Title: Genetic architectures of larval pigmentation and color pattern in the redheaded pine sawfly (*Neodiprion lecontei*)

Authors: Catherine R. Linnen[*], Claire T. O'Quin[*], Taylor Shackleford[*], Connor R. Sears[*†], Carita Lindstedt[‡]

Running title: Genetic basis of pigmentation and pattern

Keywords: larval color, carotenoids, melanin, genetic complexity, evolutionary genetics

Corresponding author: Catherine R. Linnen, 204E Thomas Hunt Morgan Building, Lexington, KY 40506, 859-323-3160, Catherine.linnen@uky.edu

[*] Department of Biology, University of Kentucky, Lexington, KY, 40506
[†] Current affiliation: Department of Biological Sciences, University of Cincinnati, Cincinnati, OH, 45221
[‡] Centre of Excellence in Biological Interactions, Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland, FI-40014

ABSTRACT

Evolutionary biologists have long debated the contribution of large-effect mutations to phenotypic evolution. Although theoretical work suggests that developmental, demographic, and ecological contexts can have profound and predictable impacts on trait genetic architectures, there are few empirical tests of these predictions. Here, we test the hypothesis that, due to differences in pleiotropy in the underlying genetic pathways, the genetic architecture of overall body color (pigmentation) is predictably simpler (*i.e.*, large-effect mutations explain proportionally more of the phenotypic variance) than that of color patterning. To test this prediction, we crossed divergent populations of the redheaded pine sawfly (*Neodiprion lecontei*) that differ in larval body color and melanic spotting pattern and measured these traits in their recombinant haploid male progeny. Using a combination of interval mapping and polygenic association mapping, we identified large-effect QTL for both traits. Consistent with the pigmentation/patterning hypothesis, we found that compared to spotting pattern, body color had a larger percentage of genetic variance attributable to large-effect loci. Additionally, by combining mapping results with a linkage group-anchored genome assembly for *N. lecontei*, we identified several promising candidate genes for both carotenoid-based yellow pigmentation and melanin-based spotting pattern. Because few studies have investigated the genetic basis of naturally occurring variation in larval color and carotenoid-based pigmentation, our study helps fill a void in the invertebrate pigmentation literature. Finally, we argue that when leveraged to test explicit *a priori* hypotheses regarding trait genetic architectures, polygenic association mapping has the potential to shed new light on the 150-year-old micromutationist-macromutationist debate.

INTRODUCTION

One of the longest running debates in evolutionary biology—tracing its roots back to disagreements between Darwin (1859) and Huxley (1860)—centers on the contribution of large-effect mutations to evolutionary change (Mayr 1982; Orr and Coyne 1992). At one extreme, "micromutationists" (*e.g.*, Darwin 1859; Pearson 1897; Fisher 1930) have argued that adaptation results from the accumulation of many alleles of individually small effect. At the other extreme, "macromutationists" (*e.g.*, Huxley 1860; Bateson 1913; Morgan 1932; Goldschmidt 1940) have emphasized the role of a small number of large-effect mutations as the primary drivers of evolutionary change. As often happens when there is a debate between two conceptual extremes, empirical data on trait genetic architectures point to a continuum rather than a strict dichotomy (Mackay *et al.* 2009; Rockman 2012; Remington 2015; Dittmar *et al.* 2016). These data indicate that it is time to move beyond extreme caricatures of the evolutionary process. Instead, with the development of novel theory (Orr 2005; Dittmar *et al.* 2016) and powerful new tools for linking genotype to phenotype in non-model organisms (Davey *et al.* 2011; Gaj *et al.* 2013; Goodwin *et al.* 2016; Huang *et al.* 2016), we can start to make and test explicit predictions about where on the micromutationist-macromutationist continuum different traits, organisms, and evolutionary scenarios will fall.

Evolutionary theory predicts that the developmental pathways that give rise to a particular trait, the type of selection acting on that trait, and the demography of the evolving population can all have profound impacts on the expected contribution of large-effect mutations to adaptation (Remington 2015; Dittmar *et al.* 2016). Collectively, theoretical work to date suggests that large-effect mutations are more likely to contribute to adaptation when: pleiotropy is minimal (Fisher 1930; but see Matuszewski *et al.* 2014), the effective population size is small (Kimura 1983), the population is adapting to a rapidly moving fitness optimum (Matuszewski *et al.* 2014), the fitness landscape has multiple optima (Orr 1998; Matuszewski *et al.* 2015), the population is far from the phenotypic optimum (Orr 1998), migration occurs between locally adapted populations (Griswold 2006; Yeaman and Whitlock 2011), and/or adaptation proceeds via new mutations (Hermisson and Pennings 2005; Matuszewski *et al.* 2015). While additional theoretical work is needed to more fully explore possible evolutionary scenarios, there is also a dearth of empirical tests of existing theoretical predictions (but see Baxter *et al.* 2009; Rogers *et al.* 2012; Martin *et al.* 2017). Given the many factors that influence trait genetic architectures, rigorous tests of these predictions require the integrated study of the genetic basis of trait variation, the evolutionary history of populations, and the selective pressures shaping trait variation.

As one of the best studied morphological characteristics in nature—both from an ecological and a genetic perspective—color features prominently in the adaptation genetics literature (True 2003; Protas and Patel 2008; Wittkopp and Beldade 2009; Manceau *et al.* 2010; Nadeau and Jiggins 2010; Kronforst *et al.* 2012). For many organisms, overall color is determined by two different types of color traits: (1) the type (and amount) of pigment synthesized ("pigmentation") and (2) the distribution of pigmentation across the body ("color pattern") (Manceau *et al.* 2010). The abundance of discrete pigmentation phenotypes in nature, coupled with the identification of many large-effect pigmentation mutations via candidate gene analysis, have led some to argue that the genetic architecture of pigmentation traits—but not color pattern traits—is atypically simple (Rockman 2012). This simplicity is thought to arise as a consequence of relatively small genetic pathways with at least some minimally pleiotropic genes that would be permissive of large-effect pigmentation mutations (Rockman 2012).

Testing the hypothesis that the genetic basis of pigmentation is predictably "simpler" (*i.e.*, more likely to fall on the macro- end of the micromutationist-macromutationist continuum) than that of color pattern will require describing the genetic architecture of both types of traits in many different organisms. Although some relevant data exist (*e.g.*, Martin and Orgogozo 2013), experimental and publication biases make it difficult to draw strong conclusions. For example, good *a priori* candidates and a focus on discrete pigmentation phenotypes that are easy to score may have biased identified pigmentation loci towards those of large effect (Kopp 2009; Manceau *et al.* 2010; Rockman 2012). Thus, to better understand consistent differences that may exist between pigmentation and color pattern traits, more unbiased, genome-wide analyses of continuously varying color traits are needed (*e.g.*, Signor *et al.* 2016).

With these considerations in mind, pine sawflies in the genus *Neodiprion* provide an excellent empirical system for systematically investigating the genetic architecture of different color traits. First, there is extensive variation in larval pigmentation and larval color pattern both within and between species (Figures 1-2). Second, there is information

available on evolutionary relationships between species and, for some species, demographic histories of populations within species (Linnen and Farrell 2007, 2008a; b; Bagley *et al*. 2017). Third, it is possible to rear and cross many different *Neodipirion* species in the lab, making genetic mapping approaches tractable (Knerer and Atwood 1972, 1973; Kraemer and Coppel 1983; Bendall *et al.* 2017). Fourth, a growing list of genomic resources for *Neodiprion*—including an annotated genome and methylome for *N. lecontei* (Vertacnik *et al.* 2016; Glastad *et al.* 2017)—will facilitate fine-mapping and identification of causal genes and mutations. And finally, we have some understanding of the ecological function of color variation in pine sawflies and, more generally, in insects, which we review briefly before returning to the topic of trait genetic architecture.

Under natural conditions, pine sawfly larvae are attacked by a diverse assemblage of arthropod and vertebrate predators, by a large community of parasitoid wasps and flies, and by fungal, bacterial, and viral pathogens (Coppel and Benjamin 1965; Wilson *et al*. 1992; Codella and Raffa 1993). To defend against predators and parasites, pine sawflies have evolved responsive chemical defenses: when threatened, larvae regurgitate a resinous defensive fluid, which they sequester from the host during feeding in a specialized pair of esophageal diverticula (Codella and Raffa 1993). This defensive regurgitation, which is often accompanied by a characteristic "U-bend" posture (Figure 1) and rhythmic jerking, is an effective repellant against many different predators and parasitoids (Eisner *et al*. 1974; Codella and Raffa 1995; Lindstedt *et al*. 2006, 2011). Although most *Neodiprion* species appear to be chemically defended and exhibit similar defensive displays, they vary from a green striped morph that is cryptic against a background of pine foliage to highly conspicuous aposematic morphs with dark spots or stripes overlaid on a bright yellow or white background (Figure 1). Thus, larval color is likely to confer protection against predators either via preventing detection (crypsis) or advertising unpalatability (aposematism) (Ruxton *et al*. 2004).

Beyond contributing to cryptic or aposematic coloration, *Neodiprion* larval color traits are likely shaped by several additional abiotic and biotic selection pressures. For example, coloration plays diverse ecological roles in insects, including thermoregulation, protection against UV damage, desiccation tolerance, and resistance to abrasion (True 2003; Lindstedt *et al.* 2009; Wittkopp and Beldade 2009). In addition to acting on specific color traits, selection may also act on color alleles via their pleiotropic effects on other traits, such as behavior, immune function, diapause/photoperiodism, fertility, and developmental timing (True 2003; Wittkopp and Beldade 2009; Heath *et al.* 2013; Lindstedt *et al.* 2016). In short, there are many direct and indirect selection pressures acting on larval coloration, and temporal and spatial variation in these pressures likely contribute to the abundant intraspecific and interspecific variation in the genus *Neodiprion.*

As a first step to testing the hypothesis that the genetic basis of pigmentation is predictably simpler than the genetic basis of color pattern (hereafter, the "pigmentation/pattern hypothesis"), we conducted a quantitative trait locus (QTL) mapping study of larval pigmentation and larval spotting pattern in the redheaded pine sawfly, *Neodiprion lecontei* (Figure 2A). This species is widespread across eastern North America, where it feeds on multiple pine species. A recent population genomic study identified three main genetic clusters within *N. lecontei* that diverged during the Pleistocene and exhibit consistent differences in larval coloration: (1) a southeastern

lineage ("South") with heavily spotted, yellow larvae and dark head capsules in early instars, (2) a central/mid-Atlantic lineage ("Central") with heavily spotted larvae that are predominantly white in some populations and predominantly yellow in others, and (3) a northern lineage ("North") with reduced spotting and yellow bodies (Bagley *et al.* 2017). We focus here on the genetic basis of color differences between a white, heavily spotted "Central" population and a yellow, lightly spotted "North" population (Figure 2).

The pigmentation/pattern hypothesis predicts that, compared to the distribution of effect sizes underlying larval color pattern, the distribution of effect sizes underlying larval pigmentation will be shifted towards mutations of larger effect. Although detection limits on small-effect QTL make it impossible to estimate the full effect-size distribution (Otto and Jones 2000; Mackay *et al.* 2009; Rockman 2012), we can nevertheless determine whether there are obvious differences among traits in the number and effect size of QTL that pass the detection threshold in a single mapping population.

Here, we use two complimentary approaches to test the prediction that the genetic architecture of larval pigmentation will be shifted towards large-effect mutations. First, we perform traditional interval mapping analyses in R/qtl (Broman and Sen 2009). Using this method, we determine both the number and effect sizes of QTL detected for each trait. The pigmentation/pattern hypothesis predicts that, compared to larval spotting, larval body color will have the largest-effect QTL and/or more QTL of moderate to large effect. Second, to provide a more direct description of genetic architecture, we employ a Bayesian Sparse Linear mixed model (BSLMM) developed for genome-wide association (GWA) mapping (Zhou *et al.* 2013). This approach yields quantitative estimates of trait genetic architecture, including a parameter that describes the proportion of genetic variance that is attributable to large-effect ("sparse") alleles. The pigmentation/pattern hypothesis predicts that estimates for this parameter will be higher for larval pigmentation than for larval color pattern. Finally, as a first step to identifying casual loci, we use our linkage map data to anchor the current *N. lecontei* genome assembly, then identify potential candidate genes within each QTL interval.

MATERIALS AND METHODS

Cross Design

To investigate the genetic architecture underlying sawfly color traits, we crossed *Neodiprion lecontei* females from a white-bodied, dark spotted population (collected from Valley View, VA; 37°54'47"N, 79°53'46"W) to *N. lecontei* males from a yellow-bodied, light-spotted population (collected from Bitely, MI; 43°47'46"N, 85°44'24"W). Both populations had been collected from the field in 2012 and reared on *Pinus banksiana* (jack pine) for at least two (VA population) or three (MI population) generations in the lab via standard rearing protocols (described in more detail in Harper *et al.* 2016; Bendall *et al.* 2017). Our mapping families were derived from four grandparental pairs, which produced 10 F_1 females. Like most hymenopterans, *N. lecontei* adults reproduce via arrhenotokous haplodiploidy, in which unfertilized eggs develop into haploid males and fertilized eggs develop into diploid females (Heimpel and de Boer 2008; Harper *et al.* 2016). Therefore, to produce an F_2 haploid generation, we allowed virgin F_1 females to lay eggs and reared their haploid male progeny on *P. banksiana*

foliage until they reached a suitable size for phenotyping. In total, we collected phenotypic and genotypic data from 429 F_2 male progeny for QTL mapping.

Color phenotyping

N. lecontei larvae pass through five (males) or six (females) feeding instars and a single non-feeding instar, which are distinguishable on the basis of color pattern and size (head capsule width and body length) (Benjamin 1955; Coppel and Benjamin 1965; Wilson *et al.* 1992). For phenotyping, we chose only mature feeding larvae, which have an orange-red head capsule with a black ring around each eye and up to four paired rows of gray to black spots (Wilson *et al.* 1992). We excluded any individuals that had molted to the final non-feeding instar, which have a very different head capsule color and spotting pattern. To generate digital images for phenotyping, we photographed CO_2-immobilized larvae (lateral surface) with a Canon EOS Rebel t3i camera equipped with an Achromat S 1.0X FWD 63mm lens. We then preserved each larva in 100% ethanol for molecular work. In total, we generated color phenotype data for 30 individuals from the VA population (mixed sex), 30 individuals from the MI population (mixed sex), 47 F_1 females, and 429 F_2 males (progeny of 10 virgin F_1 females).

To quantify larval body color from our digital photos, we used Adobe Photoshop CC 2014 or 2015 (Adobe Systems Incorporated, San Jose, CA) to ascertain the amount of yellow present, following O'Quin *et al.* (2013). First, we converted each digital image to CMYK color mode. Next, we selected the eye dropper tool (set to a size of 5x5 pixels) as the color sampler tool, which we used to sample three different body locations: the body just behind the head and parallel to the eye, the first proleg, and the anal proleg. For each of the three regions, this procedure yielded an estimate of the proportion of the selected area that was yellow. We then averaged the three measurements to produce a single final measurement of yellow pigmentation (hereafter referred to as "yellow").

To quantify larval spotting pattern, we used Adobe Photoshop's quick-selection tool to measure the area of the larval body (minus the head capsule) and the area of each row of lateral black spots. To control for differences in larval size, we divided the summed area of all lateral black spots by the area of the larval body. We refer to this phenotypic measure as "spotting". We also used the larval images to calculate the area of the head capsule, which we used as a covariate in some analyses to control for larval size (see below). We used a custom Perl script to process Photoshop measurement output files in bulk (written by John Terbot II; available upon request).

To determine whether mean phenotypic values for yellow and spotting differed between the two populations and among the three generations of our cross, we performed Welch's two-tailed *t*-tests. To determine the extent to which yellow and spotting co-varied in the F_2 males, we calculated Pearson's correlation coefficient (r). To determine which covariates to include in our QTL models, we performed ANOVAs to evaluate the relationship between the two phenotypes in the F_2 males (429 total) and their F_1 mothers (10 total) and head capsule sizes (a proxy for larval size/developmental stage). These and all other statistical analyses were performed in R version 3.3.2 (R Core Team 2013)

Genotyping

We extracted DNA from ethanol-preserved larvae using a modified CTAB method (Chen *et al.* 2010) and prepared barcoded and indexed double-digest RAD

(ddRAD) libraries using methods described elsewhere (Peterson *et al.* 2012; Bagley *et al.* 2017). We chose *NlaIII* and *EcoRI* as our restriction enzyme pair based on our previous experience using this enzyme pair for a population genomic analysis of *N. lecontei* (Bagley *et al.* 2017). We prepared a total of 10 indexed libraries: one consisting of the eight grandparents and 10 F_1 females (18 adults total), and the remaining nine consisting of F_2 haploid male larvae (~48 barcoded males per library). After digestion, adapter ligation, and pooling, we performed automated size selection of a 376-bp fragment (+/- 38 bp) from each library on a Pippin Prep (Sage Science, Beverly MA). Following size selection, we performed 12 rounds of high-fidelity PCR amplification using a unique Illumina multiplex read index for each library (adapter and primer sequences were as described in Bagley *et al.* 2017). After verifying library quality using a Bioanalyzer 2100 (Agilent, Santa Clara, CA), we sent all 10 libraries to the University of Illinois Urbana-Champaign Roy J. Carver Biotechnology Center (Urbana, IL), where the libraries were pooled and sequenced using 100-bp single-end reads on two Illumina HiSeq2500 lanes. In total, we generated 400,621,900 reads.

We demultiplexed and quality-filtered raw reads using the protocol described in Bagley *et al.* 2017. We then used Samtools v0.1.19 (Li *et al.* 2009) to map our reads to our *N. lecontei* reference genome (Vertacnik *et al.* 2016) and STACKS v1.37 (Catchen *et al.* 2013) to extract loci from our reference alignment and to call SNPs. We called SNPs in two different ways. First, for interval mapping analyses, our goal was to recover markers that represented fixed differences between the grandparental lines. To do so, we first called SNPs in our eight grandparents and 10 F_1 mothers. For these 18 individuals, which included both haploid males and diploid females, we required that SNPs had a minimum of 7x coverage and no more than 12% missing data. We then examined the resulting grandparental and F_1 genotypes to compile a list of SNPs that represented fixed differences between the two populations and, as an additional quality check, confirmed that all F_1 females were heterozygous at these SNPs. We then used STACKS to call SNPs in the F_2 haploid males, requiring that each SNP had a minimum of 5x coverage (we required a lower coverage for haploid males because we did not need to distinguish between homozygous and heterozygous genotype calls), no more than 10% missing data, and was present in the curated list produced from the grandparents. Filtering in STACKs produced a total of 559 SNPs genotyped in 429 F_2 males.

Second, to maximize the number of SNPs available for GWA mapping analyses and genome scaffolding, we ran an additional STACKS run using only the F_2 haploid males, requiring that each SNP had a minimum of 4x coverage. By removing the requirement that SNPs were called in all grandparents, we could recover many more SNPs. We then filtered the data in VCFtools v0.1.14 (Danecek *et al.* 2011) to remove individuals with a mean depth of coverage less than one, retaining 408 F_2 males. After removing low-coverage individuals, we used VCFtools to remove sites with a minor allele frequency (MAF) less than 0.05 (as these are unlikely to recover significant genotype-phenotype associations), sites with >5 heterozygotes (in haploid males, high heterozygosity is a clear indication of genotyping error), and sites with more than 50% missing data. To examine the impact of data completeness threshold and SNP number on our GWA mapping results, we also produced MAF- and heterozygote-filtered datasets with more stringent (0% missing data) and less stringent (<75% missing data)

completeness filters. In total, our three filtered datasets contained 1205 SNPs (0% missing data), 3069 SNPs (<50% missing data), and 4162 SNPs (<75% missing data).

Linkage map construction and genome scaffolding

To construct a linkage map for interval mapping, we started with 559 SNPs scored in 429 F_2 males. After an additional round of filtering in R/qtl (Broman and Sen 2009), we removed 11 haploid males that had >50% missing data, for a total of 418 F_2 males. Additionally, after removing SNPs that were genotyped in <70% of individuals, had identical genotypes to other SNPs, and had distorted segregation ratios, we recovered a final dataset of 503 SNPs. To assign these markers to linkage groups, we then used the "formLinkageGroups" function, requiring a minimum logarithm of odds (LOD) score of 6.0 and a maximum recombination frequency of 0.35. To order markers on linkage groups, we used the "orderMarkers" function, with the Kosambi mapping function to allow for crossovers. Following this initial ordering, we performed rippling on each linkage group to check whether switching marker order could improve LOD scores.

Anchoring genome scaffolds to linkage groups requires that scaffolds contain markers in the genetic linkage map. Our initial map included 503 SNPs spread across 358 scaffolds (out of 4523 scaffolds; Vertacnik *et al.* 2016). To increase the number of scaffolds and bases that we could place on our linkage groups, we therefore performed additional linkage mapping analyses with a larger SNP dataset that was called in F_2 males without any constraints on parental genotypes (filtered to remove individuals with depth of coverage < 1 and sites with MAF<0.05%, missing data > 50%, and >5 heterozygotes; remaining heterozygous sites were then treated as missing data). We then constructed a linkage map for each of our four grandparental families (N = 54, 73, 120, and 161).

For each grandparental family, we first performed additional data filtering in R/qtl to remove duplicate SNPs, SNPs with >50% missing data, and SNPs with distorted segregation ratios (which enabled us to remove SNPs that were monomorphic within families or that did not segregate in all of the F_2 families from a given grandparental pair). We then used the "formLinkageGroups" command, variable LOD thresholds (range: 5-15), and a maximum recombination frequency of 0.35. Because SNPs were not coded according to grandparent of origin, many alleles were "switched". We therefore performed an iterative process of linkage group formation, visualization of pairwise recombination fractions and LOD scores ("plotRF" command), and allele switching ("switchAlleles" command) until we obtained seven linkage groups (the number of *N. lecontei* chromosomes; Smith 1942, Maxwell 1958; Sohi and Ennis 1981) and a recombination/LOD plot indicative of linkage within, but not between, linkage groups. At this point, we ran into a second dilemma—with a denser panel of SNPs, allele ordering and examination of alternative SNP orders became prohibitively slow in R/qtl. To overcome this limitation, we used the more efficient MSTmap algorithm, implemented in R/ASMap v0.4-7 (Taylor and Butler 2017), to order our markers along their assigned linkage groups.

Finally, to order and orient our genome scaffolds along linkage groups (chromosomes), we used ALLMAPS (Tang *et al.* 2015) to combine information from our five maps (initial map with all individuals, but limited markers; plus four additional maps, each with more markers, but fewer individuals). Because maps constructed from

larger families are likely to be more accurate than those constructed from small families, we weighted the maps according to their sample sizes.

Interval mapping analysis

After linkage map construction, we used R/qtl to map QTL for our two color traits. Based on our phenotypic analyses, we included F_1 mother and head capsule size as covariates in our analysis of spotting and no covariates in our analysis of yellow. For each trait, we performed interval mapping using multiple imputation mapping. We first used the "sim.geno" function with a step size of 0 (*i.e.*, genotypes only drawn at marker locations) and 64 replicates. We then used the "stepwiseqtl" command to detect QTL and select the multiple QTL model that optimized the penalized LOD score (Broman, Manichaikul et al. 2009). To obtain penalties for the penalized LOD scores, we used the "scantwo" function to perform 1,000 permutations under a two-dimensional, two-QTL model that allows for interactions between QTL and the "calc.penalties" function to calculate penalties from these permutation results, using a significance threshold of $\alpha = 0.05$. Finally, for each QTL retained in the final model, we calculated a 1.5-LOD support interval.

Polygenic association mapping analysis

Although interval mapping has long been the analysis of choice for QTL mapping in experimental crosses, a growing number of studies are employing regression based approaches that were originally developed for genome-wide association mapping studies (Yi and Banerjee 2009; Huang *et al.* 2015; Li *et al.* 2017). For example, in a recent study of the genetic architecture of stickleback brain size, Li and colleagues (2017) demonstrated that compared to a traditional interval mapping approach, a polygenic modeling approach had increased statistical power for QTL detection, a reduced false positive rate, was better able to handle a large number of markers, and provided parameter estimates describing trait genetic architecture (genomic heritability). Given that our overall goal is to compare trait genetic architectures, the ability to estimate parameters describing trait genetic architecture is an especially attractive feature of polygenic association mapping approaches.

With these advantages in mind, we used GEMMA v0.94.1 to fit a Bayesian Sparse Linear Mixed Model (BSLMM) to our data (Zhou *et al.* 2013). The BSLMM is essentially a hybrid between a polygenic linear mixed model (LMM) and polygenic sparse regression model, which make opposing assumptions regarding trait genetic architecture: whereas LMMs generally assume that all genetic variants impact the phenotype, with normally distributed effect sizes, sparse regression models assume that only a small proportion of variants impact the phenotype. In combining these approaches, the BSLMM enables the genetic architecture to be estimated from the data and, as a consequence, performs well across a wide range of genetic architectures. This approach can also control for uneven relatedness among samples (*e.g.*, due to population stratification or, in our case, different grandparents and mothers) via inclusion of a relatedness matrix.

To ensure that our phenotypic data fit the model assumptions, both larval color traits were normal-quantile transformed in R prior to analysis. Additionally, because GEMMA cannot include covariates when fitting the BSLMM, we controlled for head

capsule size via analyzing the residuals of the linear regression between head capsule size and spotting (variation among mothers was controlled for via inclusion of a relatedness matrix). The GEMMA algorithm also requires complete (or imputed) genotype data. To explore the robustness of our genetic architecture results to different SNP numbers and missingness thresholds, we ran GEMMA analyses for each of the three SNP datasets generated from our F_2 males (no missing data, <50% missing data, and <75% missing data). For each of the two datasets that contained missing genotypes, we imputed missing data with BIMBAM v1.0 (Scheet and Stephens 2006) and used the resulting "best guess genotype" in our GEMMA analyses. File conversion between different input formats was accomplished via a combination of VCFtools v0.1.14 (Danecek *et al.* 2011), PLINK v1.90b3.46 (Purcell *et al.* 2007), and FCGene v1.0.7 (Roshyara and Scholz 2014).

For both traits and each of the three SNP datasets (six datasets total), we performed 10 independent GEMMA runs with the "-bslmm 1" option, each consisting of 25 million generations, with the first five million generations discarded as burn-in. To ensure convergence on the posterior distribution, we confirmed that parameter estimates were similar across independent runs and that parameter traces did not show any obvious increasing or decreasing trends. For each dataset, we then averaged posterior inclusion probabilities (PIP) for each SNP across all 10 runs and used the R package "qqman" (Turner 2014) to generate Manhattan plots from the averaged PIPs. To identify the most promising candidate SNPs, we ranked SNPs by their PIP scores and retained the top 1% for each dataset. We calculated the total effect size for each SNP in each run as: $\beta_i\gamma_i + \alpha_i$ (Zhou *et al.* 2013), then averaged effect size estimates across the 10 independent runs to obtain a single effect size estimate for each SNP and dataset.

We also summarized parameter estimates describing trait genetic architecture, including percent variance explained (PVE), percent of genetic variance that is due to large (or "sparse") effect alleles (PGE), and the number of SNPs in the polygenic model ("n_gamma" in GEMMA output). After ensuring that results were similar across independent runs, we computed medians and 95% credible intervals for the combined posterior distributions derived from all 10 runs (each consisting of 20 million post-burnin generations sampled every 1,000 generations, or 20,000 samples per run) for each parameter and each of our three SNP datasets.

Using our GEMMA results, we evaluated the pigmentation/pattern hypothesis in two ways. First, we asked whether the magnitudes of the estimated effect sizes for the largest-effect SNPs were consistently higher for yellow than for spotting. To obtain effect size magnitudes, we first calculated the absolute value of the average effect size for each SNP (across 10 independent runs), then used these values to calculate for each dataset the maximum effect size and median effect size for top 1% SNPs. Then, for each of the three SNP datasets, we used a nonparametric Wilcoxon rank-sum test to evaluate the null hypothesis that the effect sizes for the top 1% SNPs are equal across the two traits. Second, we used the genetic architecture parameter posterior distributions to evaluate the prediction that the percentage of genetic variance attributable to alleles of non-negligible effects (PGE) will be higher for yellow than for spotting.

Candidate gene analysis

To identify candidate QTL regions, we looked for regions of overlap between R/qtl and GEMMA analyses. We first identified the 1.5-LOD support interval

surrounding each QTL identified via interval mapping analyses. Then, we expanded this interval to include any candidate SNPs identified by GEMMA (*i.e.*, PIP scores in 99th percentile) that were within 1.5 Mb (which corresponds to ~5cM, see results) of either end of this QTL interval. Next, we used our anchored genome scaffolds to compile a list all scaffolds that fell within the physical intervals defined by the combined GEMMA and R/qtl results. Finally, we compiled a list of all remaining scaffolds with SNPs identified as candidates in our GEMMA analyses.

After compiling candidate regions in the *N. lecontei* genome, we compiled a list of candidate color genes. For larval spotting, we included genes in the melanin synthesis pathway and genes that have been implicated in pigment patterning (Wittkopp *et al.* 2003; Protas and Patel 2008; Wittkopp and Beldade 2009; Sugumaran and Barek 2016). For larval pigmentation, we included genes implicated in the transport, deposition, and processing of carotenoid pigments derived from the diet (Palm *et al.* 2012; Yokoyama *et al.* 2013; Tsuchida and Sakudoh 2015; Toews *et al.* 2017). Although several pigments can produce yellow coloration in insects (*e.g.*, melanins, pterins, ommochromes, and carotenoids), we focused on carotenoids because a heated pyridine test (McGraw *et al.* 2005) was consistent with carotenoid-based coloration in *N. lecontei* larvae (Figure S1).

Once we had compiled a list of candidate genes, we searched for these genes by name in the *N. lecontei* v1.0 genome assembly and NCBI annotation release 100 (Vertacnik *et al.* 2016). To find missing genes and as an additional quality measure, we obtained FASTA files corresponding to each candidate protein and/or gene from NCBI (using *Apis, Drosophila melanogaster,* or *Bombyx mori* sequences, depending on availability). We then used the i5k Workspace@NAL (Poelchau *et al.* 2014) BLAST (Altschul *et al.* 1990) web application to conduct tblastn (for protein sequences) or tblastx (for gene sequences) searches against the *N. lecontei* v1.0 genome assembly, using default search settings. After identifying the top hit for each candidate gene/protein, we then used the WebApollo (Lee *et al.* 2013) JBrowse (Skinner *et al.* 2009) *N. lecontei* genome browser to identify the corresponding predicted protein coding genes (from NCBI annotation release 100) in the *N. lecontei* genome.

We took additional steps to identify genes in the *yellow* gene family, all of which contain a major royal jelly protein (MRJP) domain. First, we used the search string "major royal jelly protein *Neodiprion*" to search the NCBI database for all predicted *yellow*-like and *yellow-MRJP*-like *N. lecontei* genes. We then downloaded FASTA files for the putative *yellow* gene sequences (26 total). Next, we used the Hymenoptera Genome Database (Elsik *et al.* 2016) to conduct a blastx search of our *N. lecontei* gene sequence queries against the *Apis mellifera* v4.5 genome NCBI RefSeq annotation release 103. Finally, we recorded the top *A. mellifera* hit for each putative *N. lecontei yellow* gene.

Once we identified the location of candidate color genes in the *N. lecontei* genome, we asked whether any of these genes were located within scaffolds contained within our candidate QTL intervals.

Data availability

Short-read DNA sequences will be made available via the NCBI SRA (Bioproject PRJNA#######, Biosample numbers SAMN########-SAMN########). The linkage-group anchored assembly will be submitted to NCBI and i5k to update the existing *N.*

lecontei genome assembly and annotations (Vertacnik *et al.* 2016). All remaining files (phenotypic data, SNP datasets, and input files for R/qtl and GEMMA) will be submitted to DRYAD (doi:#######).

RESULTS AND DISCUSSION

Phenotypic variation

Lab-reared larvae derived from the two founding populations (white-bodied, dark spotted VA population and the yellow-bodied, light spotted MI population) differed significantly from one another in both pigmentation and spotting pattern (Figures 2, 3; yellow: $t_{42} = 63.52$, $P < 1 \times 10^{-15}$; spotting: $t_{41.68} = 23.13$, $P < 1 \times 10^{-15}$). Because all larvae were reared on the same host under the same laboratory conditions (*i.e.*, minimal environmental variance), these results suggest that genetic variance contributes to variance in both larval color traits. Crosses between the VA and MI lines produced diploid F_1 females that appeared intermediate in both pigmentation and spotting pattern (Figure 2), and F_1 female larvae differed significantly from both parents for both traits (Figure 3; yellow MI vs. F_1: $t_{51.65} = 15.41$, $P < 1 \times 10^{-15}$; yellow VA vs. F_1: $t_{66.29} = 18.26$, $P < 1 \times 10^{-15}$; spotting MI vs. F_1: $t_{52.33} = 10.52$, $P = 1.63 \times 10^{-14}$; spotting VA vs. F_1: $t_{68.37} = 2.92$, $P = 0.0047$). These results indicate that neither pigmentation phenotype (white vs. yellow) and neither spotting phenotype (light vs. dark) is completely dominant.

Haploid F_2 males produced by virgin F_1 mothers varied continuously in both spotting and pigmentation, and differed significantly from the two founding populations in both traits (Figure 3; yellow MI vs. F_2: $t_{427.06} = 42.55$, $P < 1 \times 10^{-15}$; yellow VA vs. F_2: $t_{167.05} = 8.05$, $P = 1.48 \times 10^{-13}$; spotting MI vs. F_2: $t_{165.28} = 20.85$, $P < 1 \times 10^{-15}$; spotting VA vs F_2: $t_{53.49} = 8.20$, $P = 4.96 \times 10^{-11}$). F_2 males also differed significantly from F_1 females in pigmentation, but not spotting (yellow F_1 vs. F_2: $t_{91.69} = 11.61$, $P < 1 \times 10^{-15}$; spotting F_1 vs. F_2: $t_{56.79} = 1.87$, $P = 0.066$). We also found a weak, but significant and negative correlation between spotting area and percent yellow in F_2 males (*i.e.*, yellower individuals tend to be less heavily spotted; Pearson's $r = -0.12$, $P = 0.013$). This correlation, which could be explained by pleiotropy or physical linkage, suggests that these two traits do not evolve completely independently of one another. Nevertheless, the correlation is relatively weak and we observed many different combinations of spotting and pigmentation in the recombinant F_2 males (Figure 2).

Because F_2 males are haploid, interactions between alleles at a single locus (dominance effects) are eliminated. Thus, the range of phenotypic variation observed in F_2 males is determined by the number and effect sizes of color alleles and epistatic interactions between them. For both color traits, phenotypic variation observed in the F_2 males spanned—and even exceeded—the full range of variation observed in the grandparental populations and F_1 females (Figure 3). The observation that grandparental pigmentation and spotting phenotypes are recapitulated in the F_2 males suggests that both traits are controlled by a relatively small number of loci. There are multiple, non-mutually exclusive explanations for the transgressive color phenotypes in our haploid F_2 males, including: variation in the grandparental lines, reduced developmental stability in hybrids, epistasis, unmasking of recessive alleles in haploid males, and the complementary action of additive alleles from the two grandparental lines (Rieseberg *et al.* 1999).

Linkage mapping and genome scaffolding

For our full F_2 SNP dataset, which consisted of 503 fixed differences scored in 429 F_2 haploid males, our markers were spread across seven linkage groups (LG), which matches the number of *N. lecontei* chromosomes (Smith 1941; Maxwell 1958; Sohi and Ennis 1981). The total map length was 1169 cM, with an average marker spacing of 2.4 cM and maximum marker spacing of 24.3 cM (Table S1; Figure S2). Together, these results indicate that this linkage map is of sufficient quality and coverage for interval mapping. Additionally, with an estimated genome size of 340 Mb (estimated via flow cytometry; C. Linnen, personal observation), these mapping results yield a recombination density estimate of 3.43 cM/Mb. This recombination rate is lower than that observed in social Hymenoptera, which have among the highest rates of recombination in eukaryotes (Wilfert *et al.* 2007). Nevertheless, this rate is on par with that reported in other (non-eusocial) hymenopterans, which lends support to the hypothesis that elevated recombination rates in eusocial hymenopterans species is a derived trait and possibly an adaptation to a social lifestyle (Gadau *et al.* 2000; Schmid-Hempel 2000; Crozier and Fjerdingstad 2001).

Linkage maps estimated for the four grandparental families, each of which contained >2000 markers, ranged in length from 1072 cM to 3064 cM (Table S1). This variation in map length is likely attributable to both decreased mapping accuracy in smaller families and decreased genotyping accuracy in these less-stringently filtered SNP datasets. Nevertheless, our scaffolding analysis revealed that marker ordering was highly consistent across linkage maps (Figures S3-S9). Additionally, via including SNPs that were variable only in some families, we were able to more than triple the number of mapped scaffolds (from 358 to 1005) and increase the percentage of mapped bases from 41.2% to 78.9% (Tables S2-S3). Anchored genome scaffolds, coupled with existing *N. lecontei* gene annotations, are a valuable resource for identification of candidate genes within QTL.

Detection of color QTL via interval and polygenic association mapping

Using an interval mapping approach, we obtained significant QTL for both traits. For yellow, the full stepwise model recovered six QTL, with a significant interaction between QTL on LGs 3 and 5 (Table 1; Figure 4A). This model, which had a LOD (logarithm of the odds) score of 182.03, explained 85.8% of the total variance in larval pigmentation. The two largest-effect QTL (Yellow-4 and Yellow-5) reside on LGs 3 and 5, each explaining ~16% of the variance in larval pigmentation. These QTL also accounted for a substantial fraction of the phenotypic difference between the grandparental lines (Yellow-4: 27%; Yellow-5: 52%; Table 1, Figure 4A, Figure 5A-B), and the interaction between these two QTL was highly significant ($P = 6.5 \times 10^{-14}$, Figure 5C). Examination of the interaction plot reveals that individuals carrying the VA allele for the Yellow-5 QTL have drastically reduced yellow pigmentation, making the additional impact of the Yellow-4 QTL less pronounced (Figure 5C). Possible reasons why the effects of the Yellow-4 "VA-white" allele are most pronounced on the Yellow-5 "MI-yellow" background are considered further in our discussion of candidate genes.

For spotting, we detected only two QTL via interval mapping, both of which were located on LG 2 (Table 1, Figure 4C). The first peak (Spot-1) explained 13% of the

phenotypic variance in F_2 males and 35% of the difference in spotting between the grandparental strains (Figure 5D). The second peak (Spot-2) explained 35% of the variance in F_2 males and 57% of the difference in yellowness between the grandparents (Figure 5E). The full stepwise model, which included head capsule area and mother as covariates, had a LOD score of 94.9 and explained 64.0% of the total variance in larval spotting pattern. This model also included a subtle, but significant ($P = 0.0051$) interaction between the two spotting QTL. Specifically, the effects of the dark-spotting VA allele for the Spot-1 QTL are more pronounced on a genetic background containing the dark-spotting VA allele for the Spot-2 QTL (Figure 5F). Notably, the 1.5-LOD support intervals for the two large-effect spotting QTL overlap with those of two small-effect yellow QTL (Yellow-2 and Yellow-3) (Table 1). This co-localization of spotting and yellow QTL is consistent with the weak phenotypic correlation observed between these traits in F_2 males.

We also recovered strong associations between genotype and phenotype for both traits using a polygenic association mapping approach. Because effect sizes, PIP estimates, and genetic architecture parameter estimates were highly consistent across independent GEMMA runs (Table S4), we combined results from each set of 10 runs by (1) averaging per-SNP PIP and effect scores across runs, and (2) combining parameter posterior distributions into a single distribution for each trait/SNP dataset. Overall, there was considerable overlap between the location of QTL indicated by the GEMMA analyses and those implicated by interval mapping analyses (Figure 4, Tables S5-S6). Specifically, GEMMA analyses recovered candidate SNPs in or in close proximity to (*i.e.*, within 1.5 million base pairs or < 5 cM) all QTL intervals identified by interval mapping (Tables S5-S6). Moreover, QTL with high LOD scores had correspondingly high PIP estimates (PIP > 0.80). That said, the precise location of the QTL peaks (highest LOD score or PIP value) differed slightly among interval mapping and GEMMA analyses and among GEMMA analyses utilizing different missing data thresholds. These differences are likely attributable to differences in the SNPs included in the analysis (*e.g.*, many SNPs in GEMMA analyses with 50% and 75% missing data were not present in the 0%-missing GEMMA and R/qtl analyses). Additionally, the stringently curated R/qtl and 0% missing GEMMA datasets are less likely to contain genotyping error that may obscure genotype-phenotype associations.

In addition to the eight color QTL detected via interval mapping, GEMMA analyses identified several additional regions of the genome associated with larval color variation (Tables S5-S6). Although the 99th percentile PIP threshold is somewhat arbitrary, it is nevertheless quite stringent. Depending on the dataset, our top 1% PIP scores represented a 3- to 340-fold increase over the genome-wide average PIP value. Although it is certainly possible that the GEMMA candidate SNPs represent false positives, identification of plausible candidate genes linked to at least some of these SNPs suggests that at least some GEMMA candidates may represent true positives (see below). One possible explanation for why GEMMA picked up regions that were not detected via interval mapping is that these regions contain genetic variants contributing to phenotypic variation segregating within one of the grandparental lines. Additionally, compared to the one- and two-SNP scans implemented in interval mapping analyses, the multi-SNP association mapping method implemented in GEMMA may have more power to detect QTL of small effect. This explanation is consistent with the observation that, compared to

single-SNP GWA mapping approaches, multiple-SNP GWA mapping approaches have increased power and reduced false positive rates (Hoggart *et al.* 2008).

Comparing genetic architectures: testing the pigmentation/patterning hypothesis

According to the pigmentation/pattern hypothesis, pigmentation loci are less likely to involve negative pleiotropic consequences and therefore, the distribution of effect sizes underlying changes in pigmentation (yellow) should be shifted towards large-effect mutations compared to the effect-size distribution of patterning (spotting) (Fisher 1930; Rockman 2012). Although limited power to detect QTL of very small effect precludes us from estimating the exact number of SNPs and full effect-size distribution, we can nevertheless ask whether there are any obvious differences between the upper ends of the effect-size distributions for different traits scored in the same mapping population.

Looking first at the interval mapping analyses, the number of QTL of relatively large effect is the same for both traits: for both yellow and spotting, there are two QTL with PVE > 10% and phenotypic effect sizes that exceed 25% of phenotypic difference between the grandparental lines. However, in contrast to our predictions, the QTL with the largest observed effect size was for spotting (Spot-1), not yellow. Additionally, more QTL are detected for yellow than for spotting, which could indicate that the genetic architecture of yellow is more complex (more loci) than that of spotting. These observations are seemingly at odds with the prediction that the effect-size distribution for yellow should be shifted towards large-effect QTL. That said, an alternative interpretation for our finding that fewer QTL were detected for spotting is that most spotting QTL were not of sufficient effect size for detection. Thus, our observation of more yellow QTL could be explained by an effect-size distribution that is shifted towards larger effects (*i.e.*, there more QTL of sufficient size for detection). Taken together, our interval mapping results are equivocal with respect to the pigmentation/patterning hypothesis.

In contrast to the interval mapping results, effect size estimates from the GEMMA analyses are largely consistent with the predictions of the pigmentation/pattern hypothesis. Across all SNP datasets, the effect size of the largest-effect SNP was always higher for yellow than for spotting (Table 2). Likewise, the median effect size for SNPs that fell within the 99th percentile for PIP was consistently higher for yellow than for spotting. However, none of the Wilcoxon rank-sum tests comparing the distribution of the top 1% effect sizes were significant (0% missing data: N = 13 SNPs, W = 105, P = 0.31; <50% missing data: N = 31 SNPs, W = 505, P = 0.74; <75% missing data: N = 42 SNPs, W = 1000, P = 0.29). That said, the PIP cutoff was somewhat arbitrary and the number of SNPs analyzed was relatively small.

By providing parameter estimates that describe trait genetic architecture and are independent of arbitrary cutoffs, the BSLMM implemented in GEMMA provides a more straightforward way to compare trait genetic architectures. Despite some dependence on the SNP dataset analyzed, among-trait differences in PVE and PGE were very consistent across runs (Figure 6A-B; Table 2). First, the amount of phenotypic variation explained by genetic variance was consistently higher for yellow, mirroring our interval mapping results (PVE for full yellow model: 85.8%; PVE for full spotting model: 64.0%). Moreover, for the 0%-missing dataset, 95% credible intervals for PVE for the two traits

did not overlap (Table 2). Possible explanations for differences in PVE include a greater measurement error for spotting and/or a greater contribution of environmental variance to spotting variance. Second, although estimates of the contribution of large-effect mutations to total genetic variance were high for both traits (PGE > 0.8), yellow PGE estimates were uniformly higher than spotting PGE estimates and, for the 0%-missing dataset, 95% credible intervals for PGE did not overlap (Table 2). Our observed differences in PGE estimates are consistent with the prediction that, compared to spotting, the effect-size distribution underlying the yellow trait is shifted towards mutations of larger effect.

In contrast to PVE and PGE, we did not observe consistent differences in the estimated number of large-effect SNPs across the three datasets and 95% credible intervals for this parameter always overlapped. For the 0%-missing and 50%-missing datasets, yellow had slightly more SNPs than spotting, but spotting had more SNPs than yellow in the 75%-missing datasets (Table 2). Additionally, for all three SNP datasets, yellow and spotting had very similar posterior distributions for SNP number (Figure 6C). Nevertheless, despite some uncertainty in our SNP number estimates, our effect size estimates and PGE estimates obtained under the BSLMM implemented in GEMMA provide support for the pigmentation/pattern hypothesis.

Limitations of our data for testing the pigmentation/pattern hypothesis

Although our genetic architecture parameter estimates are consistent with the prediction that the effect-size distribution is shifted towards large-effect alleles for pigmentation, there are three main limitations of our current data that preclude a more definitive test of the pigmentation/pattern hypothesis. First and foremost, we have mapped these traits to large genomic regions, each containing many genes (~2 – 3.5 Mb for the four QTL of largest effect; 2.2 – 7.7 Mb for remaining smaller-effect QTL). It is therefore possible that individual QTL comprise multiple linked mutations of individually smaller effect (Stam and Laurie 1996; McGregor *et al.* 2007; Bickel *et al.* 2011; Linnen *et al.* 2013). With this in mind, our effect sizes and PGE estimates are best interpreted as maximum values for each trait. Under the pigmentation/pattern hypothesis, we would predict that spotting QTL are more likely to fractionate than yellow QTL. Moreover, if patterning traits involve loci that are more likely to exhibit antagonistic pleiotropy than loci involved in pigmentation traits, we would expect to see a greater contribution of cis-regulatory changes—possibly multiple linked cis-regulatory changes (*e.g.*, Rebeiz *et al.* 2009; Frankel *et al.* 2011)—to variation in spotting compared to variation in yellow. These predictions could be tested via fine-mapping QTL and functional analysis of candidate genes and mutations (see below).

A second limitation of our data is that there are many other factors beyond antagonistic pleiotropy that can impact the expected distribution of mutational effect sizes, including the demographic and selective history of the phenotypically divergent *N. lecontei* populations. In terms of selection, theoretical predictions regarding the expected distribution of effect sizes are all based on model of adaptation in which beneficial mutations are fixed as a population moves towards a new phenotypic optimum (Orr 1998). A key assumption, therefore, is that the traits under investigation are adaptive. For *N. lecontei*, experimental evidence indicates that both white and yellow larvae are

highly conspicuous to avian predators when viewed against a background of pine foliage, and both morphs facilitate rapid avoidance learning in naïve avian predators (Lindstedt and Linnen, personal observation). Consistent with theoretical predictions that avoidance learning in predators will result in stabilizing selection on warning coloration (Joron and Mallet 1998; Kapan 2001; Kronforst and Gilbert 2008), most *N. lecontei* populations are fairly uniform in their color and pattern (Linnen, personal observation). Although stabilizing selection can explain the maintenance of larval color differences between *N. lecontei* populations in the face of gene flow, it does not explain the initial divergence in color. One possibility is that initial color differences arose via genetic drift in small populations that were isolated in pine refugia during the Pleistocene glaciations (Bagley *et al.* 2017). Once a novel morph reached a critical threshold to promote avoidance learning in the local predator community, it could then increase in frequency via selection (Mallet and Singer 1987; Turner and Mallet 1996; Kronforst and Gilbert 2008). A non-mutually exclusive explanation is that among-population differences in larval color stem from geographic variation in other selection pressures, such as climate, host-plant defenses, and local communities of viruses, parasitoids, and predators (Nokelainen *et al.* 2014, Amézquita *et al.* 2017, Willmott *et al.* 2017).

Beyond demonstrating that a particular trait is adaptive, other details of the selective history are also important to predicting effect-size distributions. For example, theoretical work indicates that adaptation to a distant phenotypic optimum, adaptation to a rapidly moving optimum, adaptation to a multi-peaked fitness surface, and adaptation from new beneficial mutations can all shift the predicted effect-size distribution towards mutations of larger effect (Orr 1998; Hermisson and Pennings 2005; Matuszewski *et al.* 2014, 2015; Dittmar *et al.* 2016). In short, additional work is needed to test our assumption that larval color traits are locally adaptive and to more fully explore the targets, agents, and history of selection on larval color traits. We note, however, that even if larval color evolution is predominantly neutral, we would still expect pleiotropy to impact the expected distribution of effect sizes of color mutations fixed under genetic drift. For example, for a highly pleiotropic gene, large-effect alleles are more likely to be deleterious than small-effect alleles and therefore less likely to drift to fixation. Thus, if pigmentation genes are less pleiotropic, on average, than patterning genes, the pigmentation/patterning hypothesis should apply to both selected and neutrally evolving traits.

In terms of demographic history, theoretical work indicates that effect-size distributions will be shifted towards larger-effect mutations when effective population size is reduced (Kimura 1983) and when local adaptation is opposed by ongoing gene flow (Griswold 2006; Yeaman and Whitlock 2011). Based on a demographic analysis of genome-wide SNP data, Bagley *et al.* (2017) hypothesized that genetically distinct "North" (source of the light-spotted, yellow MI population) and "Central" (source of the dark-spotted, white VA population) lineages diverged from one another during the Pleistocene, ~25,000 years ago. Prior to this divergence, the North/Central ancestor diverged from a "South" lineage ~45,000 years ago. Using this early-diverging South lineage, which is dark-spotted and yellow, to polarize changes in larval color, we hypothesize that light spotting (MI) and white coloration (VA) are both derived character states within *N. lecontei*.

743 Compared to the Central lineage, the North lineage has drastically reduced
744 heterozygosity and an effective population size that is estimated to be ~89% lower than
745 that the Central N_e estimate (Bagley *et al.* 2017). Thus, beneficial light-spotting alleles of
746 small effect would have been more likely to be lost to drift in the North lineage than in
747 other lineages, thereby shifting the predicted effect-size distribution for spotting towards
748 larger-effect alleles. That said, the North lineage has experienced much less gene flow
749 from other lineages than the South and Central lineages (Bagley *et al.* 2017). The Central
750 lineage is also polymorphic for larval pigmentation, with a cluster of white-bodied
751 populations in the mid-Atlantic region, surrounded by genetically similar yellow-bodied
752 populations in other portions of the Central range (Bagley *et al.* 2017; Bagley and
753 Linnen, personal observation). Higher gene flow from yellow populations would have
754 favored larger-effect loci underlying locally adaptive white pigmentation compared to a
755 model lacking gene flow (Yeaman and Whitlock 2011). Consistent with these
756 demographic estimates and theoretical predictions, we found that both spotting and
757 yellow had high estimates for PGE (Table 2; Figure 6).
758 A third and final limitation is that, while a useful test of the *a priori* hypothesis
759 that the effect size of alleles contributing to pigmentation are larger, on average, than
760 those contributing to color pattern, our sample size nevertheless consists of a single
761 population pair. As the discussion above highlights, many factors can lead to differences
762 in trait genetic architectures and any one of these could have pushed the distribution of
763 pigmentation alleles towards larger effects. Because the history of adaptation—and
764 therefore the expected effect-size distribution of adaptive substitutions—is likely to be
765 highly idiosyncratic across populations and species, many phylogenetically independent
766 test cases will be needed to determine whether certain factors (*e.g.*, different levels of
767 pleiotropy associated with pigmentation and color patterning loci) are consistently
768 associated with differences in effect-size distributions.
769 To date, the most tested prediction regarding trait genetic architectures is that the
770 distribution of effect sizes will be shifted towards large-effect mutations when the
771 distance to the phenotypic optimum is large. This prediction has been supported in
772 multiple contexts. For example, Rogers *et al.* (2012) estimated QTL effect sizes
773 underlying shape and armor traits in replicate freshwater stickleback populations adapting
774 to phenotypic optima that were either close to the ancestral phenotype (predatory prickly
775 sculpin present) or far from the ancestral phenotype (no sculpin present). Consistent with
776 theory, they found that average effect sizes were larger when populations were adapting
777 to a more distant phenotypic optimum. Similarly, in specialist pupfish that have evolved
778 from a generalist ancestor, large-effect QTL contribute more to the enlarged scale-eater
779 jaw (distant phenotypic optimum) than to the molluscivore nasal protrusion (close
780 phenotypic optimum) (Martin *et al.* 2017; McGirr *et al.* 2017). Extensive intra- and
781 interspecific variation in larval pigmentation and patterning make *Neodiprion* an
782 especially promising system for testing the pigmentation/pattern hypothesis and, more
783 generally, examining the impact of pleiotropy on trait genetic architecture (Figure 1).
784
785 **Candidate genes for larval color traits**
786 Testing the pigmentation/patterning hypothesis will ultimately require identifying
787 causal genes and mutations and characterizing their phenotypic effects. To this end, we
788 asked whether our candidate QTL regions contained any genes with known or suspected

roles in carotenoid-based or melanin-based pigmentation. For larval pigmentation ("yellow"), we focused on genes known or suspected to be involved in carotenoid processing or deposition (see Figure S1). Although most animals cannot synthesize their own carotenoids, many use environmentally acquired carotenoids to produce red, orange, and yellow colors (Toews *et al.* 2017). Compared to endogenously produced pigments such as melanin, however, much less is known about the genetic pathways underlying carotenoid based pigmentation. Nevertheless, recent progress in identifying genes involved in carotenoid transport, deposition, and processing provided us with some candidates for carotenoid-based pigmentation (Table S7). These candidates include genes encoding carotenoid binding proteins (involved in binding carotenoids in gut and transporting to hemolymph), lipoproteins (involved in carotenoid transport in hemolymph), scavenger receptor proteins (involved in carotenoid uptake in specific tissues), β-carotene oxygenases (involved in carotenoid breakdown), and cytochrome p450s (possibly involved in carotenoid processing) (Bhosale and Bernstein 2007; Palm *et al.* 2012; Yokoyama *et al.* 2013; Tsuchida and Sakudoh 2015; Toews *et al.* 2017).

Both of our largest-effect yellow QTL contained promising candidate genes with known or suspected roles in carotenoid-based pigmentation (Table S7). First, within the overlapping Yellow-5 and Yellow-6 QTL intervals, we recovered a predicted protein coding gene in scaffold 164 with a high degree of similarity to the ApoLTP-1 and ApoLTP-2 protein subunits (encoded by the gene *apoLTP-II/I*) of the *Bombyx mori* lipid transfer particle (LTP) lipoprotein (e-value: 0; bitscore: 391). LTP is one of two major lipoproteins present in insect hemolymph and appears to be involved in the transport of lipids (including carotenoids) from the gut to the other major lipoprotein, lipophorin, which then transports lipids to target tissues (Tsuchida *et al.* 1998; Palm *et al.* 2012; Yokoyama *et al.* 2013). A second potential candidate in the overlapping Yellow-5 and Yellow-6 QTL intervals, located on scaffold 386, is a cytochrome P450. Although cytochrome P450s are best known for their role in detoxification, it is hypothesized that they serve diverse roles and at least one cytochrome p450 (*CYP2J19*) has been implicated in the conversion of yellow carotenoids to red ketocarotenoids in birds (Lopes *et al.* 2016; Mundy *et al.* 2016).

Within the Yellow-4 QTL, we found a predicted protein coding region in scaffold 518 with a high degree of similarity to the *Bombyx mori* Cameo2 scavenger receptor protein (e-value: 1×10^{-18}; bitscore: 92.8). *Cameo2* encodes a transmembrane protein belonging to the CD36 family that has been implicated in the selective transport of the carotenoid lutein from the hemolymph to specific tissues (Sakudoh *et al.* 2010; Tsuchida and Sakudoh 2015). In the silkworm *Bombyx mori*, *Cameo2* is responsible for the "C mutant" phenotype, which is characterized by a combination of yellow hemolymph and white cocoons that arises as a consequence of disrupted transport of lutein from the hemolymph to the middle silk gland (Sakudoh *et al.* 2010; Tsuchida and Sakudoh 2015). Based on this previous work, we hypothesize that a loss of function mutation in *Cameo2* contributes to the loss of yellow pigmentation in the integument of white-bodied *N. lecontei* larvae.

Our most promising yellow candidate genes—*apoLTP-II/I* and *Cameo2*—also provide a potential explanation for the epistatic interaction we detected between the Yellow-4 and Yellow-5 QTL (Figure 5C). We predict that reduced *apoLTP-II/I* function would reduce the amount of yellow carotenoids found in the hemolymph and reduced

Cameo2 function would reduce yellow carotenoids in the integument. In the transport of carotenoids from the gut to the integument, *Cameo2* therefore acts downstream of *apoLTP-II/I*. Thus, once the amount of carotenoids entering the hemolymph is already reduced via changes in *apoLTP-II/I* (Yellow-5), the impact of an additional reduction in the integument via changes in *Cameo2* (Yellow-4) would be reduced—this is consistent with the interaction plot in Figure 5C.

Our top two spotting QTL also yielded very promising candidate genes—this time in the well-characterized melanin biosynthesis pathway. Specifically, in both peaks, we found protein-coding genes that appear to belong to the *yellow* gene family. The *yellow* gene family encodes a functionally diverse set of proteins characterized by a shared major royal jelly protein (MRJP) domain (Ferguson *et al.* 2011). To date, *yellow* genes have been associated with diverse functions, including behavior, reproductive maturation, caste specification in honeybees, and pigmentation (Wittkopp *et al.* 2002; Drapeau *et al.* 2006; Prud'homme *et al.* 2006; Ferguson *et al.* 2011). Although much is still unknown about the function of *yellow* genes, studies in *D. melanogaster* suggest that *yellow-f* and *yellow-f2* have dopachrome-conversion enzymatic function, which is required for the production of melanic pigment (Han *et al.* 2002). Additionally, mapping and expression data indicate that deletions in *yellow-e* gene are responsible for changes in larval pigmentation patterning in two mutant strains of the *Bombyx mori* (Ito *et al.* 2010).

Among the 11,586 predicted protein coding genes in the *N. lecontei* v1.0 genome NCBI annotation release 100, we recovered 26 "*yellow*-like" or "*yellow-MRJP*-like" genes. Notably, this number is equivalent to the number of *yellow*-like/*yellow-MRJP*-like genes found in the genome of the jewel wasp, *Nasonia vitripennis,* which boasts the highest reported number of *yellow*-like genes of any insect to date (Werren *et al.* 2010). Of the 26 genes belonging to the *yellow* family, nine were most similar to *Apis yellow-MRJPs*. The remaining 17 genes appeared more similar to *Apis yellow* genes (top BLAST hits included: four *yellow-b*, one *yellow-e*, one *yellow-e3*, four *yellow-g*, one *yellow-h*, five *yellow-x*, and one *yellow-y*). Additionally, 13 of these genes (*yellow-e, yellow-e3,* four *yellow-g, yellow-h, yellow-x,* and five *MJRPs*) were found in tandem array along three adjacent scaffolds (548, 170, and 36; ~1 Mbp total) on LG 2. This genomic organization is consistent with a conserved clustering of *yellow-h, -e3, -e, -g2,* and *–g* observed across *Apis, Tribolium, Bombyx, Drosophila,* and *Nasonia* (Drapeau *et al.* 2006; Werren *et al.* 2010; Ferguson *et al.* 2011). Like *Nasonia* and *Apis*, this cluster also contains *MJRPs*; like *Heliconius*, this cluster contains a *yellow-x* gene.

The Spot-1 QTL contained 2 *yellow*-like genes that were most similar to *Apis yellow-x1* (e-value: 1.2×10^{-160}; bitscore: 471.47 and e-value: 3.9×10^{-151}; bitscore: 448.36). Notably, these genes were located within a scaffold (422) receiving very high PIP scores (>0.8) in multiple GEMMA analyses. However, there is currently little known about the function of *yellow-x* genes, which appear to be highly divergent from other *yellow* gene families (Ferguson *et al.* 2011).

The Spot-2 QTL contained the cluster of 13 *yellow* genes described above, along with two additional *MRJPs* on scaffold 769. Of these, *yellow-e* is the strongest candidate for larval spotting pattern. In two different mutant strains of *B. mori* ("*bts*" for brown head and tail spot), mutations in *yellow-e* produced a truncated gene product that results in increased reddish-brown pigmentation in the head cuticle and anal plate compared to wildtype strains (Ito *et al.* 2010). Quantitative reverse transcriptase analyses also

demonstrated that in wildtype larvae, *yellow-e* is most highly expressed in the integument of the head and the tail (Ito *et al.* 2010). Based on these observations, one possible mechanism for the reduced spotting observed in the light-spotted MI population is an increase in *yellow-e* expression.

The Spot-2 QTL also contained a predicted protein that was highly similar to tyrosine hydroxylase (TH) (e-score: 6×10^{-123}, bitscore: 406). TH catalyzes the hydroxylation of tyrosine to 3,4-dihydroxyphenylalanine (DOPA), a precursor to melanin-based pigments (Wright 1987). Work in the swallowtail butterfly *Papilio xuthus* and the armyworm *Pseudaletia separata* demonstrates that TH and another enzyme, dopa decarboxylase (DDC), are expressed in larval epithelial cells containing black pigment (Futahashi and Fujiwara 2005; Ninomiya and Hayakawa 2007). Furthermore, inhibition of either enzyme prevented the formation of melanin-based larval pigmentation patterns (Futahashi and Fujiwara 2005). Thus, a reduction in the regional expression of *TH* is another plausible mechanism underlying reduced spotting in the light-spotted MI population.

The Spot-1 and Spot-2 QTL also overlap with two minor-effect yellow QTL: Yellow-2 and Yellow-3 (table, figure). Co-localization of pigmentation and patterning QTL could be explained either by linkage or pleiotropy. As noted above, spotting and yellow values are negatively correlated in F_2 males. One explanation for this observation is that the loci in the Spot-1 and Spot-2 cluster that impact spotting area also impact overall levels of melanin throughout the integument. With increasing melanin content, larval color would appear less yellow, leading to a negative correlation between melanin content and percent yellow.

Of our eight candidate QTL regions, Yellow-1 was the only interval for which we did not find any hits to candidate pigmentation genes (Tables S5 and S7). Outside of these candidate regions, we found several additional hits in scaffolds that had high-PIP SNPs. For yellow, we identified a scaffold containing a predicted protein with a high degree of similarity to carotenoid isomerooxymerase, encoded by *NinaB* in *Drosophila* (Table S5). Work in *Drosophila* demonstrates that this protein in required for converting diet-derived carotenoids into visual pigments (Voolstra *et al.* 2010). For spotting, we found several additional candidate genes involved or potentially involved in melanin patterning, including *AbdB*, which encodes an Abdominal-B HOX protein that has been implicated in *Drosophila* pigmentation (Jeong *et al.* 2006), and several cytochrome p450s, one of which has been implicated in insect cuticle formation (Sztal *et al.* 2012) (Table S6).

Although all of our candidates require further fine-mapping and functional testing, we are encouraged to have identified multiple strong candidates for both traits. Notably, we report the first candidate genes for naturally occurring variation in carotenoid-based pigmentation in invertebrates. Additionally, some of our most promising candidate genes (*Cameo2, apoLTP-II/I, yellow-e,* and *ple*) fell just outside of the 1.5-LOD support intervals, but were included in the expanded candidate region on the basis of GEMMA and ALLMAPS results (Tables S5-S6). Should these candidates hold up to further experimental scrutiny, our findings suggest that combining QTL mapping and polygenic association mapping may be a fruitful approach for defining candidate regions.

Summary and Conclusions

In this study, we applied a combination of interval mapping and polygenic association mapping to describe the genetic architecture of two larval color traits: yellow pigmentation and spotting pattern. Both sets of analyses detect large-effect loci for both traits. Although we cannot yet rule out linked mutations of individually smaller effect, we discuss details of the demographic histories of these populations that may have favored the fixation of large-effect color alleles. Additionally, genetic architecture parameter estimates derived under the BSLMM implemented in GEMMA suggest that compared to the effect-size distribution underlying variation in spotting, the effect-size distribution for yellow is shifted towards alleles of larger effect. These findings are consistent with the pigmentation/patterning hypothesis, which argues that the genetic architecture of pigmentation should be "simpler" than that of patterning because pigmentation genes tend to be less pleiotropic than patterning genes. Verifying this hypothesis will require identifying quantitative trait nucleotides (QTNs) and assessing additional populations and species.

In addition to contrasting the genetic architecture of two color traits, we also identified several promising candidate genes that may contribute to natural variation in larval color. Although there are a growing number of studies of naturally occurring melanin-based pigmentation in adult insects (*e.g.*, Hof *et al.* 2016; Nadeau *et al.* 2016; Signor *et al.* 2016; Yassin *et al.* 2016), studies of naturally occurring larval color variation are sparse. Additionally, although carotenoids contribute to adaptive pigmentation in diverse animal taxa (Heath *et al.* 2013; Toews *et al.* 2017), ours is the first genetic study of naturally occurring variation in carotenoid-based pigmentation in a non-vertebrate. Thus, extensive intra- and interspecific variation in larval body color across the genus *Neodiprion* (Fig. 1) has the potential to provide novel insights into the molecular mechanisms underlying carotenoid-based pigmentation.

Finally, our study demonstrates the power of combining traditional interval mapping approaches with polygenic association mapping. Our combined approach not only enabled us to identify a surprisingly large number of promising candidate genes residing both within and outside of linkage-mapping-identified QTL, but also provided an intuitive way to describe and compare trait genetic architectures. We believe that this approach will prove valuable for testing additional theoretical predictions regarding trait genetic architectures. Ultimately, such studies will provide us with a more comprehensive understanding of the contribution of large-effect mutations to phenotypic evolution under different evolutionary scenarios.

ACKNOWLEDGEMENTS

For assistance with collecting and maintaining sawfly colonies, we thank past members of the Linnen laboratory, especially Robin Bagley, Adam Leonberger, Kim Vertacnik, Mary Collins, and Aubrey Mojesky. For assistance with phenotypic data collection, we thank Aubrey Mojesky, Mary Collins, and Ismaeel Siddiqqi. For advice on linkage mapping in R/qtl, we thank Karl Broman and Kelly O'Quin. For helpful comments on the manuscript, we thank Danielle Herrig, Emily Bendall, and Kim Vertacnik. Funding for

this research was provided by the National Science Foundation (DEB-1257739; CRL), the United States Department of Agriculture National Institute of Food and Agriculture (2016-67014-2475; CRL), the Academy of Finland (project no. 257581; CL), and the University of Kentucky (Ribble summer undergraduate research fellowship, honors program independent research grant, and summer research and creativity fellowship to TS).

REFERENCES

Altschul S. F., Gish W., Miller W., Myers E. W., Lipman D. J., 1990 Basic Local Alignment Search Tool. J. Mol. Biol.: 403–410.

Amézquita A., Ramos Ó., González M. C., Rodriguez C., Medina I., Simões P. I., Lima A. P., 2017 Conspicuousness, color resemblance, and toxicity in geographically diverging mimicry: The pan-Amazonian frog *Allobates femoralis.* Evolution 71: 1039-1050.

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

Bateson W., 1913 *Mendel's Principles of Heredity*. Cambridge University Press, Cambridge.

Baxter S. W., Johnston S. E., Jiggins C. D., 2009 Butterfly speciation and the distribution of gene effect sizes fixed during adaptation. Heredity 102: 57–65.

Bendall E. E., Vertacnik K. L., Linnen C. R., 2017 Oviposition traits generate extrinsic postzygotic isolation between two pine sawfly species. BMC Evol. Biol. 17: 26.

Benjamin D. M., 1955 The biology and ecology of the red-headed pine sawfly. USDA Tech. Bull. 118: 55p.

Bhosale P., Bernstein P. S., 2007 Vertebrate and invertebrate carotenoid-binding proteins. Arch. Biochem. Biophys. 458: 121–127.

Bickel R. D., Kopp A., Nuzhdin S. V, 2011 Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. PLoS Genet. 7: e1001275.

Broman K. W., Sen S., 2009 *A Guide to QTL Mapping with R/qtl*. Springer, New York.

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

Chen H., Rangasamy M., Tan S. Y., Wang H., Siegfried B. D., 2010 Evaluation of five methods for total DNA extraction from western corn rootworm beetles. PLoS One 5.

Codella S. G., Raffa K. F., 1993 Defense strategies of folivorous sawflies. In: Wagner M, Raffa KF (Eds.), *Sawfly Life History Adaptations to Woody Plants.* Academic Press Inc., San Diego, pp. 261–294.

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

Coppel H. C., Benjamin D. M., 1965 Bionomics of Nearctic pine-feeding diprionids. Annu. Rev. Entomol. 10: 69–96.

Crozier R. H., Fjerdingstad E. J., 2001 Polyandry in social Hymenoptera: disunity in diversity? Ann. Zool. Fennici 38: 267–285.

Danecek P., Auton A., Abecasis G., Albers C. A., Banks E., *et al.*, 2011 The variant call format and VCFtools. Bioinformatics 27: 2156–2158.

Darwin C., 1859 *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Davey J. W., Hohenlohe P. A., Etter P. D., Boone J. Q., Catchen J. M., *et al.*, 2011 Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. 12: 499–510.

Dittmar E. L., Oakley C. G., Conner J. K., Gould B., Schemske D., 2016 Factors influencing the effect size distribution of adaptive substitutions. Proc. R. Soc. B 283: 20153065.

Drapeau M. D., Albert S., Kucharski R., Prusko C., Maleszka R., 2006 Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of social behavior in honey bees. Genome Res.: 1385–1394.

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

Elsik C. G., Tayal A., Diesh C. M., Unni D. R., Emery M. L., *et al.*, 2016 Hymenoptera Genome Database: Integrating genome annotations in HymenopteraMine. Nucleic Acids Res. 44: D793–D800.

Ferguson L. C., Green J., Surridge A., Jiggins C. D., 2011 Evolution of the insect yellow gene family. Mol. Biol. Evol. 28: 257–272.

Fisher R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.

Frankel N., Erezyilmaz D. F., McGregor A. P., Wang S., Payre F., *et al.*, 2011 Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. Nature 474: 598–603.

Futahashi R., Fujiwara H., 2005 Melanin-synthesis enzymes coregulate stage-specific larval cuticular markings in the swallowtail butterfly, *Papilio xuthus*. Dev. Genes Evol. 215: 519–529.

Gadau J., Page R. E., Werren J. H., Schmid-Hempel P., 2000 Genome organization and social evolution in Hymenoptera. Naturwissenschaften 87: 87–9.

Gaj T., Gersbach C. A., Barbas C. F., 2013 ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31: 397–405.

Glastad K. M., Arsenault S. V., Vertacnik K. L., Geib S. M., Kay S., *et al.*, 2017 Variation in DNA methylation is not consistently reflected by sociality in Hymenoptera. Genome Biol. Evol. 9: 1687–1698.

Goldschmidt R., 1940 *The Material Basis of Evolution*. Yale University Press, New Haven.

Goodwin S., Mcpherson J. D., Mccombie W. R., 2016 Coming of age: ten years of next-generation sequencing technologies. Nat. Rev. Genet. 17: 333–351.

Griswold C. K., 2006 Gene flow's effect on the genetic architecture of a local adaptation and its consequences for QTL analyses. Heredity 96: 445–453.

Han Q., Fang J., Ding H., Johnson J. K., Christensen B. M., *et al.*, 2002 Identification of *Drosophila melanogaster* yellow-f and yellow-f2 proteins as dopachrome-conversion enzymes. Biochem. J. 368: 333–40.

Harper K. E., Bagley R. K., Thompson K. L., Linnen C. R., 2016 Complementary sex determination, inbreeding depression and inbreeding avoidance in a gregarious sawfly. Heredity 117: 326–335.

Heath J. J., Cipollini D. F., Stireman J. O., 2013 The role of carotenoids and their

derivatives in mediating interactions between insects and their environment. Arthropod. Plant. Interact. 7: 1–20.

Heimpel G. E., Boer J. G. de, 2008 Sex determination in the Hymenoptera. Annu. Rev. Entomol. 53: 209–30.

Hermisson J., Pennings P. S., 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352.

Hof A. E. van't, Campagne P., Rigden D. J., Yung C. J., Lingley J., *et al.*, 2016 The industrial melanism mutation in British peppered moths is a transposable element. Nature 534: 102–105.

Hoggart C. J., Whittaker J. C., Iorio M. De, Balding D. J., 2008 Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. PLoS Genet. 4: e1000130.

Huang A., Xu S., Cai X., 2015 Empirical Bayesian elastic net for multiple quantitative trait locus mapping. Heredity 114: 107–15.

Huang Y., Liu Z., Rong Y. S., 2016 Genome editing: From *Drosophila* to non-model insects and beyond. J. Genet. Genomics 43: 263–272.

Huxley T. H., 1860 On species, races and their origin. J. Proc. R. Inst. Gt. Britain 1858-1862 3: 151-153.

Ito K., Katsuma S., Yamamoto K., Kadono-Okuda K., Mita K., *et al.*, 2010 Yellow-e determines the color pattern of larval head and tail spots of the silkworm *Bombyx mori*. J. Biol. Chem. 285: 5624–5629.

Jeong S., Rokas A., Carroll S. B., 2006 Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. Cell 125: 1387–1399.

Joron M., Mallet J. L. B., 1998 Diversity in mimicry: Paradox or paradigm? Trends Ecol. Evol. 13: 461–466.

Kapan D. D., 2001 Three-butterfly system provides field test of Mullerian mimicry. Nature 409: 338–340.

Kimura M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Knerer G., Atwood C. E., 1972 Evolutionary trends in subsocial sawflies belonging to *Neodiprion abietis* complex (Hymenoptera: Tenthredinoidea). Am. Zool. 12: 407–418.

Knerer G., Atwood C. E., 1973 Diprionid sawflies: polymorphism and speciation. Science 179: 1090–1099.

Kopp A., 2009 Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. Evolution 63: 2771–2789.

Kraemer M. E., Coppel H. C., 1983 Hybridization of jack pine feeding sawflies (Diproinidae: *Neodiprion*). Forestry Research Notes. University of Wisconsin, Madison, WI.

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

Kronforst M. R., Barsh G. S., Kopp A., Mallet J., Monteiro A., *et al.*, 2012 Unraveling the thread of nature's tapestry: The genetics of diversity and convergence in animal pigmentation. Pigment Cell Melanoma Res. 25: 411–433.

Lee E., Helt G. A., Reese J. T., Munoz-Torres M. C., Childers C. P., *et al.*, 2013 Web

1111 Apollo: a web-based genomic annotation editing platform. Genome Biol. 14: R93.
1112 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., *et al.*, 2009 The Sequence
1113 Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.
1114 Li Z., Guo B., Yang J., Herczeg G., Gonda A., *et al.*, 2017 Deciphering the genomic
1115 architecture of the stickleback brain with a novel multilocus gene-mapping
1116 approach. Mol. Ecol. 26: 1557–1575.
1117 Lindstedt C., Carita L., Mappes J., Johanna M., Päivinen J., *et al.*, 2006 Effects of group
1118 size and pine defence chemicals on Diprionid sawfly survival against ant predation.
1119 Oecologia 150: 519–26.
1120 Lindstedt C., Lindstrom L., Mappes J., 2009 Thermoregulation constrains effective
1121 warning signal expression. Evolution 63: 469–478.
1122 Lindstedt C., Huttunen H., Kakko M., Mappes J., 2011 Disengtangling the evolution of
1123 weak warning signals: High detection risk and low production costs of chemical
1124 defences in gregarious pine sawfly larvae. Evol. Ecol. 25: 1029–1046.
1125 Lindstedt C., Schroderus E., Lindström L., Mappes T., Mappes J., 2016 Evolutionary
1126 constraints of warning signals: A genetic trade-off between the efficacy of larval and
1127 adult warning coloration can maintain variation in signal expression. Evolution 70:
1128 2562–2572.
1129 Linnen C. R., Farrell B. D., 2007 Mitonuclear discordance is caused by rampant
1130 mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies.
1131 Evolution 61: 1417–1438.
1132 Linnen C. R., Farrell B. D., 2008a Phylogenetic analysis of nuclear and mitochondrial
1133 genes reveals evolutionary relationships and mitochondrial introgression in the
1134 sertifer species group of the genus *Neodiprion* (Hymenoptera: Diprionidae). Mol.
1135 Phylogenet. Evol. 48: 240–257.
1136 Linnen C. R., Farrell B. D., 2008b Comparison of methods for species-tree inference in
1137 the sawfly genus *Neodiprion* (Hymenoptera: Diprionidae). Syst. Biol. 57: 876–90.
1138 Linnen C. R., Poh Y.-P., Peterson B. K., Barrett R. D. H., Larson J. G., *et al.*, 2013
1139 Adaptive evolution of multiple traits through multiple mutations at a single gene.
1140 Science 339: 1312–6.
1141 Lopes R. J., Johnson J. D., Toomey M. B., Ferreira M. S., Araujo P. M., *et al.*, 2016
1142 Genetic basis for red coloration in birds. Curr. Biol. 26: 1427–1434.
1143 Mackay T. F. C., Stone E. A., Ayroles J. F., 2009 The genetics of quantitative traits:
1144 challenges and prospects. Nat. Rev. Genet. 10: 565–577.
1145 Mallet J., Singer M. C., 1987 Individual selection, kin selection, and the shifting balance
1146 in the evolution of warning colours: the evidence from butterflies. Biol. J. Linn. Soc.
1147 32: 337–350.
1148 Manceau M., Domingues V. S., Linnen C. R., Rosenblum E. B., Hoekstra H. E., 2010
1149 Convergence in pigmentation at multiple levels: mutations, genes and function.
1150 Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365: 2439–2450.
1151 Martin A., Orgogozo V., 2013 The loci of repeated evolution: A catalog of genetic
1152 hotspots of phenotypic variation. Evolution 67: 1235–1250.
1153 Martin C. H., Erickson P. A., Miller C. T., 2017 The genetic architecture of novel trophic
1154 specialists: larger effect sizes are associated with exceptional oral jaw diversification
1155 in a pupfish adaptive radiation. Mol. Ecol. 26: 624–638.
1156 Matuszewski S., Hermisson J., Kopp M., 2014 Fisher's geometric model with a moving

optimum. Evolution 68: 2571–2588.

Matuszewski S., Hermisson J., Kopp M., 2015 Catch me if you can: Adaptation from standing genetic variation to a moving phenotypic optimum. Genetics 200: 1255–1274.

Maxwell D. E., 1958 Sawfly-cytology with emphasis upon Diprionidae (Hymenoptera, Symphyta). Proc 10th Int Congr Entomol. 2: 961–978.

Mayr E., 1982 *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Harvard University Press, Cambridge.

McGirr J. A., Martin C. H., Agashe D., 2017 Novel candidate genes underlying extreme trophic specialization in caribbean pupfishes. Mol. Biol. Evol. 34: 873–888.

McGraw K. J., Hudon J., Hill G. E., Parker R. S., 2005 A simple and inexpensive chemical test for behavioral ecologists to determine the presence of carotenoid pigments in animal tissues. Behav. Ecol. Sociobiol. 57: 391–397.

McGregor A. P., Orgogozo V., Delon I., Zanet J., Srinivasan D. G., *et al.*, 2007 Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature 448: 587–590.

Morgan T. H., 1932 *The Scientific Basis of Evolution*. Norton, New York.

Mundy N. I. I., Stapley J., Bennison C., Tucker R., Twyman H., *et al.*, 2016 Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene cluster. Curr. Biol. 26: 1435–1440.

Nadeau N. J., Jiggins C. D., 2010 A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. Trends Genet. 26: 484–492.

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

Ninomiya Y., Hayakawa Y., 2007 Insect cytokine, growth-blocking peptide, is a primary regulator of melanin-synthesis enzymes in armyworm larval cuticle. FEBS J. 274: 1768–1777.

Nokelainen O., Valkoen J., Lindstedt C., Mappes J., 2014 Changes in predator community structure shifts the efficacy of two warning signals in Arctiid moths. J. Anim. Ecol. 83: 598-605.

O'Quin C. T., Drilea A. C., Conte M. A., Kocher T. D., 2013 Mapping of pigmentation QTL on an anchored genome assembly of the cichlid fish, *Metriaclima zebra*. BMC Genomics 14: 287.

Orr H. A., Coyne J. A., 1992 The genetics of adaptation: a reassessment. Am. Nat. 140: 725–742.

Orr H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.

Orr H. A., 2005 The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119–127.

Otto S. P., Jones C. D., 2000 Detecting the undetected: estimating the total number of loci inderlying a quantitative trait. Genetics 156: 2093–2107.

Palm W., Sampaio J. L., Brankatschk M., Carvalho M., Mahmoud A., *et al.*, 2012 Lipoproteins in *Drosophila melanogaster*: assembly, function, and influence on tissue lipid composition. PLoS Genet. 8.

Pearson K., 1897 Mathematical contributions to the theory of evolution. On the law of

1203 ancestral heredity. Proc. R. Soc. London 62: 386–412.
1204 Peterson B. K., Weber J. N., Kay E. H., Fisher H. S., Hoekstra H. E., 2012 Double digest
1205 RADseq: an inexpensive method for de novo SNP discovery and genotyping in
1206 model and non-model species. PLoS One 7: e37135.
1207 Poelchau M., Childers C., Moore G., Tsavatapalli V., Evans J., *et al.*, 2014 The i5k
1208 Workspace@NAL: enabling genomic data access, visualization and curation of
1209 arthropod genomes. Nucleic Acids Res. 43: D714–D719.
1210 Protas M. E., Patel N. H., 2008 Evolution of coloration patterns. Annu. Rev. Cell Dev.
1211 Biol. 24: 425–446.
1212 Prud'homme B., Gompel N., Rokas A., Kassner V. A., Williams T. M., *et al.*, 2006
1213 Repeated morphological evolution through cis-regulatory changes in a pleiotropic
1214 gene. Nature 440: 1050–1053.
1215 Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A. R., *et al.*, 2007 PLINK:
1216 A tool set for whole-genome association and population-based linkage analyses. 81:
1217 559–575.
1218 Rebeiz M., Pool J. E., Kassner V. A., Aquadro C. F., Carroll S. B., 2009 Stepwise
1219 modification of a modular enhancer underlies adaptation in a *Drosophila* population.
1220 Science 326: 1663–1667.
1221 Remington D. L., 2015 Alleles versus mutations: Understanding the evolution of genetic
1222 architecture requires a molecular perspective on allelic origins. Evolution 69: 3025–
1223 3038.
1224 Rieseberg L. H., Archer M. A., Wayne R. K., 1999 Transgressive segregation, adaptation
1225 and speciation. Heredity 83: 363-372.
1226 Rockman M. V., 2012 The QTN program and the alleles that matter for evolution: All
1227 that's gold does not glitter. Evolution 66: 1–17.
1228 Rogers S. M., Tamkee P., Summers B., Balabahadra S., Marks M., *et al.*, 2012 Genetic
1229 signature of adaptive peak shift in threespine stickleback. Evolution 66: 2439–2450.
1230 Roshyara N. R., Scholz M., 2014 fcGENE: A versatile tool for processing and
1231 transforming SNP datasets. PLoS One 9.
1232 Ruxton G. D., Sherratt T. N., Speed M. P., 2004 *Avoiding Attack: The Evolutionary*
1233 *Ecology of Crypsis, Warning Signals and Mimicry*. Oxford University Press, New
1234 York.
1235 Sakudoh T., Iizuka T., Narukawa J., Sezutsu H., Kobayashi I., *et al.*, 2010 A CD36-
1236 related transmembrane protein is coordinated with an intracellular lipid-binding
1237 protein in selective carotenoid transport for cocoon coloration. J. Biol. Chem. 285:
1238 7739–7751.
1239 Scheet P., Stephens M., 2006 A fast and flexible statistical model for large-scale
1240 population genotype data: applications to inferring missing genotypes and
1241 haplotypic phase. Am. J. Hum. Genet. 78: 629–44.
1242 Schmid-Hempel P., 2000 Mating, parasites and other trials of life in social insects.
1243 Microbes Infect. 2: 515–520.
1244 Signor S. A., Liu Y., Rebeiz M., Kopp A., 2016 Genetic convergence in the evolution of
1245 male-specific color patterns in *Drosophila*. Curr. Biol. 26: 2423–2433.
1246 Skinner M. E., Uzilov A. V, Stein L. D., Mungall C. J., Holmes I. H., 2009 JBrowse: a
1247 next-generation genome browser. Genome Res 19: 1630–1638.
1248 Smith S. G., 1941 A new form of spruce sawfly identified by means of its cytology and

1249 parthenogenesis. Sci. Agric. 21: 245–305.
1250 Sohi S. S., Ennis T. J., 1981 Chromosomal characterization of cell lines of *Neodiprion*
1251 *lecontei* (Hymenoptera: Diprionidae). Proc. Entomol. Soc. Ontario 112: 45–48.
1252 Stam L. F., Laurie C. C., 1996 Molecular dissection of a major gene effect on a
1253 quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila*
1254 *melanogaster*. Evolution 1564: 1559–1564.
1255 Sugumaran M., Barek H., 2016 Critical analysis of the melanogenic pathway in insects
1256 and higher animals. Int. J. Mol. Sci. 17: 1–24.
1257 Sztal T., Chung H., Berger S., Currie P. D., Batterham P., *et al.*, 2012 A cytochrome
1258 P450 conserved in insects is involved in cuticle formation. PLoS One 7: 1–9.
1259 Tang H., Zhang X., Miao C., Zhang J., Ming R., *et al.*, 2015 ALLMAPS: robust scaffold
1260 ordering based on multiple maps. Genome Biol. 16: 3.
1261 Taylor J., Butler D., 2017 R Package ASMap: Efficient genetic linkage map construction
1262 and diagnosis. J. Stat. Softw. 79: 1–29.
1263 R Core Team, 2013 R: A language and environment for statistical computing.
1264 Toews D. P. L., Hofmeister N. R., Taylor S. A., 2017 The evolution and genetics of
1265 carotenoid processing in animals. Trends Genet. 33: 171–182.
1266 True J. R., 2003 Insect melanism: The molecules matter. Trends Ecol. Evol. 18: 640–647.
1267 Tsuchida K., Arai M., Tanaka Y., Ishihara R., Ryan R. O., *et al.*, 1998 Lipid transfer
1268 particle catalyzes transfer of carotenoids between lipophorins of *Bombyx mori*.
1269 Insect Biochem. Mol. Biol. 28: 927–934.
1270 Tsuchida K., Sakudoh T., 2015 Recent progress in molecular genetic studies on the
1271 carotenoid transport system using cocoon-color mutants of the silkworm. Arch.
1272 Biochem. Biophys. 572: 151–157.
1273 Turner J. R. G., Mallet J. L. B., 1996 Did forest islands drive the diversity of warningly
1274 coloured butterflies? Biotic drift and the shifting balance. Philos. Trans. R. Soc. B
1275 Biol. Sci. 351: 835–845.
1276 Turner S. D., 2014 qqman: an R package for visualizing GWAS results using QQ and
1277 manhattan plots. BioRxiv: 5165.
1278 Vertacnik K. L., Geib S., Linnen C. R., 2016 *Neodiprion lecontei* genome assembly
1279 Nlec1.0. NCBI/GenBank. http://dx.doi.org/10.15482/USDA.ADC/1235565.
1280 Voolstra O., Oberhauser V., Sumser E., Meyer N. E., Maguire M. E., *et al.*, 2010 NinaB
1281 is essential for Drosophila vision but induces retinal degeneration in opsin-deficient
1282 photoreceptors. J. Biol. Chem. 285: 2130–2139.
1283 Werren J. H., Richards S., Desjardins C. A., Niehuis O., Gadau J., *et al.*, 2010 Functional
1284 and evolutionary insights from the genomes of three parasitoid *Nasonia* species.
1285 Science 327: 343-348.
1286 Wilfert L., Gadau J., Schmid-Hempel P., 2007 Variation in genomic recombination rates
1287 among animal taxa and the case of social insects. Heredity 98: 189–197.
1288 Willmott K. R., Willmott J. C. R., Elias M., Jiggins C. D., 2017 Maintaining mimicry
1289 diversity: optimal warning colour patterns differ among microhabitats in Amazonian
1290 clearwing butterflies. Proc. R. Soc. B 284: 20170744.
1291 Wilson L. F., Wilkinson R. C., Averill R. C., 1992 *Redheaded Pine Sawfly: Its Ecology*
1292 *and Management*. USDA, Washington DC.
1293 Wittkopp P. J., Vaccaro K., Carroll S. B., 2002 Evolution of *yellow* gene regulation and
1294 pigmentation in *Drosophila*. Curr. Biol. 12: 1547–1556.

Wittkopp P. J., Carroll S. B., Kopp A., 2003 Evolution in black and white: Genetic control of pigment patterns in *Drosophila*. Trends Genet. 19: 495–504.

Wittkopp P. J., Beldade P., 2009 Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. Semin. Cell Dev. Biol. 20: 65–71.

Wright T. R. F., 1987 The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. Adv. Genet. 24: 127–222.

Yassin A., Delaney E. K., Reddiex A. J., Seher T. D., Bastide H., *et al.*, 2016 The *pdm3* locus is a hotspot for recurrent evolution of female-limited color dimorphism in *Drosophila*. Curr. Biol. 26: 2412–2422.

Yeaman S., Whitlock M. C., 2011 The genetic architecture of adaptation under migration-selection balance. Evolution 65: 1897–1911.

Yi N., Banerjee S., 2009 Hierarchical generalized linear models for multiple quantitative trait locus mapping. Genetics 181: 1101–1113.

Yokoyama H., Yokoyama T., Yuasa M., Fujimoto H., Sakudoh T., *et al.*, 2013 Lipid transfer particle from the silkworm, *Bombyx mori*, is a novel member of the apoB/large lipid transfer protein family. J. Lipid Res. 54: 2379–90.

Zhou X., Carbonetto P., Stephens M., 2013 Polygenic Modeling with Bayesian Sparse Linear Mixed Models. PLoS Genet. 9.

Figure 1. Interspecific variation in *Neodiprion* larval color. Top row (left to right): *Neodiprion nigroscutum*, *N. rugifrons*, *N. virginianus*. Middle row (left to right): *N. pinetum*, *N. lecontei*, *N. merkeli*. Bottom row (left to right): *N. pratti*, *N. compar*, *N. swainei*. Larvae in the first and last columns are exhibiting a defensive "U-bend" posture (a resinous regurgitant is visible in *N. virginianus*, top right). *N. pratti* photo is by K. Vertacnik, all others are by R. Bagley.

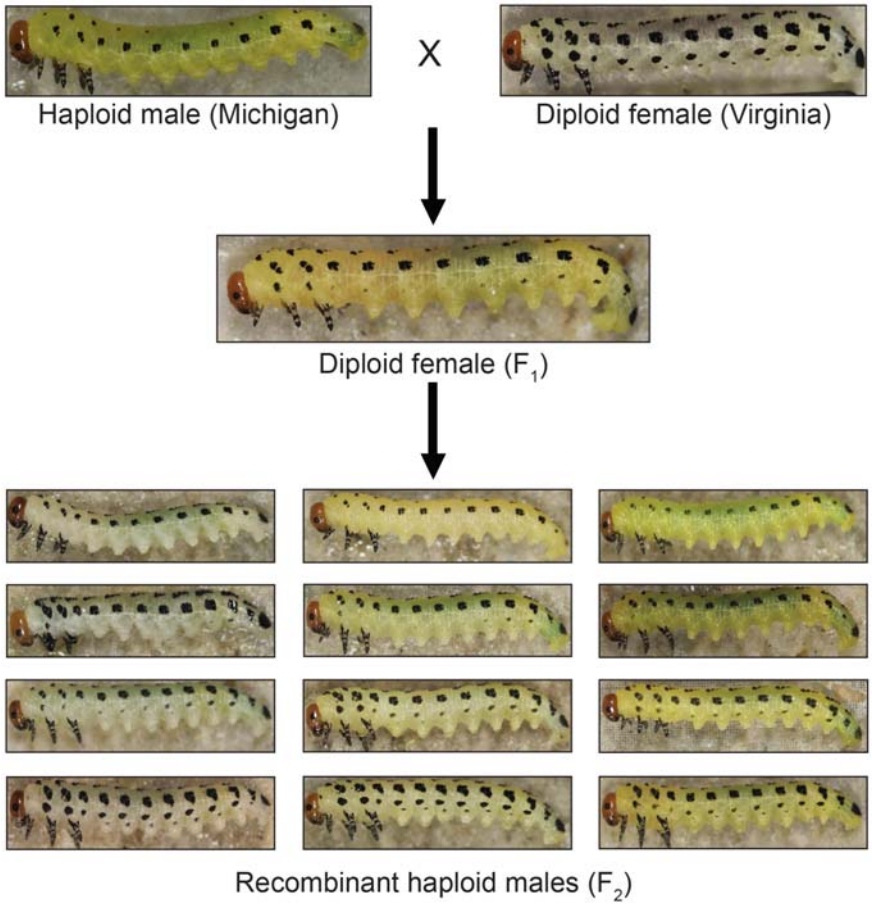

Figure 2. Intraspecific variation in *Neodiprion lecontei* larval color and cross design. We crossed white, dark-spotted diploid females from Virginia to yellow, light-spotted haploid males from Michigan. This produced haploid males with the VA genotype and phenotype (not shown) and diploid females (F_1) with intermediate spotting and color. Virgin F_1 females produced recombinant haploid males (F_2) with a wide range of body color and spotting pattern (a representative sample is shown).

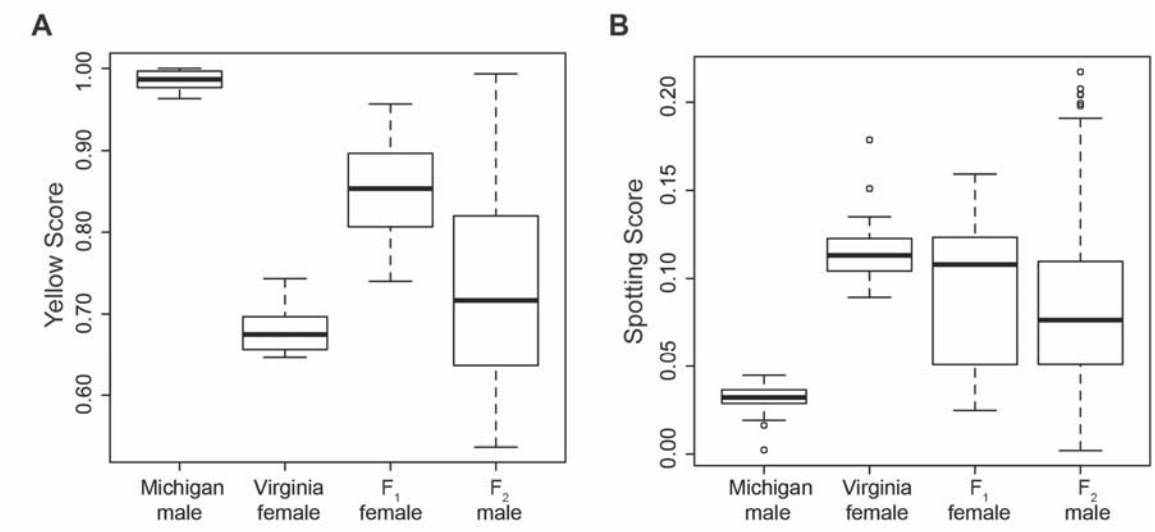

Figure 3. Larval color variation across generations. Variation in larval body color (A; higher yellow scores indicate more yellow pigment) and spotting pattern (B; higher spotting scores indicate more melanic spotting), both measured from digital images as described in the text. Boxes represent interquartile ranges (median ± 2 s.d.), with outliers indicated as points. All comparisons were significantly different after correction for multiple testing (adjusted $\alpha = 0.0042$) except F_1 female vs. F_2 male spotting score (see text).

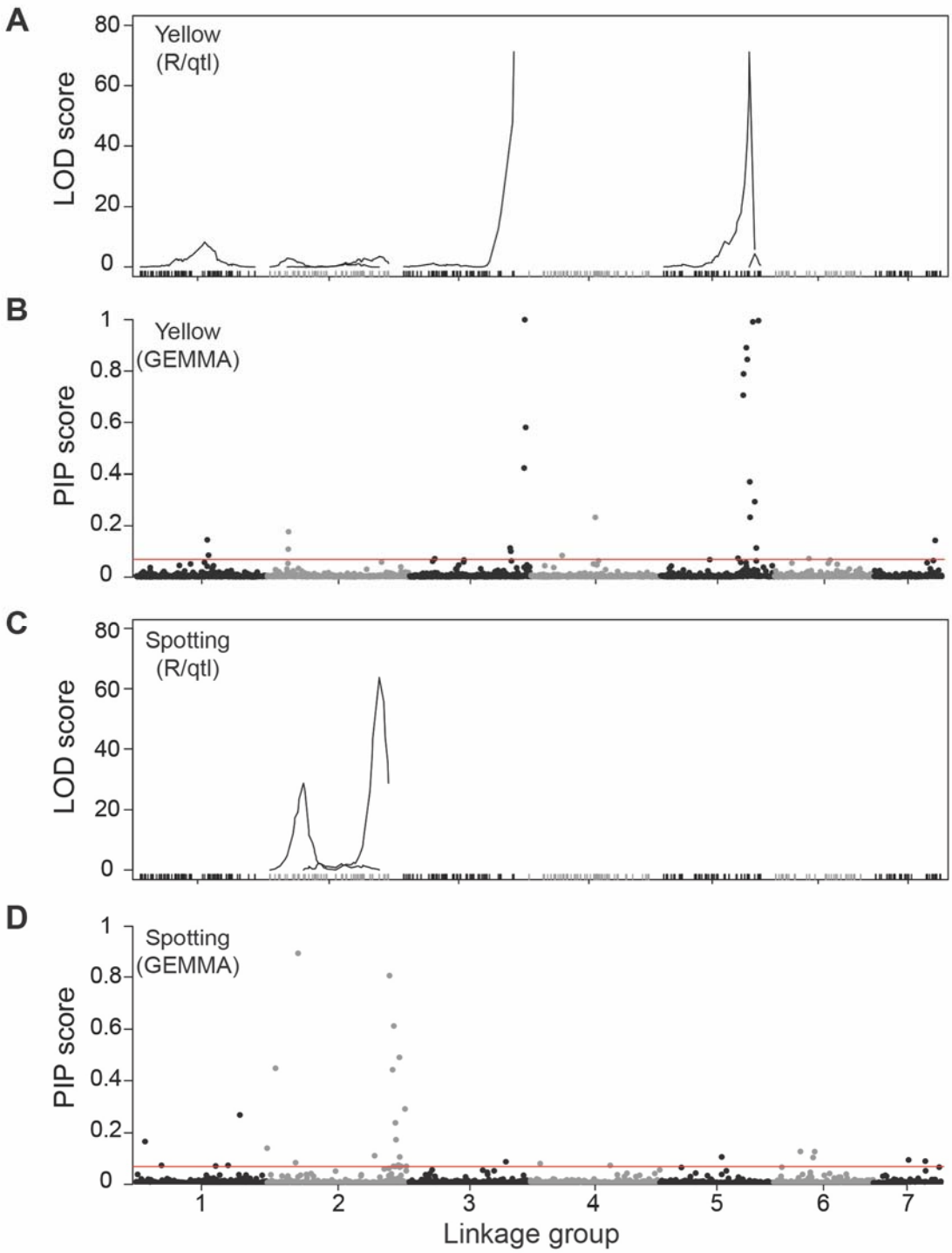

Figure 4. Linkage mapping and polygenic association mapping results. Linkage mapping analyses recovered QTL for larval body color (Yellow) on linkage groups (LGs) 1, 2,3, and 5 (A) and QTL for spotting pattern (Spotting) on LG 2 (B). These same regions were recovered in polygenic association mapping analyses as SNPs with high posterior inclusion probabilities (PIPs) (B and D; results shown are for the SNP dataset with <50% missing data). Polygenic association mapping also recovered additional high-PIP SNPs (99th percentile threshold is indicated as a horizontal line in B and D).

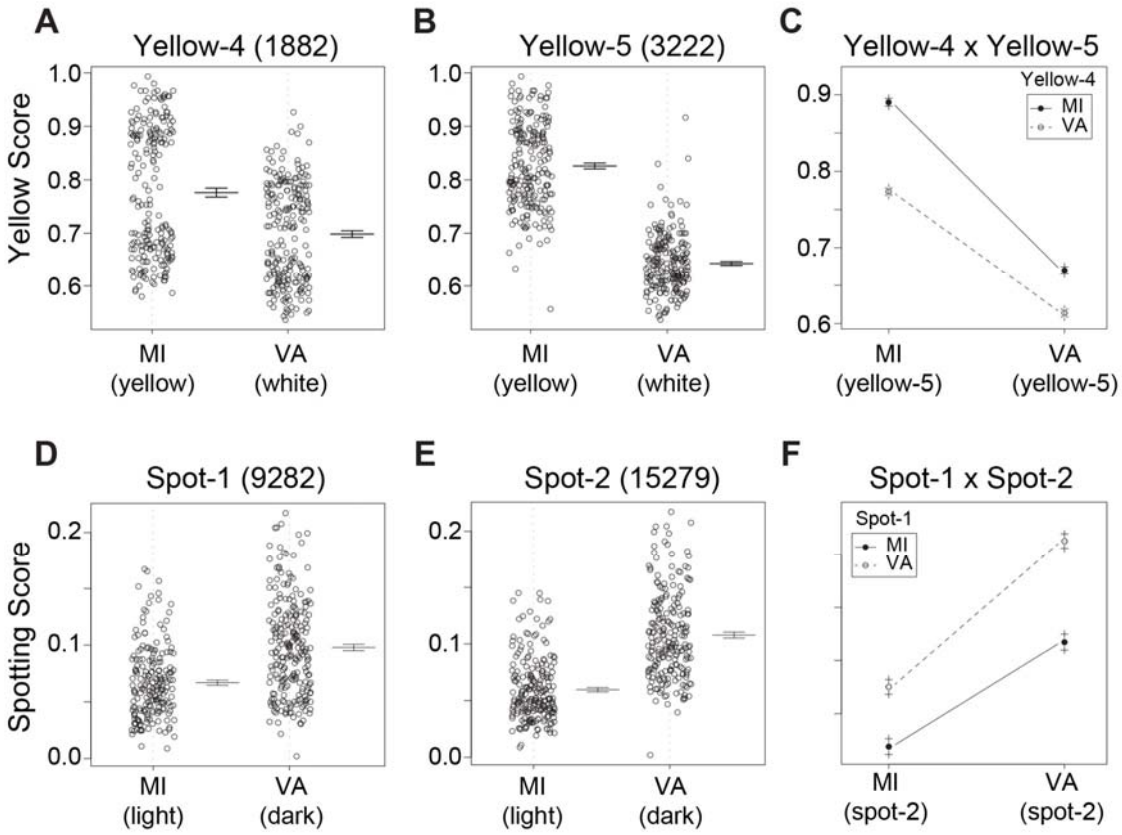

Figure 5. Effect and interaction plots for larval body color and spotting pattern. For each trait, phenotypic effect plots are given for the SNPs closest to the QTL peak for each of the two largest-effect QTL (names are as in Table 1). Yellow-4 (A) and Yellow-5 (B) are from LG 3 and LG 5, respectively. Spot-1 (D) and Spot-2 (E) are from LG 2. For both traits, the two largest-effect QTL also have a significant interaction term (Table 1). In both cases (C and F), the magnitude (but not the direction) of the allelic effects at one locus depends on the genotype of the interacting locus.

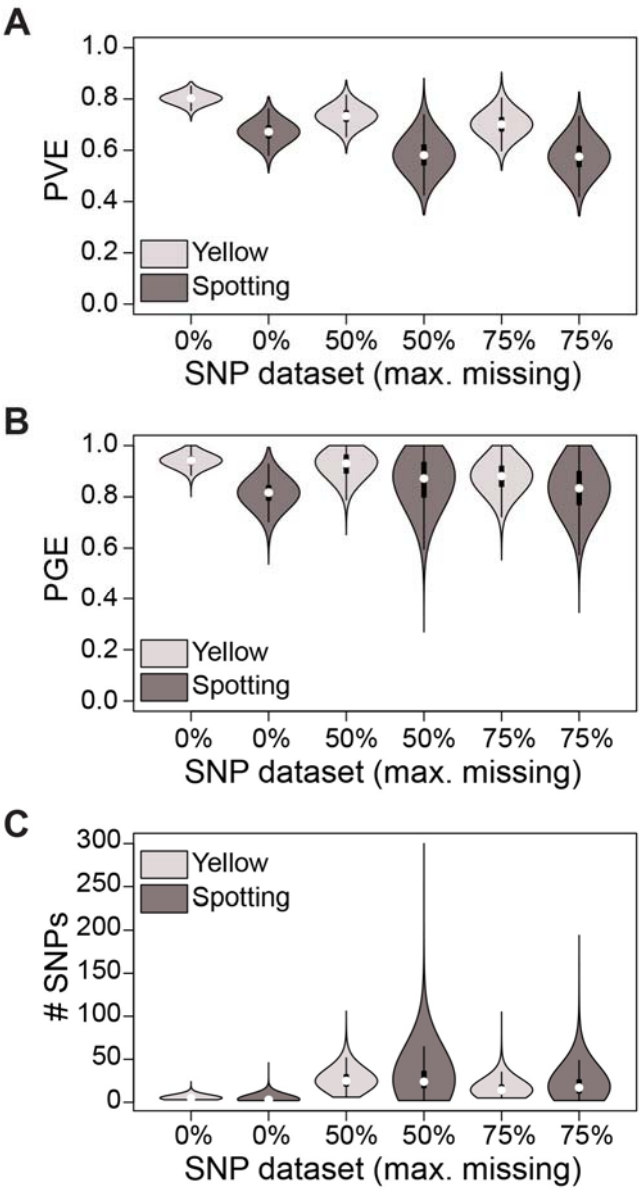

Figure 6. Genetic architecture parameter estimates for larval pigmentation (yellow) and larval color pattern (spotting). Violin plots display posterior probability distributions for genetic architecture parameters estimated under the BSLMM implemented in GEMMA. Genetic architecture parameters include: (A) PVE= total proportion of phenotypic variance explained by genetic variance; (B) PGE = total proportion of genetic variance explained by sparse (major) effects; and (C) # SNPS ("n_gamma" in GEMMA output) = number of sparse (major) effect SNPs. SNP dataset refers to the maximum percent missing data in the SNP dataset analyzed (0%: 1205 SNPs; 50% 3070 SNPs; 75%: 4162 SNPs). Distributions are combined across 10 independent runs, each consisting of 20 million post-burnin generations sampled every 1,000 generations. Violin plots depict probability densities (colored area), medians (white dot), interquartile ranges (thick black line), and 95% confidence intervals (thin black lines). For all datasets, PVE and PGE are consistently higher for yellow than for spotting (see also Table 2).

Table 1. QTL locations and effect sizes for larval pigmentation (Yellow) and larval color pattern (Spotting).

Trait	QTL name	LG[*]	Position (interval)[†]	Marker[‡]	LOD	PVE	Effect size[§] (SE)	% difference[**]
Yellow	Yellow-1	1	105.9 (84.4-112.5)	5744	8.29	1.32	-0.025 (0.0043)	8.06
Yellow	Yellow-2	2	28.2 (16.2-55.1)	19162	2.92	0.45	-0.014 (0.0043)	4.63
Yellow	Yellow-3	2	179.9 (140.8-194.0)	15279	3.51	0.54	-0.016 (0.0043)	5.13
Yellow	Yellow-4	3	181.9 (179.8-181.9)	1882	71.23	16.27	-0.084 (0.0043)	27.19
Yellow	Yellow-5	5	140.6 (139.9-149.6)	3222	71.17	16.25	-0.16 (0.0081)	51.64
Yellow	Yellow-6	5	149.6 (140.6-154.8)	19661	4.37	0.68	-0.035 (0.0082)	11.43
Yellow		Interaction	Yellow-3 x Yellow-5		12.45	2.03	0.061 (0.0086)	19.58
Spotting	Spot-1	2	55.1 (48.0-57.6)	9282	28.62	12.99	0.029 (0.0025)	34.86
Spotting	Spot-2	2	179.9 (169.1-187.5)	15279	63.73	35.47	0.048 (0.0024)	57.04
Spotting		Interaction	Spot-1 x Spot-2		1.76	0.69	0.014 (0.0049)	16.49

[*] Linkage group number

[†] Position in cM (1.5-LOD support intervals)

[‡] Marker closest to QTL peak

[§] Effect sizes as the difference in the phenotypic averages of among F_2 males carrying a VA allele and F_2 males carrying a MI allele (± standard error).

[**] Effect sizes as a percentage of the difference between average trait values for the two grandparental lines (VA and MI).

Table 2. Effect size and genetic architecture parameter estimates for larval pigmentation (Yellow) and larval color pattern (Spotting).

Parameter*	% miss.†	Yellow‡	Spotting‡
effect size (maximum)	0%	0.69	0.37
effect size (top 1%)	0%	0.023 (±0.088)	0.013 (±0.014)
PVE	0%	0.80 (0.77,0.83)	0.67 (0.60,0.73)
PGE	0%	0.94 (0.90,0.98)	0.82 (0.72,0.89)
# SNPs	0%	5 (3,10)	3 (2,9)
effect size (maximum)	50%	0.45	0.33
effect size (top 1%)	50%	0.042 (±0.15)	0.03 (±0.055)
PVE	50%	0.73 (0.67,0.79)	0.58 (0.47,0.70)
PGE	50%	0.93 (0.82,1)	0.87 (0.66,0.99)
# SNPs	50%	25 (12,52)	24 (8,75)
effect size (maximum)	75%	0.57	0.37
effect size (top 1%)	75%	0.015 (±0.031)	0.013 (±0.016)
PVE	75%	0.70 (0.63,0.78)	0.58 (0.47,0.69)
PGE	75%	0.88 (0.76,0.98)	0.83 (0.65,0.99)
# SNPs	75%	14 (6,35)	17 (4,56)

* Genetic architecture parameter estimates for the BSLMM implemented in GEMMA are as follows: PVE= total proportion of phenotypic variance explained by genetic variance; PGE = total proportion of genetic variance explained by sparse (major) effects; # SNPS ("n_gamma" in GEMMA output) = number of sparse (major) effect SNPs.

† Maximum percent missing data in the SNP dataset (0%: 1205 SNPs; 50% 3070 SNPs; 75%: 4162 SNPs).

‡ For PVE, PGE, and # SNPS, medians (and 95% credible intervals) are reported for the combined posterior distributions of 10 independent GEMMA runs (results for individual runs are given in Table S4). For the top 1% SNPs, medians (and interquartile ranges) are reported.