleterious when it first arose, we can ask whether patterns of genomic variation are consistent with a historical scenario that would have facilitated the spread of a deleterious allele. To answer this question, we used double-digest restriction-associated DNA (ddRAD) sequencing to genotype white- and yellow-bodied larvae collected across the eastern United States.

Sampling for population genomic analyses

A previous range-wide analysis revealed that *N. lecontei* populations fall into three main genetic and geographic clusters that were isolated in different pine refugia during the Pleistocene, which were dubbed North, Central, and South (Bagley et al. 2017). Because white-bodied populations are restricted to the Central cluster, we focused on increasing sampling in this region only to avoid confounding color-associated divergence with refugia-associated divergence. In total, we sampled larvae from 29 locations throughout the Central region, for a total of 65 larvae. This sample consisted of 19 white-bodied larvae, 2 larvae from mixed-color colonies, and 44 yellow-bodied larvae (Table S3). Because colonies typically consist of siblings, each larva came from a different colony. Because *Neodiprion*, like all Hymenoptera, are haplodiploid and many of the analyses we used assume diploid data, we preferentially chose large larvae (which tend to be female) for DNA extraction and used genome-wide heterozygosity estimates (*-het* option in VCFTOOLS (version 0.1.15 [Danecek et al. 2011]) to confirm that each individual was diploid (as in Bagley et al. 2017).

DNA extraction, library preparation, and genotyping

We extracted DNA and generated ddRAD sequencing libraries following protocols outlined in Bagley et al. (2017), but used a new set of sequencing adapters and PCR primers. Briefly, we digested DNA with the restriction enzymes *EcoRI* and *NlaIII*. We then ligated an adapter containing one of 48 unique 5 to 10 bp variable length barcodes to each sample (Burford Reiskind et al. 2016) (Table S4). Following pooling and size selection

(379 ± 76 bp) on a PippinPrep (SageScience, Beverly, MA), we amplified libraries using high-fidelity polymerase (Phusion; NEB, Ipswich, MA) with primers that incorporated an Illumina multiplexing index (Table S5), as well as a string of four degenerate bases for PCR duplicate detection. We then sequenced these libraries with 150-bp reads on two replicate lanes of an Illumina HiSeq 4000 housed at the University of Illinois Roy J. Carver Biotechnology Center.

To quality-filter and demultiplex raw sequencing reads, we used the default settings of the *process_radtags* module in STACKS (version 1.46 [Catchen et al. 2013]). We then aligned the resulting reads to a high-coverage, linkage-group anchored *N. lecontei* genome assembly (version 1.1; GenBank assembly Accession no. GCA_001263575.2 [Vertanick et al. 2016; Linnen et al. 2018]) using the “very sensitive” end-to-end alignment mode in BOWTIE2 (version 2.3.1 [Langmead and Salzberg 2012]). We then used SAMTOOLS (version 1.3 [Li et al. 2009]) to retain only uniquely mapping reads with Mapping Quality (MAPQ) scores ≥ 30 . Putative PCR duplicates were identified based on the sequence of the four degenerate bases in the index read (provided as a second fastq file) and removed using a custom python script (defRemove_ddRAD_PCRduplicates.py, files in Supporting Information). We then constructed RAD loci from the filtered alignments in STACKS’ *ref_map.pl* pipeline (version 1.46 [Catchen et al. 2013]). To ensure high-confidence genotype calls (Kenny et al. 2010; Peterson et al. 2012), we required a minimum stack depth of 10 (*-m 10*). We then used STACKS’ *populations* module to call SNPs present in $\geq 70\%$ of individuals (*-r 0.7*). Depending on the analysis (see below), we also used STACKS to retain SNPs with a minimum minor allele frequency of 0.01 (*-min_maf 0.01*) and/or to randomly sample one SNP per locus (*-write_random_snp* function in STACKS) to minimize linkage disequilibrium between markers. Finally, as an additional step to remove paralogously mapping loci, we performed exact tests of Hardy-Weinberg equilibrium in VCFTOOLS (version 0.1.15 [Danecek et al. 2011]) and excluded markers displaying significant heterozygote excess ($P < 0.01$). Summaries of the number of reads and loci per sample are included in Table S6.

Historical context of white allele spread

We evaluated the historical demographic context in which the white allele spread in two ways. First, we asked whether there was evidence of discrete population structure separating white-bodied and yellow-bodied larvae. The presence of strong color-associated structure would be indicative of a corresponding historical event that could have facilitated the spread of a deleterious white allele, such as a rapid range expansion, population bottleneck, barrier to gene exchange, or rare long-distance migration. To evaluate population structure, we used both model-based and model-free clustering approaches. For the model-based ap-

proach, we used the program ADMIXTURE (version 1.23 [Alexander et al. 2009]) to assign the proportion of ancestry for each individual from K ancestral populations. We performed 100 independent runs for each K value between $K = 1$ through $K = 10$. The optimal K was selected by comparing five-fold cross-validation (CV) error values for each value of K as suggested in the ADMIXTURE manual. We then used CLUMPAK (version 1.1 [Kopelman et al. 2015]) to summarize and evaluate stability of assignment solutions across the 100 replicates of each K . For the model-free approach, we used a discriminant analysis of principal components (DAPC) implemented using the *dapc* function in the ADEGENET R package (version 1.3-9.2 [Jombart 2008; Jombart et al. 2010]). We identified the optimal number of clusters from $K = 1$ through $K = 10$ using a K -means clustering algorithm. To avoid overfitting of discriminant functions, following α -score optimization (Fig. S2), we performed this analysis using one principal component (PC) that explained $\sim 7\%$ of total variation (Fig. S3). The clustering solutions were then compared using Bayesian information criterion (BIC), following Jombart et al. (2010). For both population-structure analyses, we used an SNP dataset filtered on minor allele frequency (MAF > 0.01) and linkage disequilibrium (LD, 1 SNP per RAD locus). This dataset contained 11,603 SNPs.

Second, we asked whether there was evidence to support a scenario in which a novel white allele spread by allele surfing, the rapid propagation of a new—and possibly deleterious—mutation at the edge of a range expansion via successive founder events (Edmonds et al. 2004; Peischl et al. 2013). This scenario predicts that there will be a decline in genetic diversity corresponding to the path of range expansion (due to successive founder events), with white-bodied populations exhibiting reduced genetic diversity and located at the furthest edges of the range. To evaluate this prediction, we used the *-het* option in VCFTOOLS to compute observed heterozygosity for each individual from a SNP dataset that included all variable sites and no MAF filters (42,262 SNPs). To determine whether white-bodied, mixed-color, and yellow-bodied colonies differ in heterozygosity, we used a Kruskal-Wallis test (“kruskal.test” in R). To determine whether heterozygosity correlates with either latitude or longitude, we used a Spearman’s rank correlation test (cor.test with “method = spearman”). Because we detected strong population structure between a western yellow-only group and an eastern group with all colors (see below), we restricted our heterozygosity analyses to the eastern group to avoid confounding different sources of population structure.

Evolutionary processes maintaining color differentiation in extant populations

If body-color alleles evolve neutrally, among-population differentiation at color loci should reflect a balance between genetic

drift and migration. Under this scenario, differentiation at color loci should be similar to neutral differentiation across the rest of the genome. Alternatively, if natural selection maintains geographic variation in larval color in the face of gene flow, differentiation at color loci should exceed genome-wide levels of differentiation. To determine whether any SNPs exhibited evidence of selection between white-bodied and yellow-bodied populations collected from different geographic locations, we conducted an F_{ST} outlier analysis with BayeScan version 2.1 (Foll and Gaggiotti 2008). This method assumes a model in which subpopulations share a common migrant gene pool from which they differ due to varying degrees of isolation and drift, and uses a Bayesian approach to estimate the probability that each locus is under selection (Foll and Gaggiotti 2008). We converted the original VCF file to GESTE/BayeScan format using PGDSpider (version 2.1.1.5 [Lischer and Excoffier 2011]). Our pilot analysis consisted of 20 runs with 5000 iterations each; and our Markov chain was run for a total of 100,000 iterations, with a burn-in of 50,000 generations, a thinning interval of 10, and a prior odds ratio of 100. To maximize our chances of retaining informative SNPs in linkage disequilibrium with a color locus, we retained all SNPs with MAF > 0.01 (38,852 SNPs). As in the heterozygosity analysis, we restricted this analysis to the eastern mixed-color group.

CORRELATION BETWEEN LARVAL COLOR AND OTHER DEFENSIVE TRAITS

Because dietary carotenoids serve several important functions in insects, there may be trade-offs between the use of carotenoids for warning coloration and allocation to essential functions such as immunity and protection of an insect’s own tissues from the toxic compounds used in chemical defense (Cornet et al. 2007; Blount et al. 2009, 2012; Babin et al. 2010; Dhinaut et al. 2017). This hypothesis predicts that there will be genetic correlations between larval body color and other carotenoid-related traits. To test this prediction, we evaluated trait correlations in recombinant F_2 progeny produced from a cross between *N. lecontei* females from a laboratory line derived from a white-bodied population (Valley View, VA; $37^{\circ}54'47''N$, $79^{\circ}53'46''W$) to males from a laboratory line derived from a yellow-bodied population (Bitley, MI; $43^{\circ}47'46''N$, $85^{\circ}44'24''W$) (same cross as in Linnen et al. 2018, but different families and individuals). Because *Neodiprion* are haplodiploid, these crosses generated hybrid, diploid F_1 females and nonhybrid haploid males. To generate haploid F_2 males, we reared the offspring of virgin F_1 females. The haploid F_2 male offspring, the products of recombination in virgin hybrid F_1 females, therefore represent the whole diversity of different combinations of color alleles and their associated traits. This allowed us to test whether white/yellow alleles are likely to have correlated effects on other defense traits through pleiotropy or linkage

(i.e., whether there are genetic correlations). Because estimation of genetic correlations requires prohibitively large sample sizes, we used phenotypic correlations as a proxy (Cheverud 1988; Roff 1996; but see Willis et al. 1991). Although the magnitude of phenotypic correlations should be interpreted with caution because they can differ substantially from genetic correlations (Kruuk et al. 2008), they are nevertheless informative with respect to the presence and direction of underlying genetic correlations (Roff 1996; Kruuk et al. 2008; Dochtermann 2011; but see Hadfield et al. 2007).

All larvae were fed *Pinus banksiana* (jack pine) foliage ad libitum. Conditions for the insect cultures are described in Linnen et al. (2018). We measured color as a proxy for carotenoid content and defensive traits (presence, quantity, and quality of defensive regurgitant; encapsulation response) in $N = 212$ F_2 male larvae from 10 F_1 mothers. We measured these traits in mature feeding instars only (Linnen et al. 2018).

Carotenoid content of recombinant F_2 males

To estimate carotenoid content for each larva, we took five measurements with the spectrophotometer from each of three body regions: the dorsum, lateral side, and ventrum. All other details of spectra collection were as described above. We then used the program CLR: Color Analysis Programs version 1.05 (Montgomerie 2006) to process raw spectra and compute color saturation (S1B) values for each larva, which correlate negatively with carotenoid content in multiple taxa (Butler et al. 2011). To obtain a single S1B value per larva, we averaged across the 15 reflectance spectra.

Chemical defense of recombinant F_2 males

To evaluate larval chemical defense, we gently poked each larva twice with a capillary tube (the capillary tube did not break the skin of the larva) between the front legs on the ventral side and recorded presence/absence of a defensive regurgitant (see Lindstedt et al. 2018). For consistency, all pokes were performed by the same researcher and larva were chosen at random (without respect to color) from their rearing boxes. If a defensive regurgitant was produced, we collected it in a 5- μ L capillary tube. To estimate defense fluid quantity, we measured the length of the regurgitant in the capillary tube with digital calipers. To control for the effect of body size on the amount of defense fluid produced, we also measured larval body length. Larval defense fluid samples in capillary tubes were placed in 1.5-mL microcentrifuge tubes containing 500 μ L *n*-hexane and stored in a -20°C freezer until analysis.

To determine monoterpenes and other terpene compounds in the defensive regurgitates, a Shimadzu, GC-2010 Plus (Shimadzu Corp., Kyoto, Japan) gas chromatograph fitted with a flame ionization detector (FID) was used. Samples dissolved in hexane

were injected (splitless injection of 1 μ L sample, inlet temperature 290°C) into a Zebron ZB-5MSi capillary column (length 30 m, inner diameter 0.25 mm, film thickness 0.25 μm ; Phenomenex Inc., Torrance, CA). Helium was used as the carrier gas at a flow rate of 1 mL min^{-1} (83.6 kPa). The temperature program started at 50°C (1.5-min hold) and the column oven was heated with a rate of $10^\circ\text{C min}^{-1}$ to 180°C , and then at a rate of $2.5^\circ\text{C min}^{-1}$ to 290°C (10-min hold). The temperature of FID was 290°C . We used L-fenchone as an internal standard. We identified the compounds with separate runs using a GC equipped with a mass spectrophotometer (Shimadzu 2010 GC/MS; Shimadzu Corp., Kyoto, Japan) and otherwise similar conditions as in the GC/FID analyses. For the identification of terpenes, we compared the mass spectra (electron ionization 70 eV) of the obtained peaks to those of the library. We used the retention times of terpene peaks in GC/MS chromatograms to assign the compounds in GC/FID chromatograms. Masses (μg) of different compounds in the defense fluid samples were calculated by dividing the area of a compound's peak by area of the internal standard's peak, and then multiplying it with the concentration of the internal standard (20 ppm). Concentration (%) of monoterpene compounds (total mass of all the monoterpene compounds/sample) and other terpene compounds (total mass of all the other terpene compounds/sample) in the defense fluid sample was calculated with the following formula: (HPLC value in mg/defense fluid sample size [μL]) $\times 100$.

Encapsulation response of recombinant F_2 males

To evaluate immune defense, we measured encapsulation response as described in Lindstedt et al. (2011b) after chemical defense measurements were taken. Larvae were first anaesthetized with CO_2 (Lindstedt et al. 2018). A sterilized hypodermic needle was then used to puncture the skin on the dorsal part of the individual and a nylon implant (length 3 mm, diameter 0.11 mm) was inserted into the resulting hole. After 24 h, the implant was removed, dried, and photographed under a Zeiss DiscoveryV8 microscope equipped with an Axiocam 105 camera and ZEN lite 2012 software (Carl Zeiss Microscopy, LLC, Thornwood, NY). Every implant was photographed three times from three different angles (Rantala et al. 2000). Image analysis software ImageJ 1.47v (National Institutes of Health, USA) was then used to quantify the gray value of the implant, which was used as a measure of the encapsulation reaction. The gray value of the background was subtracted from the gray value of the implant to correct for potential variation in light source. The darker the implant, the stronger the immune response.

Statistical analyses for phenotypic correlations

Phenotypic correlations between color (carotenoid content), chemical defense (defensive behavior, quantity, and terpene con-

centration of defense fluid), and immune defense (encapsulation response) were evaluated in R (version 3.6.2 [R Core Team 2020]). We used generalized linear mixed models using the function “glmer” from the package “lme4” (version 1.1-23 [Bates et al. 2014]) and linear mixed models using the function “lme” from the package “nlme” (version 3.1-147 [Pinheiro et al. 2020]). For all models, colony identity was included as a random effect and the continuous predictors were centered, so that the intercept term estimates the outcome variable at the average predictor levels. Chemical concentration variables were modeled using total volume as an offset term and Gaussian errors. Encapsulation response was also modeled using Gaussian errors. To account for large skew and excess zeros in defense fluid quantity, this trait was transformed (0.00001 added to all data points) and modeled using Gamma family. As a binary outcome variable, the probability of defense was modeled using binomial family.

CORRELATION BETWEEN LARVAL COLOR AND HOST-PLANT CAROTENOID CONTENT

Because the yellow body color of *N. lecontei* larvae is derived from carotenoids acquired from the host plants (Linnen et al. 2018), a reduction in carotenoid availability in the larval diet (adults are nonfeeding) could potentially favor the loss of yellow body color in some populations. To evaluate this possibility, we first asked whether the prevalence of white-bodied colonies differed among host plant species. To do so, we examined color and host species for the 823 colonies described in “GEOGRAPHIC DISTRIBUTION OF LARVAL COLOR IN *N. Lecontei*” section. We restricted our focus to the Central region only and, within this region, the three primary pine hosts for which we had substantial sample sizes: *P. echinata*, shortleaf pine ($N = 63$ colonies); *P. rigida*, pitch pine ($N = 50$ colonies); and *P. virginiana*, Virginia pine ($N = 167$ colonies). To determine whether the counts of different colony colors differed among the three hosts, we used a Fisher’s exact test (“fisher.test”) followed by post hoc tests (“fisher.multcomp” function from the RVAideMemoire version 0.9-73 R package).

Because the prevalence of white-bodied colonies differed among the three pine hosts (see *Results*), we next asked whether the carotenoid content of these hosts differed. To measure carotenoid content in *P. rigida*, *P. echinata*, and *P. virginiana*, we sampled ~ 10 –20 needles from each of the 45 clippings (five clippings from three individual trees per species) that were used to measure host reflectance spectra (see “Conspicuousness of yellow- and white-bodied *N. lecontei* larvae against different host plants” section). We then cut the needles into 2- to 5-mm segments and stored them in 1.7-mL microcentrifuge tubes in the -20°C freezer until pigment extraction. To extract pigments and quantify carotenoid content, we followed protocols outlined in Minocha et al. (2009). First, we placed 15.0 mg (± 0.3 mg) of

chopped needle into a 2.0-mL tube, to which we added 1.5 mL of 200-proof ethanol. Next, samples were vortexed and stored in the dark at 65°C for 24 h. After cooling to room temperature, samples were vortexed for approximately 1 min at medium speed and centrifuged at $13,500 \times g$ (rcf) for 5 min. For each sample, we then transferred 200 μl of the supernatant to each of three wells (replicates, to reduce measurement error) in a clear, flat-bottomed 96-well plate. Using a Synergy HT plate reader (Biotek Instruments, Winooski, VT) and Gen5 software, we read the absorbance of each sample at 664, 649, and 470 nm. We then used these values and an equation for ethanol solvent (Lichtenthaler 1987; Minocha et al. 2009) to compute carotenoid concentration. We used the *lme4* (version 1.1-21), *lmerTest* (version 3.1-1), and *multcomp* (version 1.4-12) R (version 3.6.2) packages to fit a linear mixed model to the data (with host species and individual tree as fixed and random effects, respectively) and, because the host effect was significant, to compare pairs of hosts with post hoc Tukey contrasts.

Results

WHITE-BODIED LARVAE ARE RESTRICTED TO THE EASTERN COAST OF THE UNITED STATES

Our collecting records (summarized in Table S1) confirmed that all-white colonies are restricted to the east of the Appalachian Mountains, primarily in coastal states in the eastern United States (Fig. 1c). Although some states had records of both all-white and all-yellow colonies, these were never collected at the same site. Mixed-color colonies were rare and 4 out of 7 were collected on nonnative pines in disturbed areas (e.g., parking lots). Notably, a single mixed-color colony was collected west of the Appalachians. This mixed-color colony was the only colony out of 278 colonies collected in Lexington, KY to contain any white-bodied larvae (Table S1). Given the rarity of white larvae in the area, the mixed-color colony may have been the result of a rare long-distance migration event or a de novo mutation. Overall, these results reveal that much of the variation in larval color is distributed among rather than within populations.

WHITE AND YELLOW *N. lecontei* LARVAE ARE APOEMATIC

Regardless of the larval body color (white or yellow), body region, host species, and host tissue type, *N. lecontei* larvae were highly conspicuous against a pine background in terms of color contrast (all JND values > 5 ; Fig. 2a). Most larval/host combinations were also highly conspicuous in terms of luminance, but some individual larvae had JND < 5 for some host background types (Fig. S4). These results demonstrate that a blue tit could distinguish between larvae and their host plant, which

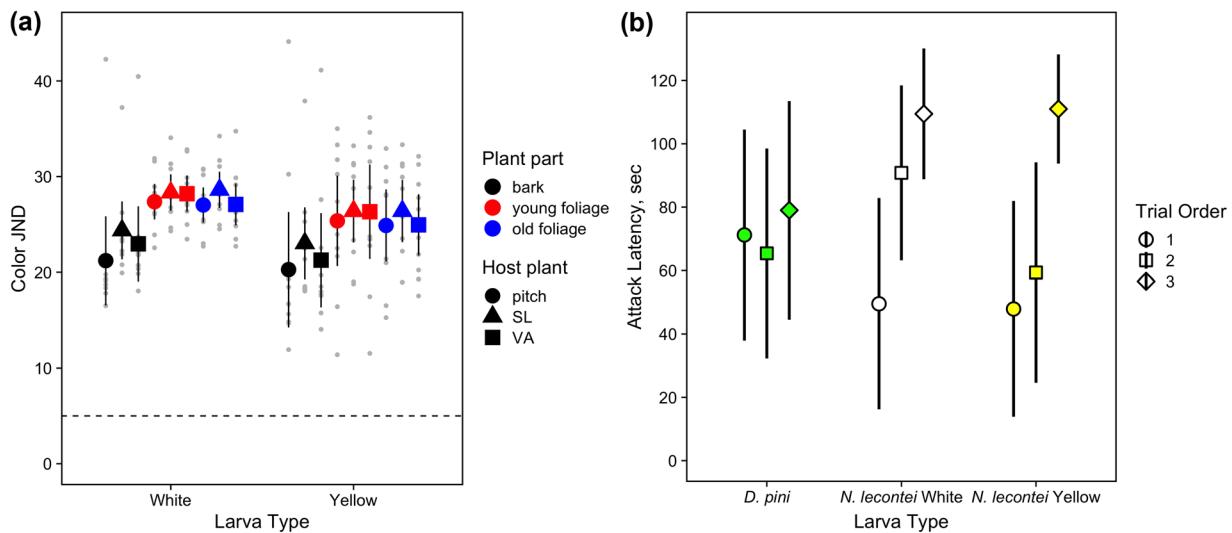


Figure 2. White and yellow *Neodiprion lecontei* larvae are aposematic and do not differ in signal efficacy. (a) Color contrasts against different host plants and host-plant parts. For reference, JND values for prey/background combinations that are <1 are indistinguishable, values between <1 and 3 are hard to distinguish unless under optimal conditions, and values >5 are easy to tell apart under most conditions (Vorobyev and Osorio 1998). Dashed line shows the threshold value for $JND = 5$ above which objects should appear clearly conspicuous for blue tits (*Cyanistes caeruleus*). Changes in attack latency (seconds) (b) across three trials for three different types of pine sawfly in avoidance learning assays. Error bars are 95% confidence intervals.

Table 1. Predator response with respect to larval type. Output table from a linear model (y = change in attack latency, gaussian errors, t statistic) using planned contrasts (*D. pini* vs. pooled *N. lecontei*; white vs. yellow *N. lecontei*).

Change in attack latency	Estimate	SE	T or z	P
Intercept	11.428	22.778	0.502	0.6200
<i>D. pini</i> vs. <i>N. lecontei</i>	56.175	23.972	2.348	0.0265
White <i>N. lecontei</i> vs. Yellow <i>N. lecontei</i>	-4.356	25.668	-0.170	0.8665
Hunger (mg mealworms eaten)	-14.043	33.213	-0.423	0.6758

supports the hypothesis that *N. lecontei* larvae are conspicuous to their predators. Moreover, blue tits should be able to distinguish between white larvae and yellow larvae (color contrast JND 7.4, luminance contrast JND 3.6).

We also found that captive great tits learned to avoid brightly colored white and yellow *N. lecontei* larvae more quickly than light green and chemically defended *D. pini* sawfly larvae. Specifically, the change in the mean attack latency from the first to third trial was significantly larger for *N. lecontei* larvae than for *D. pini* larvae (Table 1; Fig. 2b). By contrast, predators learned to avoid white and yellow *N. lecontei* larvae equally quickly (Table 1; Fig. 2b). Combined with our finding that both morphs are conspicuous to an avian predator against pine backgrounds (Fig. 2a), our predation assay results suggest that white and yellow larvae do not differ in signal efficacy. Finally, hunger level of birds did not affect significantly attack latencies (Table 1).

HISTORICAL DEMOGRAPHY WOULD NOT HAVE FACILITATED THE SPREAD OF A DELETERIOUS ALLELE, AND DIFFERENTIATION AT A SINGLE COLOR LOCUS IS ELEVATED BETWEEN WHITE-BODIED AND YELLOW-BODIED POPULATIONS

Both model-based and model-free analyses of discrete population structure recovered $K = 2$ as the optimal number of clusters (Fig. S5). However, assignment patterns under $K = 2$ corresponded to geographical regions, not color. Specifically, individuals west of the Appalachian Mountains ("West"; i.e., those in Kentucky) belonged to one cluster and individuals sampled from within or east of the Appalachian Mountains assigned to the other ("East") (Fig. 3a, b). Individual assignments were stable across all 100 ADMIXTURE runs, and similar for most individuals between the ADMIXTURE and DAPC methods (Fig. 3b). Thus, there is no discrete population structure indicative of a historical event that separated white-bodied and yellow-bodied populations.

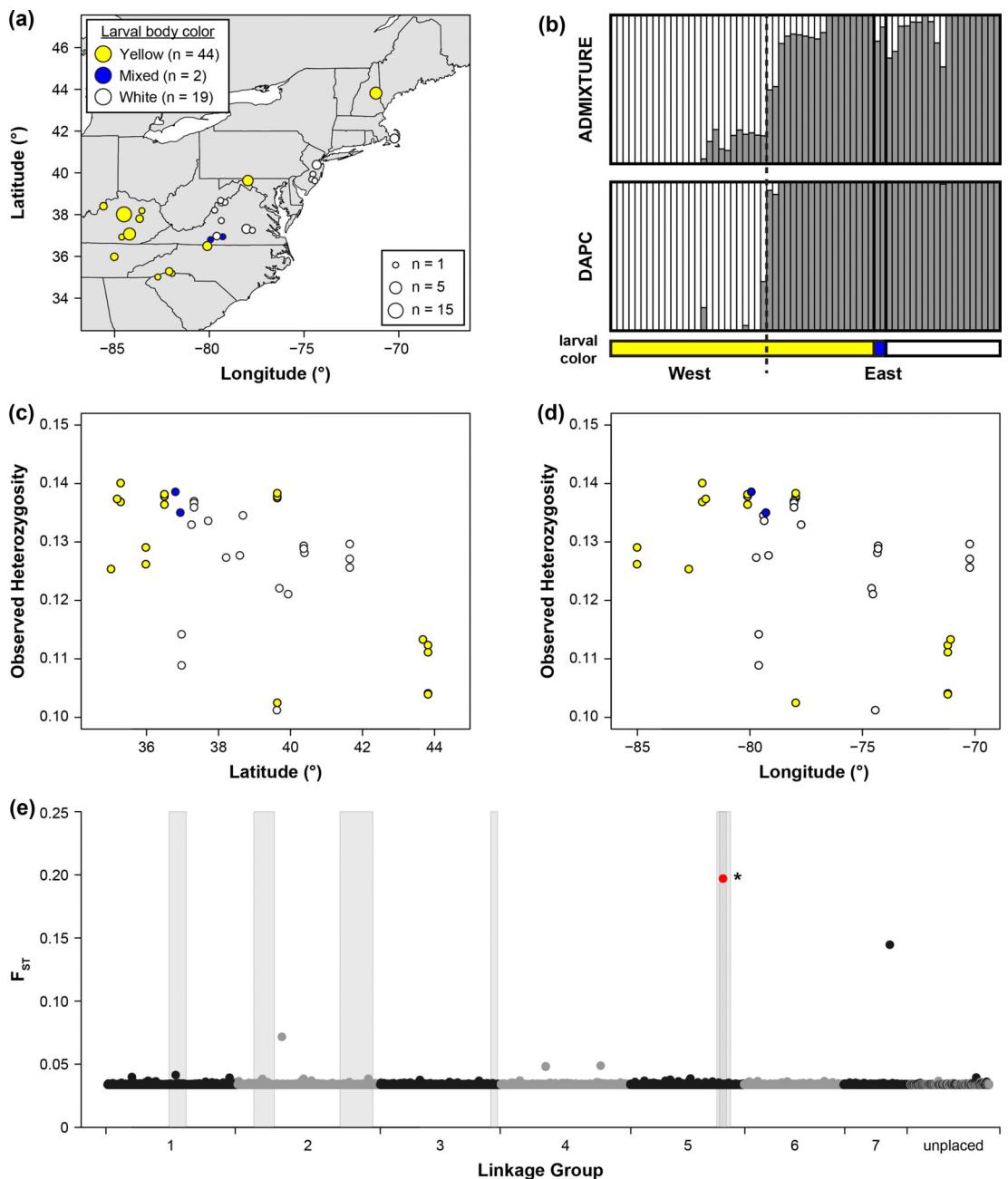


Figure 3. Genome-wide markers reveal no evidence of color-associated population structure in *Neodiprion lecontei*, but differentiation at a single color locus is elevated between white-bodied and yellow-bodied populations. (a) Color, location, and sample size for 65 individuals and 29 locations included in the population structure analyses. Points are colored based on the color of all larvae collected at that location, with size of the point reflecting number of individuals (colonies) sampled. (b) Inferred ancestry for each individual sorted by the color of the colony from which it was collected under a model of $K = 2$ based on the program *admixture* (top) and *aDPC* (bottom). Colors are denoted by the “larval color” bar, with colors indicated as in (a). Within each color, individuals are sorted in geographical order. The “West” genetic cluster contains only yellow larvae, whereas the “East” genetic cluster contains yellow, white, and mixed-color colonies. (c) Relationship between observed heterozygosity and latitude for “East” individuals with colony color indicated as in panel A. Heterozygosity declines with latitude, and the lowest-heterozygosity individuals are from yellow-bodied populations collected in the northernmost part of the range. (d) Relationship between observed heterozygosity and longitude for “East” individuals with colony color indicated as in panel A. Heterozygosity declines with longitude. (e) Per-site differentiation (F_{ST}) between white-bodied and yellow-bodied populations of *N. lecontei* (“East” populations only) across seven linkage groups and remaining unplaced scaffolds. Shaded gray boxes denote the locations of six QTL associated with larval body color in a cross between white-bodied and yellow-bodied populations (Linnen et al. 2018). The red point and asterisk indicate the only significant F_{ST} outlier, which falls within two major-effect QTL intervals.

Table 2. Correlation between a proxy for carotenoid concentration (S1B) and defensive behavior, quality and quantity of chemical defense against predators, and encapsulation response (defense against parasitoids). *t*-statistic for Gaussian error LMMs (B, C, and E) and *z*-statistic for GLMMs (A and D).

	Estimate \pm SE	<i>t</i> or <i>z</i>	<i>P</i>
A: Defensive fluid volume (<i>N</i> = 91)			
Intercept	0.603 \pm 0.251	2.40	0.0163
Body length	-0.072 \pm 0.094	-0.76	0.445
S1B	1.506 \pm 1.930	0.78	0.435
B: Monoterpene content of defensive fluid (per μ L of fluid) (<i>N</i> = 66)			
Intercept	0.074 \pm 0.024	6.20	<0.0001
S1B	0.084 \pm 0.097	0.86	0.391
C: Total terpene content of defensive fluid (per μ L of fluid) (<i>N</i> = 66)			
Intercept	0.118 \pm 0.017	7.14	<0.0001
S1B	0.177 \pm 0.130	1.36	0.180
D: Defensive behavior (deploys fluid or not) (<i>N</i> = 91)			
Intercept	1.94 \pm 0.45	4.34	<0.0001
S1B	3.99 \pm 4.62	0.864	0.388
E: Encapsulation response (<i>N</i> = 83)			
Intercept	51.21 \pm 8.71	5.88	<0.0001
S1B	-13.39 \pm 36.42	-0.368	0.714

In larvae collected from the “East” cluster, we found no relationship between heterozygosity and larval colony color (Kruskal-Wallis chi-squared = 3.17, *df* = 2, *P* = 0.20). Consistent with a post-glacial range expansion to the north and east from a southern refugium, we found a significant decline in heterozygosity with both latitude (Spearman’s rho = -0.51; *P* = 0.00084; Fig. 3c) and longitude (Spearman’s rho = -0.46; *P* = 0.0032; Fig. 3d). However, the low-heterozygosity populations at the northern edge of the range were yellow, not white (Fig. 3a, c). These data do not support an allele-surfing scenario for the white allele. Instead, these data are consistent with a scenario in which the white allele arose and spread sometime after the eastern range expansion was complete.

Our *F_{ST}* outlier analysis yielded a single SNP on chromosome 5 (position 21793198) with significantly elevated differentiation between white-bodied and yellow-bodied populations (*F_{ST}* = 0.197, *q*-value = 0.012; Fig. 3e). Notably, this SNP fell within the intervals of two QTL peaks previously identified for larval body color in *N. lecontei*, including one QTL that explains ~52% of the difference in color between white- and yellow-bodied populations (Linnen et al. 2018). This finding is consistent with the hypothesis that differences in pigmentation among *N. lecontei* populations are currently maintained by selection in the face of gene flow. We note, however, that our *F_{ST}* outlier results should be interpreted with caution because the demographic history of white-bodied and yellow-bodied populations likely deviates from the simple model assumed by BayeScan. Although this method is generally robust to more complex demographic scenarios (Foll and Gaggiotti 2008), additional data are needed

to evaluate the impact of violating BayeScan assumptions on the rate of false positives and false negatives and to apply alternative tests of selection.

LARVAL BODY COLOR DOES NOT CORRELATE WITH OTHER DEFENSIVE TRAITS

Our analysis of over 200 recombinant haploid males revealed that there was no correlation between body color (a proxy for carotenoid content) and any of the defensive traits (behavioral, chemical, immune) measured (Table 2). Together, these results do not support the hypothesis that the loss of yellow pigmentation was favored by natural selection because it enabled white-bodied larvae to allocate more resources to defense and immune functions (but see *Discussion*).

WHITE-BODIED LARVAL COLONIES ARE ASSOCIATED WITH LOW-CAROTENOID HOSTS

The proportion of colonies that were white bodied differed significantly across host plants (Fisher’s exact test, *P* < 1 \times 10⁻⁹), with white-bodied colonies significantly more common on *P. rigida* than on either *P. echinata* (*P* = 9.6 \times 10⁻⁶) or *P. virginiana* (*P* = 2.4 \times 10⁻⁹) (Fig. 4a). Prevalence of white-bodied colonies did not differ between *P. virginiana* and *P. echinata* (*P* = 0.66). We also found that the three hosts differed significantly in carotenoid content (*F₂* = 20.12, *P* = 0.0022), with *P. rigida* having the lowest carotenoid content of all three hosts (Fig. 4b; *P. rigida* vs. *P. echinata*: *z* = 6.108, *P* = 3.02 \times 10⁻⁹; *P. rigida* vs. *P. virginiana*: *z* = 4.538, *P* = 1.13 \times 10⁻⁵; *P. echinata* vs. *P. virginiana*: *z* = -1.570, *P* = 0.116). Notably, although *P.*

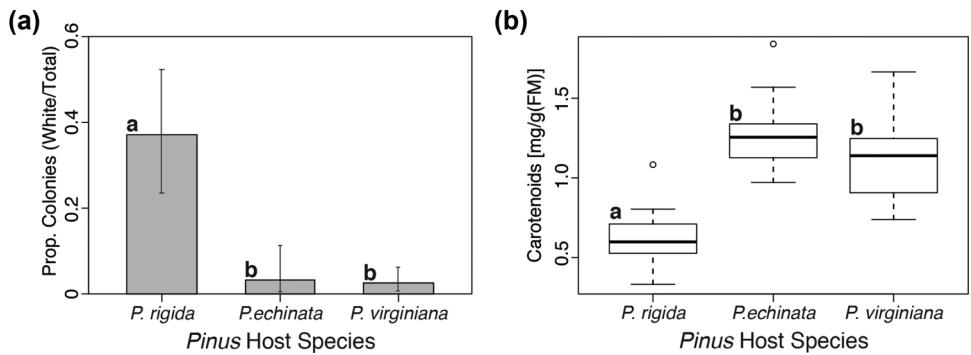


Figure 4. Relationship between host use and larval body color in *N. lecontei*. (a) Compared to *Pinus echinata* and *P. virginiana*, *P. rigida* had a significantly higher proportion of white-bodied larval colonies. Error bars are Clopper-Pearson 95% confidence intervals. (b) Compared to *P. echinata* and *P. virginiana*, *P. rigida* foliage had significantly lower carotenoid content. In all panels, letters indicate significant pairwise differences between colony color or host plants at $P < 0.05$.

rigida is widely distributed throughout the eastern United States, it is most abundant in the Atlantic Coastal plain (Gleason and Cronquist 1991; Critchfield and Little 1966), where white-bodied populations are predominantly found (Fig. 1c). Overall, these results support the hypothesis that an increased tendency to use a low-carotenoid host could have favored the spread of novel white-body alleles in the eastern United States.

Discussion

Here, we investigated the evolutionary processes and selective pressures shaping geographic variation in body color in conspicuously colored *N. lecontei* larvae. We first confirmed that what we consider to be aposematic coloration does, in fact, confer protection against visual predators of *N. lecontei* larvae. Therefore, it is possible that *N. lecontei* larval coloration evolves under positive-frequency-dependent selection, favoring familiar color morphs over novel color morphs (e.g., Kapan 2001; Chouteau et al. 2016). Indeed, as expected under purifying selection against novel warning color alleles, local populations of *N. lecontei* larvae tend to be monomorphic for color (Fig. 1c). However, in contrast to many studies on aposematic adult coloration (Summers et al. 1999; Jiggins et al. 2001), a lack of color-associated population structure (Fig. 3) suggests that the demographic context in which the white allele arose would not have been conducive to the spread of a deleterious allele. Additionally, our finding that differentiation at a single SNP linked to a previously identified color locus exceeded genome-wide levels of genetic differentiation between white- and yellow-bodied populations (Fig. 3) suggests that geographical differences in color among extant populations are attributable to natural selection rather than migration-drift balance. Together, these results suggest that for aposematic larvae, genetic drift need not be invoked to explain shifts in warning color.

Having determined that white-body alleles could have evolved under positive selection, we next examined sources of selection that could have potentially offset (or exacerbated) the cost of being rare. Perhaps unexpectedly, we did not find any evidence that the loss of yellow pigmentation impacts defense against natural enemies. Body color had no impact on the efficacy of the warning signal (Fig. 2b). We also found no evidence that white-body alleles have indirect effects on chemical defense or immune function via pleiotropy or physical linkage (Table 2). Instead, we found that white-bodied larvae were disproportionately abundant on a pine species with low carotenoid content (Fig. 4). Together, these findings suggest that bottom-up selection via host-plant quality may play an important role in driving warning color polytypism in immature stages that rely on diet-derived pigmentation. Overall, these results suggest that life stage and pigment source can greatly impact the evolutionary processes and selective pressures shaping warning color variation. To explore this idea further, we contrast our results with other aposematic study systems and highlight priorities for future work on this and other aposematic systems.

The lack of color-associated population structure in *N. lecontei* contrasts with several other aposematic study systems that implicate geographic isolation and genetic drift as key facilitators of polytypic warning signals. For example, in *Heliconius* butterflies, genetic drift is thought to promote the establishment of novel wing pattern variants in accordance with Wright's shifting balance model (Wright 1948; Mallet and Singer 1987; Mallet 2010) (but see Brown et al. 1974; Sheppard et al. 1985; Turner et al. 1996). In support of this hypothesis, multiple *Heliconius* species have pronounced population genetic structure indicative of historical subdivision and drift (Kronforst and Gilbert 2008). Color-associated population structure has also been detected in other aposematic species, such as the strawberry poison frog (*Oophaga pumilio*) (Wang and Summers 2010) and the polytypic

dyeing dart poison frog (*Dendrobates tinctorius*) (Lawrence et al. 2019). One important difference between *N. lecontei* and many other aposematic study systems is the role of warning coloration in mate choice. Although larval color variation in *Neodiprion* sawflies is unlikely to have any impact on mate preferences in the non-aposematic adults, positive assortative mating based on the color of aposematic adults has been demonstrated experimentally in both *Heliconius* (Jiggins et al. 2001; Naisbit et al. 2001; Merrill et al. 2014) and *O. pumilio* (Reynolds and Fitzpatrick 2007). The absence of this added barrier to evolving a novel color morph may explain, in part, why a shift in *N. lecontei* larval coloration seemingly did not require external factors to increase isolation and drift.

In addition to a lack of sexual selection, several other features of larval defensive displays in *N. lecontei* (and other taxa) have the potential to reduce the strength of purifying selection against novel warning color alleles, thereby facilitating their spread in the absence of isolation and genetic drift. First, *N. lecontei* larvae present their defensive chemicals externally (Codella and Raffa 1995; Costa 2006), a strategy that has been shown to reduce attack rates against novel conspicuous prey as predators can estimate a prey individual's defensive capacity before tasting it (Halpin et al. 2008). Second, both yellow and white *N. lecontei* larvae have melanin-based black spots across their body, which predators could use as an additional signal of unprofitability together with the bright pigmentation (but see Aronsson and Gamberale-Stille 2008). This similarity in one aspect of the aposematic pattern could have decreased initial costs of rarity for a white larval form, thereby facilitating the shift from yellow to white pigmentation via generalization (Balogh et al. 2010; Lawrence et al. 2019; but see Rönkä et al. 2018b). Third, and perhaps most importantly, the costs of being a low frequency color morph could be mitigated by gregariousness in chemically defended prey. *N. lecontei* larvae feed in large aggregations, which have been shown to increase avoidance learning efficacy and initial wariness of predators toward both conspicuous and cryptic unprofitable prey (Sillén-Tullberg 1990; Alatalo and Mappes 1996; Riipi et al. 2001). Gregariousness could have therefore facilitated shifts in warning coloration among *N. lecontei* populations. Additional experiments are needed to determine the interactive effects of predator's generalization, larval group size, larval defensive displays, and color. To assess whether predation risk for yellow- and white-bodied larvae varies among *N. lecontei* populations, future research should also test for differences in the survival of white- and yellow-bodied larvae across their geographical range and across different visual backgrounds (Rojas et al. 2014; Rönkä et al. 2020).

In terms of possible benefits that could have facilitated the spread of novel white-body alleles via positive selection, our results suggest that geographic variation in the abundance of a

low-carotenoid host (*P. rigida*) may have favored the loss of carotenoid-based larval coloration in some *N. lecontei* populations. Thus, rather than strong top-down selection by predators and pathogens, bottom-up selection via host-plant characteristics may have driven shifts in warning color among *N. lecontei* populations. This finding contrasts with strong top-down selection pressures that may explain endogenously produced polymorphic and polytypic warning color in *Arctia plantaginis* (Rönkä et al. 2020) and polytypic warning color in *Heliconius melpomene*. Historical demographic analyses suggest that *H. melpomene* radiated into areas already occupied by another, more abundant aposematic species, *H. erato* (e.g., Turner et al. 1996; Mallet and Joron 1999; Kronforst and Gilbert 2008; Quek et al. 2010). Although warning color polytypisms were thought to have evolved via drift in *H. erato*, later radiating *H. melpomene* are thought to have evolved under positive selection to match locally abundant *H. erato* color phenotypes (Kronforst and Gilbert 2008; Quek et al. 2010). Consistent with this "advergence" hypothesis, early-radiating *erato*-clade species exhibit much more pronounced population substructure than *melpomene*-clade species (Kronforst and Gilbert 2008). Like *N. lecontei*, *H. melpomene* exhibits geographic variation in host use (Smiley 1978; Brown 1981). But unlike *N. lecontei*, *Heliconius* butterflies primarily rely on endogenous pigments for warning coloration (reviewed in Nadeau 2016). Although these differences are intriguing, comparative analysis of color evolution in diverse taxa is needed to test our hypothesis that pigment source impacts the primary selective agents shaping pigment variation.

In the *N. lecontei* system, additional experimental work is also needed to clarify causal links between differences in host use and differences in fitness between *N. lecontei* color morphs. Our observation that *P. rigida*, a low-carotenoid host, harbors disproportionately more white-bodied larvae (Fig. 4) is consistent with the hypothesis that white alleles were favored due to carotenoid limitation. However, pine species differ in other characteristics besides carotenoid content that could potentially favor different carotenoid allocation strategies in the larvae. For example, differences in pine species color result in differences in larval luminance contrast values (Fig. S4). Nevertheless, larvae are conspicuous on all three hosts and therefore host color seems unlikely to drive color divergence among larval populations. *Pinus rigida* could also differ from the other pine species in nutritional profile, secondary compounds (terpene content), and/or abiotic environments in ways that increase the demand for antioxidant functions of carotenoid compounds (Blount et al. 2009). However, based on previous studies, *N. lecontei* seem relatively insensitive to variation in the host-plant terpene content (Codella and Raffa 1995).

If carotenoid limitation is the primary selection pressure favoring white body color, there are two nonmutually exclusive mechanisms through which host carotenoid content could differ-

entially impact the fitness of yellow and white *N. lecontei* larvae. First, production of conspicuous yellow coloration could be constrained under a carotenoid-limited diet, potentially weakening the signal efficacy (saturation and brightness) of yellow genotypes that develop on a low-carotenoid host. Second, because carotenoids are also involved in physiological processes such as immune defense (Cornet et al. 2007; Babin et al. 2010; Blount et al. 2012; Dhinaut et al. 2017; Koch et al. 2018), trade-offs between warning signaling and these other functions could result in divergent selection on pigmentation on different host plants. These trade-offs and constraints should be most evident when the availability of pigments or their precursors are limited in the diet (Lindstedt et al. 2020). However, such trade-offs should also generate negative genetic (and phenotypic) correlations between levels of carotenoid-based pigmentation and other fitness traits (especially immune or antioxidant function), which we did not find in this study (Table 2; see also Siiri-Lii et al. 2007). That said, an important limitation of our data is that we did not measure trait correlations (trade-offs) in larvae that developed on *P. rigida*. Because both diet-derived pigmentation and defensive compounds (Blount et al. 2012) as well as immune responses (Cotter et al. 2003; Lindstedt et al. 2020) are sensitive to the pool of resources allocated to these different functions, they can be more easily detected under limited resources (King et al. 2011)—in this case, on the *P. rigida* host. Thus, to rigorously test the carotenoid trade-off hypotheses due to host plant shifts, rearing experiments of white- and yellow-bodied larvae on alternative host-plant species are needed.

To conclude, our integrated analysis of color variation in *Neodiprion* sawflies provides a valuable point of comparison for other classic aposematic systems, such as *Heliconius* butterflies. Although both taxa exhibit striking among-population variation in color attributable to major-effect loci, our data indicate that the observed polytypisms likely evolved under very different evolutionary scenarios and selection pressures. We hypothesize that these differences ultimately stem from differences in the source of pigments and the role of color in mate choice. To evaluate these hypotheses more rigorously, comparable analyses of diverse aposematic taxa that vary in pigment sources (dietary vs. endogenous), life stages (immature vs. adult), and behavior (mobility and gregariousness in larval stage vs. adult stage) are needed (Willmott et al. 2011; Linnen et al. 2018; Galarza et al. 2019; Lindstedt et al. 2019). More generally, this kind of integrative study approach will give us a more comprehensive understanding of how organisms adapt to changes in their environment (Zaman et al. 2019) and how ecological selection during immature life stages impacts adaptation and speciation.

AUTHOR CONTRIBUTIONS

CRL and CL conceived the original idea. CRL, CL, RKB, and MJ designed methodology and collected the data. CRL, CL, RKB, and SC analyzed the data. CRL and CL led the writing of the manuscript. All authors contributed critically to the drafts and writing and gave final approval for publication.

ACKNOWLEDGMENTS

We are grateful to D. Westneat (University of Kentucky) and J. Mappes (University of Jyväskylä) for offering facilities, equipment, and technical help for the avian experiments and bird maintenance; J. Kirvesoja and H. Pakkanen for the help with chemical analyses; D. Herrig for custom python scripts; V. Sousa for helpful discussion; and members of the Linnen lab for rearing and collecting sawflies (especially K. Vertacnik and A. Leonberger). This study was funded by the National Science Foundation (DBI-1062890 to the University of Kentucky site-based Research Experience for Undergraduates summer program; DEB-1257739 to CRL; DEB-CAREER-1750946 to CRL) and the Academy of Finland via Centre of Excellence in Biological Interactions and individual grants (#257581 and #336336 to CL) to CL. SC was funded by an Academy of Finland Fellowship (#314219).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS STATEMENT FOR THE EXPERIMENTS WITH

BIRDS

Predation experiments with wild great tits at Konnevesi Research Station (Jyväskylä, Finland) were carried out with permission from the Central Finland Centre for Economic Development, Transport and Environment (KESELY/1017/07.01/2010) and license from the National Animal Experiment Board (ESAVI-2010-087517Ym-23).

DATA ARCHIVING

All the datasets used in this study are available as source data files named and linked for each figure and table accordingly. Statistical models with the full list of parameters used are defined in the methods section and for some of the analyses as Supporting Information (referred in text as Fig. Sx or Table Sx). R-codes and scripts (<https://doi.org/10.5281/zenodo.5772251>) included for the accepted final version of the manuscript together with the datasets (<https://doi.org/10.5061/dryad.cz8w9gj4w>) are openly available. Demultiplexed sequence files for all individuals will be available on the SRA (NCBI BioSample numbers: SAMN23893938-SAMN23894002).

LITERATURE CITED

Alatalo, R. V., and J. Mappes. 1996. Tracking the evolution of warning signals. *Nature* 382:708–710.
 Alexander, D. H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1655–1664.

Arias, M., Y. le Poul, M. Chouteau, R. Boisseau, N. Rosser, M. Théry, and V. Llaurens. 2016. Crossing fitness valleys: empirical estimation of a fitness landscape associated with polymorphic mimicry. *Proc. R. Soc. B Biol. Sci.* 283:20160391.

Aronsson, M., and G. Gamberale-Stille. 2008. Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Anim. Behav.* 75:417–423.

Aubier, T. G., and T. N. Sherratt. 2015. Diversity in Müllerian mimicry: the optimal predator sampling strategy explains both local and regional polymorphism in prey. *Evolution* 69:2831–2845.

Babin, A., C. Biard, and Y. Moret. 2010. Dietary supplementation with carotenoids improves immunity without increasing its cost in a crustacean. *Am. Nat.* 176:234–241.

Bagley, R. K., V. C. Sousa, M. L. Niemiller, and C. R. Linnen. 2017. History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (*Neodiprion lecontei*). *Mol. Ecol.* 26:1022–1044.

Balogh, A. C. V., G. Gamberale-Stille, B. S. Tullberg, and O. Leimar. 2010. Feature theory and the two-step hypothesis of Müllerian mimicry evolution. *Evolution* 64:810–822.

Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker. 2014. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1–48.

Bezzerides, A. L., K. J. McGraw, R. S. Parker, and J. Husseini. 2007. Elytra color as a signal of chemical defense in the Asian ladybird beetle *Harmonia axyridis*. *Behav. Ecol. Sociobiol.* 61:1401–1408.

Blount, J. D., M. P. Speed, G. D. Ruxton, and P. A. Stephens. 2009. Warning displays may function as honest signals of toxicity. *Proc. R. Soc. B Biol. Sci.* 276:871–877.

Blount, J. D., H. M. Rowland, F. P. Drijfhout, J. A. Endler, R. Inger, J. J. Sloggett, G. D. D. Hurst, D. J. Hodgson, and M. P. Speed. 2012. How the ladybird got its spots: effects of resource limitation on the honesty of aposematic signals. *Funct. Ecol.* 26:334–342.

Briolat, E. S., E. R. Burdfield-Steel, S. C. Paul, K. H. Rönkä, B. M. Seymour, T. Stankowich, and A. M. M. Stuckert. 2018. Diversity in warning coloration: selective paradox or the norm? *Biol. Rev.* 94:388–414.

Brown, K. S. 1981. The biology of *Heliconius* and related genera. *Annu. Rev. Entomol.* 26:427–457.

Brown, K. S., P. M. Sheppard, and J. R. G. Turner. 1974. Quaternary refugia in tropical America: evidence from race formation in *Heliconius* butterflies. *Proc. R. Soc. Lond. B. Biol. Sci.* 187:369–378.

Burford Reiskind, M. O., K. Coyle, H. V. Daniels, P. Labadie, M. H. Reiskind, N. B. Roberts, R. B. Roberts, J. Schaff, and E. L. Vargo. 2016. Development of a universal double-digest RAD sequencing approach for a group of nonmodel, ecologically and economically important insect and fish taxa. *Mol. Ecol. Resour.* 16:1303–1314.

Butler, M. W., M. B. Toomey, and K. J. McGraw. 2011. How many color metrics do we need? Evaluating how different color-scoring procedures explain carotenoid pigment content in avian bare-part and plumage ornaments. *Behav. Ecol. Sociobiol.* 65:401–413.

Carroll, M., A. Hanlon, T. Hanlon, A. R. Zangerl, and M. R. Berenbaum. 1997. Behavioral effects of carotenoid sequestration by the parsnip webworm *Depressaria pastinacella*. *J. Chem. Ecol.* 23:2707–2719.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22:3124–3140.

Cheverud, J. M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42:958–968.

Chouteau, M., M. Arias, and M. Joron. 2016. Warning signals are under positive frequency-dependent selection in nature. *Proc. Natl. Acad. Sci. USA* 113:2164–2169.

Codella, S., and K. Raffa. 1995. Host plant influence on chemical defense in conifer sawflies (Hymenoptera: Diprionidae). *Oecologia* 104:1–11.

Codella, S. G., and K. F. Raffa. 1996. Individual and social components of wood ant response to conifer sawfly defence (Hymenoptera: Formicidae, Diprionidae). *Anim. Behav.* 52:801–811.

Cornet, S., C. Biard, and Y. Moret. 2007. Is there a role for antioxidant carotenoids in limiting self-harming immune response in invertebrates? *Biol. Lett.* 3:284–288.

Costa, J. 2006. Other insect societies. The Belknap Press of Harvard Univ. Press, Cambridge, MA.

Cotter, S. C., L. E. B. Kruuk, and K. Wilson. 2003. Costs of resistance: genetic correlations and potential trade-offs in an insect immune system. *J. Evol. Biol.* 17:421–429.

Crutchfield, W. B., and E. L. Little. 1966. Geographic distribution of the pines of the world. U.S. Department of Agriculture, Forest Service, Washington, D.C.

Cuthill, I. C., W. L. Allen, K. Arbuckle, B. Caspers, G. Chaplin, M. E. Hauber, G. E. Hill, N. G. Jablonski, C. D. Jiggins, A. Kelber, et al. 2017. The biology of color. *Science* 357:eaan0221.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.

Dasmahapatra, K. K., J. R. Walters, A. D. Briscoe, J. W. Davey, A. Whibley, N. J. Nadeau, A. V. Zimin, D. S. T. Hughes, L. C. Ferguson, S. H. Martin, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98.

Dhinaut, J., A. Balourdet, M. Teixeira, M. Chogne, and Y. Moret. 2017. A dietary carotenoid reduces immunopathology and enhances longevity through an immune depressive effect in an insect model. *Sci. Rep.* 7:12429.

Dochtermann, N. A. 2011. Testing Cheverud's conjecture for behavioral correlations and behavioral syndromes. *Evolution* 65:1814–1820.

Edmonds, C. A., A. S. Lillie, and L. L. Cavalli-Sforza. 2004. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* 101:975–979.

Eisner, T., J. S. Johnesee, J. Carrel, L. B. Hendry, and J. Meinwald. 1974. Defensive use by an insect of a plant resin. *Science* 184:996–999.

Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180:977–993.

Friman, V.-P., C. Lindstedt, T. Hiltunen, J. Laakso, and J. Mappes. 2009. Predation on multiple trophic levels shapes the evolution of pathogen virulence. *PLoS ONE* 4:e6761.

Gaitonde, N., J. Joshi, and K. Kunte. 2018. Evolution of ontogenetic change in color defenses of swallowtail butterflies. *Ecol. Evol.* 8:9751–9763.

Galarza, J. A., O. Nokelainen, R. Ashrafi, R. H. Hegna, and J. Mappes. 2014. Temporal relationship between genetic and warning signal variation in the aposematic wood tiger moth (*Parasemia plantaginis*). *Mol. Ecol.* 23:4939–4957.

Galarza, J. A., K. Dhaygude, B. Ghaedi, K. Suisto, J. Valkonen, and J. Mappes. 2019. Evaluating responses to temperature during pre-metamorphosis and carry-over effects at post-metamorphosis in the wood tiger moth (*Arctia plantaginis*). *Philos. Trans. R. Soc. B Biol. Sci.* 374:20190295.

Gleason, H. A., and A. Cronquist. 1991. Manual of vascular plants of northeastern United States and adjacent Canada. 2nd ed. New York Botanical Garden, New York.

Gordon, S. P., H. Kokko, B. Rojas, O. Nokelainen, and J. Mappes. 2015. Colour polymorphism torn apart by opposing positive frequency-dependent selection, yet maintained in space. *J. Anim. Ecol.* 84:1555–1564.

Grill, C. P., and A. J. Moore. 1998. Effects of a larval antipredator response and larval diet on adult phenotype in an aposematic ladybird beetle. *Oecologia* 114:274–282.

Hadfield, J. D., A. Nutall, D. Osorio, and I. P. F. Owens. 2007. Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *J. Evol. Biol.* 20:549–557.

Halpin, C. G., J. Skelhorn, and C. Rowe. 2008. Naïve predators and selection for rare conspicuous defended prey: the initial evolution of aposematism revisited. *Anim. Behav.* 75:771–781.

Hegna, R. H., O. Nokelainen, J. R. Hegna, and J. Mappes. 2013. To quiver or to shiver: increased melanization benefits thermoregulation, but reduces warning signal efficacy in the wood tiger moth. *Proc. R. Soc. B Biol. Sci.* 280:20122812.

Hegna, R. H., J. A. Galarza, and J. Mappes. 2015. Global phylogeography and geographical variation in warning coloration of the wood tiger moth (*Parasemia plantaginis*). *J. Biogeogr.* 42:1469–1481.

Herrig, D. K., K. L. Vertacnik, A. R. Kohrs, and C. R. Linnen. 2021. Support for the adaptive decoupling hypothesis from whole-transcriptome profiles of a hypermetamorphic and sexually dimorphic insect, *Neodiprion lecontei*. *Mol. Ecol.* 00:1–16.

Jiggins, C. D., R. E. Naisbit, R. L. Coe, and J. Mallet. 2001. Reproductive isolation caused by colour pattern mimicry. *Nature* 411:302–305.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94.

Kapan, D. D. 2001. Three-butterfly system provides a field test of Müllerian mimicry. *Nature* 409:18–20.

Kenny, E. M., P. Cormican, W. P. Gilks, A. S. Gates, C. T. O'Dushlaine, C. Pinto, A. P. Corvin, M. Gill, and D. W. Morris. 2010. Multiplex target enrichment using DNA indexing for ultra-high throughput SNP detection. *DNA Res.* 18:31–38.

King, E. G., D. A. Roff, and D. J. Fairbairn. 2011. The evolutionary genetics of acquisition and allocation in the wing dimorphic cricket, *Gryllus firmus*. *Evolution* 65:2273–2285.

Koch, R. E., A. N. Kavazis, D. Hasselquist, W. R. Hood, Y. Zhang, M. B. Toomey, and G. E. Hill. 2018. No evidence that carotenoid pigments boost either immune or antioxidant defenses in a songbird. *Nat. Commun.* 9:491.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15:1179–1191.

Kronforst, M. R., and L. E. Gilbert. 2008. The population genetics of mimetic diversity in *Heliconius* butterflies. *Proc. R. Soc. B Biol. Sci.* 275:493–500.

Kruuk, L. E. B., J. Slate, and A. J. Wilson. 2008. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annu. Rev. Ecol. Evol. Syst.* 39:525–548.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357–359.

Lawrence, J. P., B. Rojas, A. Fouquet, J. Mappes, A. Blanchette, R. A. Saporito, R. J. Bosque, E. A. Courtois, and B. P. Noonan. 2019. Weak warning signals can persist in the absence of gene flow. *Proc. Natl. Acad. Sci. USA* 116:19037–19045.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.

Lichtenthaler, H. K. 1987. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. Pp. 350–382 in L. Packer and R. Douce, eds. *Plant cell membranes*. Academic Press, New York.

Lindstedt, C., L. Lindström, and J. Mappes. 2008. Hairiness and warning colours as components of antipredator defence: additive or interactive benefits? *Anim. Behav.* 75:1703–1713.

—. 2009. Thermoregulation constrains effective warning signal expression. *Evolution* 63:469–478.

Lindstedt, C., J. H. R. Talsma, E. Ihalainen, L. Lindström, and J. Mappes. 2010. Diet quality affects warning coloration indirectly: excretion costs in a generalist herbivore. *Evolution* 64:68–78.

Lindstedt, C., H. Eager, E. Ihalainen, A. Kahilainen, M. Stevens, and J. Mappes. 2011a. Direction and strength of selection by predators for the color of the aposematic wood tiger moth. *Behav. Ecol.* 22:580–587.

Lindstedt, C., H. Huttunen, M. Kakkio, and J. Mappes. 2011b. Disentangling the evolution of weak warning signals: high detection risk and low production costs of chemical defences in gregarious pine sawfly larvae. *Evol. Ecol.* 25:1029–1046.

Lindstedt, C., E. Schroderus, L. Lindström, T. Mappes, and J. Mappes. 2016. Evolutionary constraints of warning signals: a genetic trade-off between the efficacy of larval and adult warning coloration can maintain variation in signal expression. *Evolution* 70:2562–2572.

Lindstedt, C., A. Miettinen, D. Freitak, T. Ketola, A. López-Sepulcre, E. Mäntylä, and H. Pakkanen. 2018. Ecological conditions alter cooperative behaviour and its costs in a chemically defended sawfly. *Proc. R. Soc. B Biol. Sci.* 285:20180466.

Lindstedt, C., L. Murphy, and J. Mappes. 2019. Antipredator strategies of pupae: how to avoid predation in an immobile life stage? *Philos. Trans. R. Soc. B Biol. Sci.* 374:20190069.

Lindstedt, C., K. Suisto, and J. Mappes. 2020. Appearance before performance? Nutritional constraints on life-history traits, but not warning signal expression in aposematic moths. *J. Anim. Ecol.* 89:494–505.

Linnen, C. R., C. T. O'Quin, T. Shackleford, C. R. Sears, and C. Lindstedt. 2018. Genetic basis of body color and spotting pattern in redheaded pine sawfly larvae (*Neodiprion lecontei*). *Genetics* 209:291–305.

Lischer, H. E. L., and L. Excoffier. 2011. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28:298–299.

Mallet, J. 2010. Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. *Ecol. Entomol.* 35:90–104.

Mallet, J., and M. Joron. 1999. Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. *Annu. Rev. Ecol. Syst.* 30:201–233.

Mallet, J., and M. C. Singer. 1987. Individual selection, kin selection, and the shifting balance in the evolution of warning colours: the evidence from butterflies. *Evolution* 32:337–350.

Mappes, J., and R. V. Alatalo. 1997. Effects of novelty and gregariousness in survival of aposematic prey. *Behav. Ecol.* 8:174–177.

Mappes, J., J. Tuomi, and R. V. Alatalo. 1999. Do palatable prey benefit from aposematic neighbors? *Ecoscience* 6:159–162.

Medina, I., R. Vega-Trejo, T. Wallenius, M. R. E. Symonds, and D. Stuart-Fox. 2020. From cryptic to colorful: evolutionary decoupling of larval and adult color in butterflies. *Evol. Lett.* 4:34–43.

Merrill, R. M., A. Chia, and N. J. Nadeau. 2014. Divergent warning patterns contribute to assortative mating between incipient *Heliconius* species. *Ecol. Evol.* 4:911–917.

Merrill, R. M., K. K. Dasmahapatra, J. W. Davey, D. D. Dell'Aglio, J. J. Hanly, B. Huber, C. D. Jiggins, M. Joron, K. M. Kozak, V. Llaurens, et al. 2015. The diversification of *Heliconius* butterflies: what have we learned in 150 years? *J. Evol. Biol.* 28:1417–1438.

Minocha, R., G. Martinez, B. Lyons, and S. Long. 2009. Development of a standardized methodology for quantifying total chlorophyll and carotenoids from foliage of hardwood and conifer tree species. *Can. J. For. Res.* 39:849–861.

Montgomerie, R. 2006. *Analyzing colors*. Harvard Univ. Press, Cambridge, MA.

—. 2008. CLR, version 1.05. Queen's University, Kingston, Canada. Available via <http://post.queensu.ca/mont/color/analyze.html>.

Nadeau, N. J. 2016. Genes controlling mimetic colour pattern variation in butterflies. *Curr. Opin. Insect Sci.* 17:24–31.

Nadeau, N. J., C. Pardo-Diaz, A. Whibley, M. A. Supple, S. V. Saenko, R. W. R. Wallbank, G. C. Wu, L. Maroja, L. Ferguson, J. J. Hanly, et al. 2016. The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* 534:106–110.

Naisbit, R. E., C. D. Jiggins, and J. Mallet. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proc. R. Soc. B Biol. Sci.* 268:1849–1854.

Nokelainen, O., R. H. Hegna, J. H. Reudler, C. Lindstedt, and J. Mappes. 2012. Trade-off between warning signal efficacy and mating success in the wood tiger moth. *Proc. R. Soc. B Biol. Sci.* 279:257–265.

Nosil, P., R. Villoutreix, C. F. de Carvalho, T. E. Farkas, V. Soria-Carrasco, J. L. Feder, B. J. Crespi, and Z. Gompert. 2018. Natural selection and the predictability of evolution in *Timema* stick insects. *Science* 359:765–770.

Osorio, D., and M. Vorobyev. 2005. Photoreceptor spectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. *Proc. R. Soc. B Biol. Sci.* 272:1745–1752.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. On the accumulation of deleterious mutations during range expansions. *Mol. Ecol.* 22:5972–5982.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7:e37135.

Pinheiro, J., D. Bates, S. DebRoy, and D. Sarkar. 2020. *nlme*: linear and non-linear mixed effects models. R Foundation for Statistical Computing, Vienna.

Quek, S.-P., B. A. Counterman, P. de Moura, M. Z. Cardoso, C. R. Marshall, W. O. McMillan, and M. R. Kronforst. 2010. Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. *Proc. Natl. Acad. Sci. USA* 107:7365–7370.

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rantala, M. J., J. Koskimaki, J. Taskinen, K. Tynkkynen, and J. Suhonen. 2000. Immunocompetence, developmental stability and wingspot size in the damselfly *Calopteryx splendens* L. *Proc. R. Soc. B Biol. Sci.* 267:2453–2457.

Reynolds, R. G., and B. M. Fitzpatrick. 2007. Assortative mating in poison-dart frogs based on an ecologically important trait. *Evolution* 61:2253–2259.

Riipi, M., R. V. Alatalo, L. Lindström, and J. Mappes. 2001. Multiple benefits of gregariousness cover detectability costs in aposematic aggregations. *Nature* 413:512–514.

Roff, D. A. 1996. The evolution of genetic correlations: an analysis of patterns. *Evolution* 50:1392–1403.

Rojas, B., P. Rautiala, and J. Mappes. 2014. Differential detectability of polymorphic warning signals under varying light environments. *Behav. Processes* 109:164–172.

Rönkä, K., C. De Pasqual, J. Mappes, S. Gordon, and B. Rojas. 2018a. Colour alone matters: no predator generalization among morphs of an aposematic moth. *Anim. Behav.* 135:153–163.

Rönkä, K., J. Mappes, C. Michalis, R. Kiviö, J. Salokannas, and B. Rojas. 2018b. Can multiple-model mimicry explain warning signal polymorphism in the wood tiger moth, *Arctia plantaginis* (Lepidoptera: Erebidae)? *Biol. J. Linn. Soc.* 124:237–260.

Rönkä, K., J. K. Valkonen, O. Nokelainen, B. Rojas, S. Gordon, E. Burdfield-Steel, and J. Mappes. 2020. Geographic mosaic of selection by avian predators on hindwing warning colour in a polymorphic aposematic moth. *Ecol. Lett.* 23:1654–1663.

Rowland, H. M., E. Ihlainen, L. Lindström, J. Mappes, and M. P. Speed. 2007. Co-mimics have a mutualistic relationship despite unequal defences. *Nature* 448:64–67.

Ruxton, G. D., T. N. Sherratt, and M. P. Speed. 2004. *Avoiding attack: the evolutionary ecology of crypsis, warning signals and mimicry*. Oxford Univ. Press, Oxford U.K.

Sandre, S.-L., M. Stevens, and J. Mappes. 2010. The effect of predator appetite, prey warning coloration and luminance on predator foraging decisions. *Behaviour* 147:1121–1143.

Sheppard, P. M., J. R. G. Turner, K. S. Brown, W. W. Benson, M. C. Singer, and J. M. Smith. 1985. Genetics and the evolution of müllerian mimicry in *heliconius* butterflies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308:433–610.

Siiri-Lii, S., T. Toomas, E. Toomas, J. Riitta, and M. Johanna. 2007. Carotenoid-based colour polyphenism in a moth species: search for fitness correlates. *Entomol. Exp. Appl.* 124:269–277.

Siddiqi, A., T. W. Cronin, E. R. Loew, M. Vorobyev, and K. Summers. 2004. Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *J. Exp. Biol.* 207:2471–2485.

Sillén-Tullberg, B. 1990. Do predators avoid groups of aposematic prey? An experimental test. *Anim. Behav.* 40:856–860.

—. 1988. Evolution of gregariousness in aposematic butterfly larvae: a phylogenetic analysis. *Evolution* 42:293–305.

Smiley, J. 1978. Plant chemistry and the evolution of host specificity: new evidence from *Heliconius* and *Passiflora*. *Science* 201:745–747.

Summers, K., R. Symula, M. Clough, and T. Cronin. 1999. Visual mate choice in poison frogs. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 266:2141–2145.

Talloen, W., H. Van Dyck, and L. Lens. 2004. The cost of melanization: butterfly wing coloration under environmental stress. *Evolution* 58:360–366.

Terbot, J. W., R. L. Gaynor, and C. R. Linnen. 2017. Gregariousness does not vary with geography, developmental stage, or group relatedness in feeding redheaded pine sawfly larvae. *Ecol. Evol.* 7:3689–3702.

Thomas, R. J., L. A. Bartlett, N. M. Marples, D. J. Kelly, and I. C. Cuthill. 2004. Prey selection by wild birds can allow novel and conspicuous colour morphs to spread in prey populations. *Oikos* 106:285–294.

Turner, J. R. G., J. L. B. Mallet, B. C. Clarke, and P. R. Grant. 1996. Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting balance. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 351:835–845.

Vertanick, K., S. Geib, and C. Linnen. 2016. *Neodiprion lecontei* genome assembly Nlec1.0. NCBI/GenBank. <https://doi.org/10.15482/USDA-ADC/1235565>.

Vorobyev, M., and D. Osorio. 1998. Receptor noise as a determinant of colour thresholds. *Proc. R. Soc. B Biol. Sci.* 265:351–358.

Vorobyev, M., D. Osorio, A. T. D. Bennett, N. J. Marshall, and I. C. Cuthill. 1998. Tetrachromacy, oil droplets and bird plumage colours. *J. Comp. Physiol. A* 183:621–633.

Wang, I. A. N. J., and K. Summers. 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Mol. Ecol.* 19:447–458.

Wang, L., S. J. Cornell, M. P. Speed, and K. Arbuckle. 2021. Coevolution of group-living and aposematism in caterpillars: warning colouration may

facilitate the evolution from group-living to solitary habits. *BMC Ecol. Evol.* 21:25.

Willis, J. H., J. A. Coyne, and M. Kirkpatrick. 1991. Can one predict the evolution of quantitative characters without genetics? *Evolution* 45:441–444.

Willmott, K., A. Sourakov, and M. Elias. 2011. Two possible caterpillar mimicry complexes in neotropical danaine butterflies (Nymphalidae, Danainae: Danaini, Ithomiini). *Ann. Entomol. Soc. Am.* 104:1108–1118.

Winters, A. E., N. F. Green, N. G. Wilson, M. J. How, M. J. Garson, N. J. Marshall, K. L. Cheney, and A. E. Winters. 2017. Stabilizing selection on individual pattern elements of aposematic signals. *Proc. R. Soc. B Biol. Sci.* 284:20170926.

Wright, S. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. *Evolution* 2:279–294.

Zaman, K., M. K. Hubert, and S. D. Schoville. 2019. Testing the role of ecological selection on colour pattern variation in the butterfly *Parnassius clodius*. *Mol. Ecol.* 28:5086–5102.

Associate Editor: Prof. Thomas Flatt

Handling Editor: Prof. Tracey Chapman

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Avoidance learning rates of birds towards white and yellow *Neodiprion lecontei* larvae were compared with light green-black *Diprion pini* larvae with the similar defensive compounds (Photo Carita Lindstedt).

Figure S2. α score plot for DAPC.

Figure S3. Variance explained across PCs.

Figure S4. Luminance contrasts against different host plants and host plant parts.

Figure S5. Cluster number determined by model and non-model based methods.

Table S1. Collection location, collection host, collection date, body color, and genetic cluster for all *Neodiprion lecontei* collected between 2001 and 2016.

Table S2. Human-sorted white and yellow *Neodiprion lecontei* larvae reflected underlying differences in color. Similarly color of frozen white and yellow *N. lecontei* larvae recapitulate differences between fresh larvae.

Table S3. Sampling locations for all individuals included in population genomic analyses.

Table S4 – Sequences for adapters containing variable-length barcodes from Burford Reiskind *et al.* (2016).

Table S5 – Sequences for PCR primers, including degenerate bases.

Table S6. Per-individual sequence statistic summary.

Table S7. Random and fixed effects from the linear mixed model on *Neodiprion lecontei* and host plant (*P. virginiana* [VA], *P. echinata* [SL], and *P. rigida*) color traits.