Title: Genetic basis of body color and spotting pattern in redheaded pine sawfly larvae (*Neodiprion lecontei*)

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

Running title: Genetic basis of larval color

Keywords: pigmentation, carotenoids, melanin, genetic architecture, convergent evolution, 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

Pigmentation has emerged as a premier model for understanding the genetic basis of phenotypic evolution, and a growing catalog of color loci is starting to reveal biases in the mutations, genes, and genetic architectures underlying color variation in the wild. However, existing studies have sampled a limited subset of taxa, color traits, and developmental stages. To expand our sample of color loci, we performed quantitative trait locus (QTL) mapping analyses on two types of larval pigmentation traits that vary among populations of the redheaded pine sawfly (*Neodiprion lecontei*): carotenoid-based yellow body color and melanin-based spotting pattern. For both traits, our QTL models explained a substantial proportion of phenotypic variation and suggested a genetic architecture that is neither monogenic nor highly polygenic. Additionally, we used our linkage map to anchor the current *N. lecontei* genome assembly. With these data, we identified promising candidate genes underlying: (1) a loss of yellow pigmentation in Mid-Atlantic/northeastern populations (*Cameo2* and *apoLTP-II/I*), and (2) a pronounced reduction in black spotting in Great-Lakes populations (*yellow, TH, Dat*). Several of these genes also contribute to color variation in other wild and domesticated taxa. Overall, our findings are consistent with the hypothesis that predictable genes of large-effect contribute to color evolution in nature.

INTRODUCTION

Over the last century, color phenotypes have played a central role in our efforts to understand how evolutionary processes shape phenotypic variation in natural populations (Gerould 1921; Sumner 1926; Fisher and Ford 1947; Haldane 1948; Cain and Sheppard 1954; Kettlewell 1955). More recently, as technological advances have enabled us to link genotypes and phenotypes in non-model organisms (Davey *et al.* 2011; Gaj *et al.* 2013; Goodwin *et al.* 2016; Huang *et al.* 2016), we've begun to decipher the genetic and developmental mechanisms underlying naturally occurring color variation as well (True 2003; Protas and Patel 2008; Wittkopp and Beldade 2009; Manceau *et al.* 2010; Nadeau and Jiggins 2010; Kronforst *et al.* 2012). With a growing catalog of color loci, we are starting to assess how ecology, evolution, and development interact to bias the genetic architectures, genes, and mutations underlying the remarkable diversity of color phenotypes in nature (Hoekstra and Coyne 2007; Stern and Orgogozo 2008, 2009; Kopp 2009; Manceau *et al.* 2010; Streisfeld and Rausher 2011; Martin and Orgogozo 2013; Dittmar *et al.* 2016; Massey and Wittkopp 2016; Martin and Courtier-Orgogozo 2017). To make robust inferences about color evolution, however, we require genetic data from diverse traits, taxa, developmental stages, and evolutionary timescales.

Two long-term goals of research on the genetic underpinnings of pigment variation are: (1) to evaluate the importance of large-effect loci to phenotypic evolution (Orr and Coyne 1992; Mackay *et al.* 2009; Rockman 2012; Remington 2015; Dittmar *et al.* 2016), and (2) to determine the extent to which the independent evolution of similar phenotypes (phenotypic convergence) is attributable to the same genes and mutations (genetic convergence) (Arendt and Reznick 2008; Gompel and Prud'homme 2009; Christin *et al.* 2010; Manceau *et al.* 2010; Elmer and Meyer 2011; Conte *et al.* 2012; Rosenblum *et al.* 2014). Addressing these questions will require a large, unbiased sample

of pigmentation loci. At present, most of what we know about the genetic basis of color variation in animals comes from a handful of taxa and from melanin-based color variation in adult life stages (True 2003; Protas and Patel 2008; Wittkopp and Beldade 2009; Nadeau and Jiggins 2010; Kronforst *et al.* 2012; Sugumaran and Barek 2016). Another important source of bias in our existing sample of pigmentation loci stems from a tendency to focus on candidate genes and discrete pigmentation phenotypes (Kopp 2009; Manceau *et al.* 2010; Rockman 2012; but see O'Quin *et al.* 2013; Albertson *et al.* 2014; Signor *et al.* 2016; Yassin *et al.* 2016). Here, we describe an unbiased, genome-wide analysis of multiple, continuously varying color traits in larvae from the order Hymenoptera, a diverse group of insects that is absent from our current catalog of color loci (Martin and Orgogozo 2013).

More specifically, our study focuses on pine sawflies in the genus *Neodiprion*. Several factors make *Neodiprion* an especially promising and tractable system for investigating the genetic basis of color variation. First, there is extensive variation in many different types of color traits, and these vary both within and between *Neodiprion* species (Figures 1-2). Second, previous phylogenetic and demographic studies of this genus (Linnen and Farrell 2007, 2008a; b; Bagley *et al.* 2017) enable us to infer directions of trait change and identify instances of phenotypic convergence. Third, because many different species can be reared and crossed in the lab (Knerer and Atwood 1972, 1973; Kraemer and Coppel 1983; Bendall *et al.* 2017), unbiased genetic mapping approaches are feasible in *Neodiprion*. Fourth, a growing list of genomic resources for *Neodiprion*—including an annotated genome and a methylome for *N. lecontei* (Vertacnik *et al.* 2016; Glastad *et al.* 2017)— facilitate identification of causal genes and mutations. And finally, an extensive natural history literature (Coppel and Benjamin 1965; Knerer and Atwood 1973) provides insights into the ecological functions of color variation in pine sawflies, which we review briefly for context.

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). When threatened, larvae adopt a characteristic "U-bend" posture and regurgitate a resinous defensive fluid (Figure 1), which 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 are chemically defended, larvae 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; Lindstedt *et al.* 2011).

Beyond selection for crypsis or aposematism, there are a myriad of abiotic and biotic selection pressures that could act on *Neodiprion* larval color. 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). Color alleles may also have 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). Temporal and spatial variation in

these diverse selection pressures likely contribute to the abundant color variation in the genus *Neodiprion*.

As a first step to understanding the proximate mechanisms that generate color variation in pine sawflies, we conducted a quantitative trait locus (QTL) mapping study of larval body color and larval spotting pattern in the redheaded pine sawfly, *Neodiprion lecontei*. This species is widespread across eastern North America, where it feeds on multiple pine species. Throughout most of this range, larvae have several rows of dark black spots overlaid on a bright, yellow body (*e.g.*, center image in Figure 1). However, previous field surveys and demographic analyses indicate that there has been a loss of yellow pigmentation in some populations in the Mid-Atlantic/northeastern United States and a pronounced reduction in spotting in populations living in the Great-Lakes region of the U.S. and Canada (Bagley *et al.* 2017; Figure 2). To describe genetic architectures and identify candidate genes underlying these two reduced-pigmentation phenotypes, we crossed white-bodied, dark-spotted individuals from a Virginia population to yellow-bodied, light-spotted individuals from a Michigan population (Figure 2). After mapping larval body-color and spotting-pattern traits in recombinant F_2 haploid males, we used our linkage map to anchor the current *N. lecontei* genome assembly and identified candidate color genes within our QTL intervals. We conclude by comparing these initial findings on color genetics in pine sawflies with published studies from other insect taxa.

MATERIALS AND METHODS

Cross Design

To investigate the genetic basis of 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 were collected from the field in 2012 and reared on *Pinus banksiana* (jack pine) for at least two 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.

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 (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 pairs of rows of black spots (Wilson *et al.* 1992). Immediately after phenotyping, we 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).

Larval body color

We quantified larval body color in two ways: spectrometry and digital photography. First, we immobilized larvae with CO_2 and recorded 15 reflectance spectra—five measurements from each of three body regions: the dorsum, lateral side, and ventrum—using a USB2000 spectrophotometer with a PX-2 Pulsed Xenon light source and SpectraSuite software (Ocean Optics, Largo, FL). We then used the program CLR: Colour Analysis Programs v1.05 (Montgomerie 2008) to trim the raw reflectance data to 300-700nm, compute 1-nm bins, and to compute the following summary statistics (described in more detail in the CLR documentation): B1, B2, B3, S1R, S1G, S1B, S1U, S1Y, S1V, S3, S5a, S5b, S5c, S6, S7, S8, S9, H1, H3, H4a, H4b, and H4c. This procedure produced 15 estimates for each summary statistic for each larva, which we then averaged to obtain a single value per statistic per larva. To reduce the 22 color summary statistics to a smaller number of independent variables, we performed a principal component analysis (PCA). To ensure that each variable was normally distributed, we performed a normal-quantile transformation prior to performing the PCA. Based on examination of the resulting scree plot, we retained the first two principal components (PC1 and PC2), which together accounted for 73.7% of the variance in the spectral data. Based on factor loadings, PC1 (PVE: 48.2%) corresponds to saturation (S1G, S1Y, S8) and brightness (B1, B2), while PC2 (PVE: 25.5%) corresponds to saturation (S5A, S5B, S5C, S6) and hue (H4c) (Table S1). Unless otherwise noted, the PCA and all other statistical analyses were performed in R (Version 3.1.3 or 3.2.3).

Because it provides objective and information-rich data, spectrometry is generally considered to be the gold standard for color quantification (Endler 1990; Andersson and Prager 2006). However, despite efforts to keep background lighting and reflectance probe positioning as consistent as possible across samples, measurement of small patches of yellow on rounded larval bodies proved challenging. Given the potential for variation in reflectance probe position to introduce substantial noise into our data (Montgomerie 2006), we used digital photography as a complementary method for quantifying larval body color (Stevens *et al.* 2007). For this analysis, we photographed CO_2-immobilized larvae (dorsal and lateral surfaces) with a Canon EOS Rebel t3i camera equipped with an Achromat S 1.0X FWD 63mm lens and mounted on a copy stand. All photographs were taken in a dark, windowless room, and larvae were illuminated by two SLS CL-150 copy lamp lights, each with a 23-Watt (100W equivalent) soft white compact fluorescent light bulb. We then 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 (lateral surface) 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").

Larval spotting pattern

On each side, *N. lecontei* larvae have up to four rows of up to 11 spots extending from the mesothorax to the 9th abdominal segment (Wilson *et al.* 1992). We quantified the extent of this black spotting in two ways: the number of spots present (spot number) and the proportion of body area covered by spots (spot area). To quantify spot number, we simply counted the number of spots in digital images of the dorsal and lateral surfaces (one side only). To quantify larval spotting area, 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 also used the larval images and Adobe Photoshop 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 in File S3).

Statistical analysis of phenotypic data

In total, we produced three measures of body color (PC1, PC2, and yellow) and two measures of spotting pattern (spot number and spot area). To determine whether mean phenotypic values for the five color traits differed between the two populations and among the three generations of our cross, we performed Welch's two-tailed *t*-tests, followed by Bonferroni correction for multiple comparisons. To determine the extent to which these five traits co-varied in the F_2 males, we calculated Pearson's correlation coefficients (r) and associated P-values, followed by Bonferroni correction for multiple comparisons. To determine which covariates to include in our QTL models, we performed ANOVAs to evaluate the relationship between each phenotype 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 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 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 de-multiplexed 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 QTL mapping analyses, our goal was to recover fixed differences between the grandparental lines. To do so, we first called SNPs in our eight

243 grandparents and 10 F_1 mothers. For these 18 individuals, we required that SNPs had a
244 minimum of 7x coverage and no more than 12% missing data. We then compiled a list of
245 SNPs that represented fixed differences and, as an additional quality check, confirmed
246 that all F_1 females were heterozygous at these loci. We then used STACKS to call SNPs
247 in the F_2 haploid males, requiring that each SNP had a minimum of 5x coverage, no more
248 than 10% missing data, and was present in the curated list produced from the
249 grandparents. Filtering in STACKs produced 559 SNPs genotyped in 429 F_2 males.
250 Second, to maximize the number of SNPs available for genome scaffolding, we
251 performed an additional STACKS run using only the F_2 haploid males, requiring that
252 each SNP had a minimum of 4x coverage. By removing the requirement that SNPs were
253 called in all grandparents, we could recover many more SNPs. We then filtered the data
254 in VCFtools v0.1.14 (Danecek *et al.* 2011) to remove individuals with a mean depth of
255 coverage less than one, retaining 408 F_2 males. After removing low-coverage individuals,
256 we used VCFtools to remove sites with a minor allele frequency (MAF) less than 0.05,
257 sites with >5 heterozygotes (in a sample of haploid males, high heterozygosity is a clear
258 indication of pervasive genotyping error), and sites with more than 50% missing data.
259
260 **Linkage map construction and genome scaffolding**
261 To construct a linkage map for interval mapping, we started with 559 SNPs
262 scored in 429 F_2 males. After an additional round of filtering in R/qtl (Broman and Sen
263 2009), we removed 11 haploid males that had >50% missing data, for a total of 418 F_2
264 males. Additionally, after removing SNPs that were genotyped in <70% of individuals,
265 had identical genotypes to other SNPs, and had distorted segregation ratios (at $\alpha<0.05$,
266 after Bonferroni correction for multiple testing), we retained 503 SNPs. To assign these
267 markers to linkage groups, we then used the "formLinkageGroups" function, requiring a
268 minimum logarithm of odds (LOD) score of 6.0 and a maximum recombination
269 frequency of 0.35. To order markers on linkage groups, we used the "orderMarkers"
270 function, with the Kosambi mapping function to allow for crossovers. Following this
271 initial ordering, we performed rippling on each linkage group to check whether switching
272 marker order could improve LOD scores.
273 Our initial map included 503 SNPs spread across 358 scaffolds (out of 4523
274 scaffolds; Vertacnik *et al.* 2016). To increase the number of scaffolds and bases that we
275 could place on our linkage groups, we performed additional linkage mapping analyses
276 with a larger SNP dataset that was called as described above. We then constructed a
277 linkage map for each of our four grandparental families (N = 54, 73, 120, and 161). For
278 each grandparental family, we first performed additional data filtering in R/qtl to remove
279 duplicate SNPs and SNPs with distorted segregation ratios, retaining between 2,049 and
280 3,155 SNPs per family. We then used the "formLinkageGroups" command with variable
281 LOD thresholds (range: 5-15) and a maximum recombination frequency of 0.35. Because
282 SNPs were not coded according to grandparent of origin, many alleles were "switched".
283 We therefore performed an iterative process of linkage group formation, visualization of
284 pairwise recombination fractions and LOD scores ("plotRF" command), and allele
285 switching ("switchAlleles" command) until we obtained seven linkage groups (the
286 number of *N. lecontei* chromosomes; Smith 1942; Maxwell 1958; Sohi and Ennis 1981)
287 and a recombination/LOD plot indicative of linkage within, but not between, linkage
288 groups. Because allele ordering and examination of alternative SNP orders for these

289 larger panels of SNPs was prohibitively slow in R/qtl, we used the more efficient
290 MSTmap algorithm, implemented in R/ASMap v0.4-7 (Taylor and Butler 2017), to order
291 our markers along their assigned linkage groups.
292 Finally, to order and orient our genome scaffolds along linkage groups
293 (chromosomes), we used ALLMAPS (Tang *et al.* 2015) to combine information from our
294 five maps (initial map with all individuals, but limited markers; plus four additional
295 maps, each with more markers, but fewer individuals). Because maps constructed from
296 larger families are likely to be more accurate than those constructed from small families,
297 we weighted the maps according to their sample sizes.
298
299 **Interval mapping analysis**
300 After linkage map construction, we used R/qtl to map QTL for our five color
301 traits. Based on analyses of phenotypic covariates (Table S2), we included F_1 mothers
302 and head-capsule size as covariates in our analyses of PC1, spot number, and spot area;
303 F_1 mothers as covariates in our analysis of PC2; and no covariates in in our analysis of
304 yellow. For each trait, we performed interval mapping using multiple imputation
305 mapping. We first used the "sim.geno" function with a step size of 0 (*i.e.*, genotypes only
306 drawn at marker locations) and 64 replicates. We then used the "stepwiseqtl" command
307 to detect QTL and select the multiple QTL model that optimized the penalized LOD
308 score (Broman and Sen 2009; Manichaikul *et al.* 2009). To obtain penalties for the
309 penalized LOD scores, we used the "scantwo" function to perform 1,000 permutations
310 under a two-dimensional, two-QTL model that allows for interactions between QTL and
311 the "calc.penalties" function to calculate penalties from these permutation results, using a
312 significance threshold of $\alpha = 0.05$. Finally, for each QTL retained in the final model, we
313 calculated a 1.5-LOD support interval.
314
315 **Candidate gene analysis**
316 As a first step to moving from QTL intervals to causal loci, we compiled a list of
317 candidate color genes and determined their location in the *N. lecontei* genome. For larval
318 spotting, we included genes in the melanin synthesis pathway and genes that have been
319 implicated in pigmentation patterning (Wittkopp *et al.* 2003; Protas and Patel 2008;
320 Wittkopp and Beldade 2009; Sugumaran and Barek 2016). For larval body color, we
321 included genes implicated in the transport, deposition, and processing of carotenoid
322 pigments derived from the diet (Palm *et al.* 2012; Yokoyama *et al.* 2013; Tsuchida and
323 Sakudoh 2015; Toews *et al.* 2017). Although several pigments can produce yellow
324 coloration in insects (*e.g.,* melanins, pterins, ommochromes, and carotenoids), we
325 focused on carotenoids because a heated pyridine test (McGraw *et al.* 2005) was
326 consistent with carotenoid-based coloration in *N. lecontei* larvae (Figure S1).
327 Once we had compiled a list of candidate genes, we searched for these genes by
328 name in the *N. lecontei* v1.0 genome assembly and NCBI annotation release 100
329 (Vertacnik *et al.* 2016). To find missing genes and as an additional quality measure, we
330 obtained FASTA files corresponding to each candidate protein and/or gene from NCBI
331 (using *Apis, Drosophila melanogaster,* or *Bombyx mori* sequences, depending on
332 availability). We then used the i5k Workspace@NAL (Poelchau *et al.* 2014) BLAST
333 (Altschul *et al.* 1990) web application to conduct tblastn (for protein sequences) or tblastx
334 (for gene sequences) searches against the *N. lecontei* v1.0 genome assembly, using

335 default search settings. After identifying the top hit for each candidate gene/protein, we
336 then used the WebApollo (Lee *et al.* 2013) JBrowse (Skinner *et al.* 2009) *N. lecontei*
337 genome browser to identify the corresponding predicted protein coding genes (from
338 NCBI annotation release 100) in the *N. lecontei* genome.
339 We took additional steps to identify genes in the *yellow* gene family, all of which
340 contain a major royal jelly protein (MRJP) domain (Ferguson *et al.* 2011). First, we used
341 the search string "major royal jelly protein *Neodiprion*" to search the NCBI database for
342 all predicted *yellow*-like and *yellow-MRJP*-like *N. lecontei* genes. We then downloaded
343 FASTA files for the putative *yellow* gene sequences (26 total). Next, we used the
344 Hymenoptera Genome Database (Elsik *et al.* 2016) to conduct a blastx search of our *N.*
345 *lecontei* gene sequence queries against the *Apis mellifera* v4.5 genome NCBI RefSeq
346 annotation release 103. Finally, we recorded the top *A. mellifera* hit for each putative *N.*
347 *lecontei yellow* gene.
348
349 **Data availability**
350 Short-read DNA sequences will be made available via the NCBI SRA (Bioproject
351 PRJNA#######, Biosample numbers SAMN########-SAMN########). The linkage-
352 group anchored assembly will be submitted to NCBI and i5k to update the existing *N.*
353 *lecontei* genome assembly and annotations (Vertacnik *et al.* 2016). File S3 contains the
354 custom script used to process the raw spot-area data. File S4 contains phenotypic data
355 from all generations. File S5 contains the input file for R/qtl. Files S6 and S7 contain the
356 genotype data and linkage maps used in the scaffolding analysis, respectively.
357
358 RESULTS AND DISCUSSION
359
360 **Variation in larval color traits**
361 Lab-reared larvae derived from the two founding populations differed
362 significantly from one another for all color phenotypes (Figures 2, 3; Table S3). Because
363 all larvae were reared on the same host under the same laboratory conditions (*i.e.*,
364 minimal environmental variance), these results suggest that genetic variance contributes
365 to variance in these larval color traits. Crosses between the VA and MI lines produced
366 diploid F_1 female larvae that were intermediate in both body color and spotting pattern
367 (Figures 2, 3). With the exception of PC2, F_1 larvae differed significantly from both
368 grandparents for all color traits (Table S3), indicating a lack of complete dominance for
369 these traits. In contrast, the yellow, MI-like hue/chroma (PC2) appears to be completely
370 dominant to the white, VA-like hue/chroma. We can make additional inferences about
371 trait dominance by comparing the phenotypes of diploid F_1 female larvae (dominance
372 effects present) to those of the haploid F_2 male larvae (dominance effects absent). For all
373 three body-color measures (yellow, PC1, and PC2), average F_1 larval colors were more
374 MI-like (yellow) than those of F_2 larvae. In contrast, compared to the F_2 larvae, spotting
375 phenotypes (number and area) of F_1 larvae are more similar to the VA population (dark-
376 spotted). Taken together, these phenotypic data suggest that the more heavily pigmented
377 body-color and spotting-pattern phenotypes (*i.e.,* large values for yellow, spot number,
378 and spot area; small values for PC1 and PC2) are partially dominant to the less pigmented
379 phenotypes.

380 Larval color variation observed in the F_2 males spanned—and even exceeded—
381 the range of variation observed in the grandparental populations and F_1 females (Figure
382 3). The observation that grandparental body-color and spotting-pattern phenotypes are
383 recapitulated in the F_2 males suggests that both types of traits are controlled by a
384 relatively small number of loci. There are multiple, non-mutually exclusive explanations
385 for the transgressive color phenotypes in our haploid F_2 larvae, including: variation in the
386 grandparental lines, reduced developmental stability in hybrids, epistasis, unmasking of
387 recessive alleles in haploid males, and the complementary action of additive alleles from
388 the two grandparental lines (Rieseberg *et al.* 1999).
389 We also observed significant correlations between many of the color traits (Figure
390 S2). Encouragingly, estimates of body color derived from digital images (yellow)
391 correlated significantly with both estimates derived from spectrometry (PC1 and PC2),
392 suggesting that these independent data sources describe the same underlying larval color
393 phenotype (*i.e.,* amount of yellow pigment). Similarly, we observed strong correlations
394 between the number and area of larval spots, indicating that both measures reliably
395 characterize the extent of melanic spotting on the larval body. We also found that larvae
396 with yellower bodies (small values for PC2) tended to be less heavily spotted (Figure S2).
397 This observation suggests that body color and spotting pattern do not evolve completely
398 independently of one another. Nevertheless, the correlation between these traits was weak
399 and we observed many different combinations of spotting and pigmentation in the
400 recombinant F_2 males (Figure 2).
401
402 **Linkage mapping and genome scaffolding**
403 Our 503 SNP markers were spread across seven linkage groups (LGs), which
404 matches the number of *N. lecontei* chromosomes (Smith 1941; Maxwell 1958; Sohi and
405 Ennis 1981). The total map length was 1169 cM, with an average marker spacing of 2.4
406 cM and maximum marker spacing of 24.3 cM (Table S4; Figure 4). Together, these
407 results indicate that this linkage map is of sufficient quality and coverage for interval
408 mapping. Additionally, with an estimated genome size of 340 Mb (estimated via flow
409 cytometry; C. Linnen, personal observation), these mapping results yield a recombination
410 density estimate of 3.43 cM/Mb. This recombination rate is lower than that observed in
411 social hymenopterans, which have among the highest rates of recombination in
412 eukaryotes (Wilfert *et al.* 2007). Nevertheless, this rate is on par with that reported in
413 other (non-eusocial) hymenopterans, which lends support to the hypothesis that elevated
414 recombination rates in eusocial hymenopteran species is a derived trait and possibly an
415 adaptation to a social lifestyle (Gadau *et al.* 2000; Schmid-Hempel 2000; Crozier and
416 Fjerdingstad 2001).
417 Linkage maps estimated for the four grandparental families, each of which
418 contained >2000 markers, ranged in length from 1072 cM to 3064 cM (Table S4). This
419 variation in map length is likely attributable to both decreased mapping accuracy in
420 smaller families and decreased genotyping accuracy in these less-stringently filtered SNP
421 datasets. Nevertheless, our scaffolding analysis revealed that marker ordering was highly
422 consistent across linkage maps (Figures S3-S9). Additionally, via including more SNPs,
423 we were able to more than triple the number of mapped scaffolds (from 358 to 1005) and
424 increase the percentage of mapped bases from 41.2% to 78.9% (Tables S5-S6).

425 Anchored genome scaffolds, coupled with existing *N. lecontei* gene annotations, are a
426 valuable resource for identification of candidate genes within QTL.
427
428 **Genetic architectures of larval body color and larval spotting pattern**
429 Combining the results from our three measures of larval body-color (yellow, PC1,
430 and PC2), our interval mapping analyses detected a total of six distinct QTL regions and
431 one QTL x QTL interaction (Table 1; Figures 4,5A). Of these, the largest-effect QTL
432 were on LGs 3 and 5. We also recovered a significant interaction between these QTL for
433 yellow (Figure 5B). Compared to analyses of larval body color, analyses of larval
434 spotting pattern yielded fewer distinct QTL (four), but more QTL x QTL interactions
435 (three) (Table 1; Figures 4, 5C, 5D). Both spotting measures (spot number and spot area)
436 recovered two large-effect QTL on LG 2. These spotting QTL also overlapped with QTL
437 with small-to-modest effects on body color (Table 1; Figure 4). Co-localization of
438 spotting-pattern and body-color QTL suggests that the phenotypic correlations we
439 observed between body color and spotting pattern in F_2 males (Figure S2) are caused by
440 underlying genetic correlations via either pleiotropy or physical linkage.
441 These genetic mapping results provide us with our first glimpse into the genetic
442 architecture of naturally occurring color variation in pine sawflies. For both larval body
443 color and larval spotting pattern, we recovered a relatively small number of loci, some of
444 which had a substantial effect on the color phenotype (Table 1). Additionally, our
445 multiple-QTL models explained a considerable percentage of the observed variation in
446 larval color: up to 86% for larval body color (yellow) and up to 73% for larval spotting
447 pattern (spot number) (Table 2). Although limited statistical power precluded us from
448 quantifying the number of QTL of very small effect, our results suggest that the
449 proportion of variation explained by small, isolated QTL is small. It is possible, however,
450 that our large-effect QTL comprise many linked QTL of individually small effect (*e.g.,*
451 Stam and Laurie 1996; McGregor *et al.* 2007; Bickel *et al.* 2011; Linnen *et al.* 2013).
452 Distinguishing between oligogenic architectures (a small number of moderate to large-
453 effect loci) and polygenic architectures (a large number of small-effect loci) will require
454 fine-mapping and functionally validating genes and mutations for both color traits.
455 Nevertheless, for both types of color traits, our results clearly rule out both a monogenic
456 architecture and an architecture in which there are many unlinked, small-effect mutations.
457 The latter architecture is also highly unlikely if local color adaptation in pine sawflies
458 proceeded in the face of gene flow (Griswold 2006; Yeaman and Whitlock 2011), as
459 previous demographic analyses seem to indicate is the case (Bagley *et al.* 2017).
460 To date, QTL-mapping and candidate-gene studies of color traits have yielded
461 many examples of relatively simple genetic architectures in which one or a small number
462 of mutations have very large phenotypic effects (Rockman 2012; Martin and Orgogozo
463 2013). However, the relevance of these studies to broader patterns in phenotypic
464 evolution remains unclear due to biases that stem from a tendency to work on discrete
465 phenotypes and specific candidate genes. These biases can be minimized by focusing on
466 continuously varying color traits and employing an unbiased mapping approach as we
467 have done here (see also O'Quin *et al.* 2013; Albertson *et al.* 2014; Signor *et al.* 2016;
468 Yassin *et al.* 2016). Experimental biases aside, some have also argued that color itself—
469 but not color pattern—is atypically genetically simple (i.e., more likely to be mono- or
470 oligogenic than other traits) (Charlesworth *et al.* 1982; Rockman 2012). In our own data,

471 we see no obvious differences between the architectures of body color and spotting
472 pattern. If anything, the architecture of spotting appears slightly "simpler" (fewer QTL
473 with larger individual effects) than that of body color (Table 1). That said, more spotting
474 variation remains unexplained (Table 2), which could be attributable to a plethora of
475 undetected small-effect loci, and thus a more complex underlying architecture. A more
476 definitive test of the "color/pattern" hypothesis will therefore require fine-mapping our
477 QTL, with the prediction that the spotting QTL will "fractionate" to a greater extent (i.e.,
478 break into more QTL of individually smaller effect) than the body-color QTL. A rigorous
479 evaluation of the color/pattern hypothesis will also require data on the genetic
480 architecture of color and pattern traits that have evolved independently in other taxa. To
481 this end, ample variation in both body-color and color-pattern traits make *Neodiprion*
482 sawflies an ideal system for determining whether there are predictable differences in the
483 genetic architecture of different types of color traits.
484
485 **Candidate genes for larval color traits**
486 In total, we identified 61 candidate genes with known or suspected roles in
487 melanin-based or carotenoid-based pigmentation and for which we were able to identify
488 putative homologs in the *N. lecontei* genome (Table S7). Of these, 26 appeared to belong
489 to the *yellow* gene family. Notably, this number is equivalent to the number of *yellow*-
490 like/*yellow-MRJP*-like genes found in the genome of the jewel wasp, *Nasonia vitripennis,*
491 which boasts the highest reported number of *yellow*-like genes of any insect to date
492 (Werren *et al.* 2010). Additionally, 13 of these genes (*yellow-e, yellow-e3*, four *yellow-g*,
493 *yellow-h, yellow-x,* and five *MJRPs*) were found in tandem array along three adjacent
494 scaffolds (548, 170, and 36; ~1 Mb total) on LG 2. This genomic organization is
495 consistent with a conserved clustering of *yellow-h, -e3, -e, -g2,* and *–g* observed across
496 *Apis, Tribolium, Bombyx, Drosophila*, and *Nasonia* (Drapeau *et al.* 2006; Werren *et al.*
497 2010; Ferguson *et al.* 2011). Like *Nasonia* and *Apis*, this cluster also contains *MJRPs*;
498 like *Heliconius*, this cluster contains a *yellow-x* gene. Overall, we placed 57 of our 61
499 color genes on our LG-anchored assembly. With these data, we identified candidate
500 genes for all but two QTL (Yellow-1 on LG 1 and SpotNum-3 on LG 6) (Table S7).
501
502 *Candidate genes for larval body color*
503 Both of our largest-effect body-color QTL regions contained promising candidate
504 genes with known or suspected roles in carotenoid-based pigmentation (Table S7). First,
505 within the LG-3 QTL region (Yellow-4, PC1-1, PC2-3; Table 1, Figure 4), we found a
506 predicted protein-coding region in scaffold 518 with a high degree of similarity to the
507 *Bombyx mori* Cameo2 scavenger receptor protein (e-value: 1×10^{-18}; bitscore: 92.8).
508 *Cameo2* encodes a transmembrane protein belonging to the CD36 family that has been
509 implicated in the selective transport of the carotenoid lutein from the hemolymph to
510 specific tissues (Sakudoh *et al.* 2010; Tsuchida and Sakudoh 2015). In the silkworm
511 *Bombyx mori*, *Cameo2* is responsible for the "C mutant" phenotype, which is
512 characterized by a combination of yellow hemolymph and white cocoons that arises as a
513 consequence of disrupted transport of lutein from the hemolymph to the middle silk gland
514 (Sakudoh *et al.* 2010; Tsuchida and Sakudoh 2015).
515 Second, within the LG-5 QTL region (Yellow-5, Yellow-6, PC1-2, PC2-4; Table
516 1, Figure 4), we recovered a predicted protein coding gene in scaffold 164 with a high

517 degree of similarity to the ApoLTP-1 and ApoLTP-2 protein subunits (encoded by the
518 gene *apoLTP-II/I*) of the *Bombyx mori* lipid transfer particle (LTP) lipoprotein (e-value:
519 0; bitscore: 391). LTP is one of two major lipoproteins present in insect hemolymph and
520 appears to be involved in the transport of hydrophobic lipids (including carotenoids) from
521 the gut to the other major lipoprotein, lipophorin, which then transports lipids to target
522 tissues (Tsuchida *et al.* 1998; Palm *et al.* 2012; Yokoyama *et al.* 2013). Based on these
523 observations, we hypothesize that loss-of-function mutations in both *Cameo2* and
524 *apoLTP-II/I* contribute to the loss of yellow pigmentation in white-bodied *N. lecontei*
525 larvae.
526
527 *Candidate genes for larval spotting pattern*
528 Our two largest-effect spotting-pattern QTL also yielded promising candidate
529 genes—this time in the well-characterized melanin biosynthesis pathway (Table S7). In
530 both LG-2 QTL regions, we found protein-coding genes that belong to the *yellow* gene
531 family. The first LG-2 QTL region (SpotNum-1 and SpotArea-1; Table 1, Figure 4)
532 contained 2 *yellow*-like genes that were most similar to *Apis yellow-x1* (e-value: $1.2 \times 10^{-}$
533 ^{160}; bitscore: 471.47 and e-value: 3.9×10^{-151}; bitscore: 448.36). At present, there is little
534 known about the function of *yellow-x* genes, which appear to be highly divergent from
535 other *yellow* gene families (Ferguson *et al.* 2011). The second LG-2 QTL region
536 (SpotNum-2 and SpotArea-2; Table 1, Figure 4) contained a cluster of 13 *yellow* genes,
537 including *yellow-e.* In two different mutant strains of *B. mori*, mutations in *yellow-e*
538 produced a truncated gene product that results in increased reddish-brown pigmentation
539 in the head cuticle and anal plate compared to wildtype strains (Ito *et al.* 2010). In
540 wildtype larvae, *yellow-e* is most highly expressed in the integument of the head and the
541 tail (Ito *et al.* 2010). Based on these observations, one possible mechanism for the
542 reduced spotting observed in the light-spotted MI population is an increase in *yellow-e*
543 expression.
544 The second LG-2 spotting QTL region also contained a predicted protein that was
545 highly similar to tyrosine hydroxylase (TH) (e-score: 6×10^{-123}, bitscore: 406). TH
546 catalyzes the hydroxylation of tyrosine to 3,4-dihydroxyphenylalanine (DOPA), a
547 precursor to melanin-based pigments (Wright 1987). Work in the swallowtail butterfly
548 *Papilio xuthus* and the armyworm *Pseudaletia separata* demonstrates that TH and
549 another enzyme, dopa decarboxylase (DDC), are expressed in larval epithelial cells
550 containing black pigment (Futahashi and Fujiwara 2005; Ninomiya and Hayakawa 2007).
551 Furthermore, inhibition of either enzyme prevented the formation of melanin-based larval
552 pigmentation patterns (Futahashi and Fujiwara 2005). Thus, a reduction in the regional
553 expression of *TH* (aka *pale*) is another plausible mechanism underlying reduced spotting
554 in the light-spotted MI population.
555 We also found candidate genes in some of our QTL of more modest effect. As
556 noted above, the Yellow-2 and Yellow-3 QTL overlap with the two large-effect spotting
557 QTL (Table 1, Figures 4,5). One possible explanation for this observation is that in
558 addition to impacting spotting phenotypes, genes in the melanin biosynthesis pathway
559 also impact overall levels of melanin throughout the integument and therefore overall
560 body color. Finally, the SpotNum-4 QTL contained a gene encoding dopamine N-acetyl
561 transferase (*Dat*). Dopamine N-acetyl transferase, which catalyzes the reaction between
562 dopamine and N-acetyl dopamine and ultimately leads to the production of a colorless

563 pigment, is responsible for a difference in pupal case color between two *Drosophila*
564 species (Ahmed-Braimah and Sweigart 2015).
565
566 *Comparison with color loci in other taxa*
567 Identification of candidate genes within our QTL peaks provides an opportunity
568 to compare our results with those from different insect taxa. For melanin-based traits, the
569 bulk of existing genetic studies involve *Drosophila* fruit flies (Massey and Wittkopp
570 2016), with a handful of additional studies in wild and domesticated lepidopterans (*e.g.,*
571 Ito *et al.* 2010; Hof *et al.* 2016; Nadeau *et al.* 2016). A recent synthesis of *Drosophila*
572 studies suggests that *cis*-regulatory changes in a handful of genes controlling melanin
573 biosynthesis (*yellow*, *ebony*, *tan, Dat*) and patterning (*bab1*, *bab2*, *omb*, *Dll*, and *wg*)
574 explain much of the intra- and interspecific variation in pigmentation (Massey and
575 Wittkopp 2016). Although our data cannot speak to the contribution of *cis*-regulatory
576 versus protein-coding changes, we do recover some of the same melanin biosynthesis
577 genes (*yellow* and *Dat*). By contrast, all of the *Drosophila* melanin patterning genes fell
578 outside of our spotting-pattern QTL intervals (Table S7). While we cannot rule these
579 genes out completely, our results suggest that none of them plays a major role in this
580 intraspecific comparison. Whether the same is true of patterning differences within and
581 between other *Neodiprion* species remains to be seen.
582 Compared to melanin-based pigmentation, much less is known about the genetic
583 underpinnings of variation in carotenoid-based pigmentation, which is widespread in
584 nature (Heath *et al.* 2013; Toews *et al.* 2017). Nevertheless, studies in the domesticated
585 silkworm (Sakudoh *et al.* 2010; Tsuchida and Sakudoh 2015), salmonid fish (Sundvold *et*
586 *al.* 2011), and scallops (Liu *et al.* 2015) collectively suggest that scavenger receptor
587 genes such as our candidate *Cameo2* may be a common and taxonomically widespread
588 source carotenoid-based color variation (Toews *et al.* 2017). Beyond evaluating patterns
589 of gene reuse, extensive intra- and interspecific variation in larval body color across the
590 genus *Neodiprion* (Figure 1) has the potential to provide novel insights into the molecular
591 mechanisms underlying carotenoid-based pigmentation.
592
593 **Summary and Conclusions**
594 Our study, which focuses on naturally occurring color variation in an under-
595 sampled life stage (larva), taxon (Hymenoptera), and pigment type (carotenoids),
596 represents a valuable addition to the invertebrate pigmentation literature. Our results also
597 have several implications for both hymenopteran evolution and color evolution. First, we
598 provide the first recombination-rate estimate for the Eusymphyta, the sister group to all
599 remaining Hymenoptera (*e.g.,* bees, wasps, and ants) (Peters *et al.* 2017). Our estimate
600 suggests that the exceptionally high recombination rates observed in some social
601 hymenopterans represents a derived state, supporting the hypothesis that increased
602 recombination is an adaptation for increasing genetic diversity in a colony setting (Gadau
603 *et al.* 2000; Schmid-Hempel 2000; Crozier and Fjerdingstad 2001; Wilfert *et al.* 2007).
604 Second, for both larval body color and spotting pattern, we can conclude that
605 intraspecific variation is due neither to a single Mendelian locus nor to a large number of
606 unlinked, small-effect loci. It remains to be seen, however, whether color, color pattern,
607 and other types of phenotypes differ predictably in their genetic architectures (Rockman
608 2012; Dittmar *et al.* 2016). Third, our QTL intervals contain several key players in

melanin-based and carotenoid-based pigmentation in other taxa, including *yellow, Dat*, and *Cameo2* (Massey and Wittkopp 2016; Toews *et al.* 2017). Although functional testing is needed to establish causal links, we can rule out a major role for several core melanin-patterning genes in this intraspecific comparison. Finally, although much work remains—including fine-mapping QTL, functional analysis of candidate genes, and genetic dissection of diverse color traits in additional *Neodiprion* populations and species—our rapid progression from phenotypic variation to strong candidate genes demonstrates the tremendous potential of this system for addressing fundamental questions about the genetic basis of color variation in natural populations.

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 constructive comments on earlier versions of this manuscript, we thank Danielle Herrig, Emily Bendall, Kim Vertacnik, Katie Peichel, and two anonymous reviewers. 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

Ahmed-Braimah Y. H., Sweigart A. L., 2015 A single gene causes an interspecific difference in pigmentation in *Drosophila*. Genetics 200: 331–342.

Albertson R. C., Powder K. E., Hu Y., Coyle K. P., Roberts R. B., *et al.*, 2014 Genetic basis of continuous variation in the levels and modular inheritance of pigmentation in cichlid fishes. Mol. Ecol. 23: 5135–5150.

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

Andersson S., Prager M., 2006 Quantifying Colors. In: Hill GE, McGraw KJ (Eds.), *Bird Coloration Volume 1: Mechanisms and Measurements*, Harvard University Press, Cambridge, MA, pp. 41–89.

Arendt J., Reznick D., 2008 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23: 26–32.

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.

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.

655 Bickel R. D., Kopp A., Nuzhdin S. V, 2011 Composite effects of polymorphisms near
656 multiple regulatory elements create a major-effect QTL. PLoS Genet. 7: e1001275.
657 Broman K. W., Sen S., 2009 *A Guide to QTL Mapping with R/qtl*. Springer, New York.
658 Cain A. J., Sheppard P. M., 1954 Natural selection in *Cepaea*. Genetics 39: 89–116.
659 Catchen J., Hohenlohe P. A., Bassham S., Amores A., Cresko W. A., 2013 Stacks: An
660 analysis tool set for population genomics. Mol. Ecol. 22: 3124–3140.
661 Charlesworth B., Lande R., Slatkin M., 1982 A Neo-Darwinian commentary on
662 macroevolution. Evolution 36: 474–498.
663 Chen H., Rangasamy M., Tan S. Y., Wang H., Siegfried B. D., 2010 Evaluation of five
664 methods for total DNA extraction from western corn rootworm beetles. PLoS One 5.
665 Christin P. A., Weinreich D. M., Besnard G., 2010 Causes and evolutionary significance
666 of genetic convergence. Trends Genet. 26: 400–405.
667 Codella S. G., Raffa K. F., 1993 Defense strategies of folivorous sawflies. In: Wagner M,
668 Raffa KF (Eds.), *Sawfly Life History Adaptations to Woody Plants,* Academic Press
669 Inc., San Diego, pp. 261–294.
670 Codella S. G., Raffa K. F., 1995 Host plant influence on chemical defense in conifer
671 sawflies (Hymenoptera: Diprionidae). Oecologia 104: 1–11.
672 Conte G. L., Arnegard M. E., Peichel C. L., Schluter D., 2012 The probability of genetic
673 parallelism and convergence in natural populations. Proc. R. Soc. B 279: 5039–47.
674 Coppel H. C., Benjamin D. M., 1965 Bionomics of Nearctic pine-feeding diprionids.
675 Annu. Rev. Entomol. 10: 69–96.
676 Crozier R. H., Fjerdingstad E. J., 2001 Polyandry in social Hymenoptera: disunity in
677 diversity? Ann. Zool. Fennici 38: 267–285.
678 Danecek P., Auton A., Abecasis G., Albers C. A., Banks E., *et al.*, 2011 The variant call
679 format and VCFtools. Bioinformatics 27: 2156–2158.
680 Davey J. W., Hohenlohe P. A., Etter P. D., Boone J. Q., Catchen J. M., *et al.*, 2011
681 Genome-wide genetic marker discovery and genotyping using next-generation
682 sequencing. Nat. Rev. Genet. 12: 499–510.
683 Dittmar E. L., Oakley C. G., Conner J. K., Gould B., Schemske D., 2016 Factors
684 influencing the effect size distribution of adaptive substitutions. Proc. R. Soc. B 283:
685 20153065.
686 Drapeau M. D., Albert S., Kucharski R., Prusko C., Maleszka R., 2006 Evolution of the
687 Yellow / Major Royal Jelly Protein family and the emergence of social behavior in
688 honey bees. Genome Res.: 1385–1394.
689 Eisner T., Johnesse J. S., Carrel J., Hendry L. B., Meinwald J., 1974 Defensive use by an
690 insect of a plant resin. Science. 184: 996–999.
691 Elmer K. R., Meyer A., 2011 Adaptation in the age of ecological genomics: insights from
692 parallelism and convergence. Trends Ecol. Evol. 26: 298–306.
693 Elsik C. G., Tayal A., Diesh C. M., Unni D. R., Emery M. L., *et al.*, 2016 Hymenoptera
694 Genome Database: Integrating genome annotations in HymenopteraMine. Nucleic
695 Acids Res. 44: D793–D800.
696 Endler J. A., 1990 On the measurement and classification of colour in studies of animal
697 colour patterns. Biol. J. Linn. Soc. 41: 315–352.
698 Ferguson L. C., Green J., Surridge A., Jiggins C. D., 2011 Evolution of the insect yellow
699 gene family. Mol. Biol. Evol. 28: 257–272.
700 Fisher R., Ford E., 1947 The spread of a gene in natural conditions. Heredity. 1: 143–174.

701 Futahashi R., Fujiwara H., 2005 Melanin-synthesis enzymes coregulate stage-specific
702 larval cuticular markings in the swallowtail butterfly, *Papilio xuthus*. Dev. Genes
703 Evol. 215: 519–529.
704 Gadau J., Page R. E., Werren J. H., Schmid-Hempel P., 2000 Genome organization and
705 social evolution in Hymenoptera. Naturwissenschaften 87: 87–9.
706 Gaj T., Gersbach C. A., Barbas C. F., 2013 ZFN, TALEN, and CRISPR/Cas-based
707 methods for genome engineering. Trends Biotechnol. 31: 397–405.
708 Gerould J. H., 1921 Blue-green caterpillars: The origin and ecology of a mutuation in
709 hemolymph color in *Colias (Eurymus) philodice*. J. Exp. Zool. Part A Ecol. Genet.
710 Physiol. 34: 384–415.
711 Glastad K. M., Arsenault S. V., Vertacnik K. L., Geib S. M., Kay S., *et al.*, 2017
712 Variation in DNA methylation is not consistently reflected by sociality in
713 Hymenoptera. Genome Biol. Evol. 9: 1687–1698.
714 Gompel N., Prud'homme B., 2009 The causes of repeated genetic evolution. Dev. Biol.
715 332: 36–47.
716 Goodwin S., Mcpherson J. D., Mccombie W. R., 2016 Coming of age : ten years of next-
717 generation sequencing technologies. Nat. Rev. Genet. 17: 333–351.
718 Griswold C. K., 2006 Gene flow's effect on the genetic architecture of a local adaptation
719 and its consequences for QTL analyses. Heredity. 96: 445–453.
720 Haldane J., 1948 The theory of a cline. J. Genet. 48: 277–284.
721 Harper K. E., Bagley R. K., Thompson K. L., Linnen C. R., 2016 Complementary sex
722 determination, inbreeding depression and inbreeding avoidance in a gregarious
723 sawfly. Heredity. 117: 326–335.
724 Heath J. J., Cipollini D. F., Stireman J. O., 2013 The role of carotenoids and their
725 derivatives in mediating interactions between insects and their environment.
726 Arthropod. Plant. Interact. 7: 1–20.
727 Heimpel G. E., Boer J. G. de, 2008 Sex determination in the Hymenoptera. Annu. Rev.
728 Entomol. 53: 209–30.
729 Hoekstra H. E., Coyne J. A., 2007 The locus of evolution: Evo devo and the genetics of
730 adaptation. Evolution. 61: 995–1016.
731 Hof A. E. van't, Campagne P., Rigden D. J., Yung C. J., Lingley J., *et al.*, 2016 The
732 industrial melanism mutation in British peppered moths is a transposable element.
733 Nature 534: 102–105.
734 Huang Y., Liu Z., Rong Y. S., 2016 Genome editing: from *Drosophila* to non-model
735 insects and beyond. J. Genet. Genomics 43: 263–272.
736 Ito K., Katsuma S., Yamamoto K., Kadono-Okuda K., Mita K., *et al.*, 2010 Yellow-e
737 determines the color pattern of larval head and tail spots of the silkworm B*ombyx*
738 *mori*. J. Biol. Chem. 285: 5624–5629.
739 Kettlewell H. B. D., 1955 Selection experiments on industrial melanism in the
740 Lepidoptera. Heredity. 9: 323–342.
741 Knerer G., Atwood C. E., 1972 Evolutionary trends in subsocial sawflies belonging to
742 *Neodiprion abietis* complex (Hymenoptera: Tenthredinoidea). Am. Zool. 12: 407–
743 418.
744 Knerer G., Atwood C. E., 1973 Diprionid sawflies: polymorphism and speciation.
745 Science. 179: 1090–1099.
746 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., 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 Apollo: a web-based genomic annotation editing platform. Genome Biol 14: R93.
Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.
Lindstedt C., Carita L., Mappes J., Johanna M., Päivinen J., *et al.*, 2006 Effects of group size and pine defence chemicals on Diprionid sawfly survival against ant predation. Oecologia 150: 519–26.
Lindstedt C., Lindstrom L., Mappes J., 2009 Thermoregulation constrains effective warning signal expression. 63: 469–478.
Lindstedt C., Huttunen H., Kakko M., Mappes J., 2011 Disengtangling the evolution of weak warning signals: High detection risk and low production costs of chemical defences in gregarious pine sawfly larvae. Evol. Ecol. 25: 1029–1046.
Lindstedt C., Schroderus E., Lindström L., Mappes T., Mappes J., 2016 Evolutionary constraints of warning signals: A genetic trade-off between the efficacy of larval and adult warning coloration can maintain variation in signal expression. Evolution. 70: 2562–2572.
Linnen C. R., Farrell B. D., 2007 Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. Evolution. 61: 1417–1438.
Linnen C. R., Farrell B. D., 2008a Phylogenetic analysis of nuclear and mitochondrial genes reveals evolutionary relationships and mitochondrial introgression in the sertifer species group of the genus *Neodiprion* (Hymenoptera: Diprionidae). Mol. Phylogenet. Evol. 48: 240–257.
Linnen C. R., Farrell B. D., 2008b Comparison of methods for species-tree inference in the sawfly genus *Neodiprion* (Hymenoptera: Diprionidae). Syst. Biol. 57: 876–90.
Linnen C. R., Poh Y.-P., Peterson B. K., Barrett R. D. H., Larson J. G., *et al.*, 2013 Adaptive evolution of multiple traits through multiple mutations at a single gene. Science. 339: 1312–6.
Liu H., Zheng H., Zhang H., Deng L., Liu W., *et al.*, 2015 A de novo transcriptome of the noble scallop, *Chlamys nobilis*, focusing on mining transcripts for carotenoid-based coloration. BMC Genomics 16: 1–13.
Mackay T. F. C., Stone E. A., Ayroles J. F., 2009 The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. 10: 565–577.
Manceau M., Domingues V. S., Linnen C. R., Rosenblum E. B., Hoekstra H. E., 2010 Convergence in pigmentation at multiple levels: mutations, genes and function. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365: 2439–2450.
Manichaikul A., Moon J. Y., Sen Ś., Yandell B. S., Broman K. W., 2009 A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. Genetics 181: 1077–1086.

793 Martin A., Orgogozo V., 2013 The loci of repeated evolution: a catalog of genetic
794 hotspots of phenotypic variation. Evolution. 67: 1235–50.
795 Martin A., Courtier-Orgogozo V., 2017 Morphological evolution repeatedly caused by
796 mutations in signaling ligand genes. In: *Diversity and Evolution of Butterfly Wing*
797 *Patterns*, Springer, Singapore, pp. 59–87.
798 Massey J. H., Wittkopp P. J., 2016 The genetic basis of pigmentation differences within
799 and between *Drosophila* species. In: *Current topics in developmental biology*,
800 Elsevier, pp. 27–61.
801 Maxwell D. E., 1958 Sawfly cytology with emphasis upon Diprionidae (Hymenoptera,
802 Symphyta). Proc 10th Int Congr Entomol. 2: 961–978.
803 McGraw K. J., Hudon J., Hill G. E., Parker R. S., 2005 A simple and inexpensive
804 chemical test for behavioral ecologists to determine the presence of carotenoid
805 pigments in animal tissues. Behav. Ecol. Sociobiol. 57: 391–397.
806 McGregor A. P., Orgogozo V., Delon I., Zanet J., Srinivasan D. G., *et al.*, 2007
807 Morphological evolution through multiple cis-regulatory mutations at a single gene.
808 Nature 448: 587–590.
809 Montgomerie R., 2006 Analyzing Colors. In: Hill GE, McGraw KJ (Eds.), *Bird*
810 *Coloration Volume 1: Mechanisms and Measurement,* Harvard University Press,
811 Cambridge, MA, pp. 90–147.
812 Montgomerie R., 2008 CLR, version 1.05.
813 Nadeau N. J., Jiggins C. D., 2010 A golden age for evolutionary genetics? Genomic
814 studies of adaptation in natural populations. Trends Genet. 26: 484–492.
815 Nadeau N. J., Pardo-Diaz C., Whibley A., Supple M. A., Saenko S. V., *et al.*, 2016 The
816 gene *cortex* controls mimicry and crypsis in butterflies and moths. Nature 534: 106–
817 110.
818 Ninomiya Y., Hayakawa Y., 2007 Insect cytokine, growth-blocking peptide, is a primary
819 regulator of melanin-synthesis enzymes in armyworm larval cuticle. FEBS J. 274:
820 1768–1777.
821 O'Quin C. T., Drilea A. C., Conte M. a, Kocher T. D., 2013 Mapping of pigmentation
822 QTL on an anchored genome assembly of the cichlid fish, *Metriaclima zebra*. BMC
823 Genomics 14: 287.
824 Orr H. A., Coyne J. A., 1992 The genetics of adaptation: a reassessment. Am. Nat. 140:
825 725–742.
826 Palm W., Sampaio J. L., Brankatschk M., Carvalho M., Mahmoud A., *et al.*, 2012
827 Lipoproteins in *Drosophila melanogaster*: assembly, function, and influence on
828 tissue lipid composition. PLoS Genet. 8.
829 Peters R. S., Krogmann L., Mayer C., Donath A., Gunkel S., *et al.*, 2017 Evolutionary
830 history of the Hymenoptera. Curr. Biol. 27: 1013–1018.
831 Peterson B. K., Weber J. N., Kay E. H., Fisher H. S., Hoekstra H. E., 2012 Double-digest
832 RADseq: an inexpensive method for de novo SNP discovery and genotyping in
833 model and non-model species. PLoS One 7: e37135.
834 Poelchau M., Childers C., Moore G., Tsavatapalli V., Evans J., *et al.*, 2014 The i5k
835 Workspace@NAL: enabling genomic data access, visualization and curation of
836 arthropod genomes. Nucleic Acids Res. 43: D714–D719.
837 Protas M. E., Patel N. H., 2008 Evolution of coloration patterns. Annu. Rev. Cell Dev.
838 Biol. 24: 425–446.

Remington D. L., 2015 Alleles versus mutations: Understanding the evolution of genetic architecture requires a molecular perspective on allelic origins. Evolution. 69: 3025–3038.

Rieseberg L. H., Archer M. a, Wayne R. K., 1999 Transgressive segregation, adaptation and speciation. Heredity. 83: 363–372.

Rockman M. V., 2012 The QTN program and the alleles that matter for evolution: All that's gold does not glitter. Evolution. 66: 1–17.

Rosenblum E. B., Parent C. E., Brandt E. E., 2014 The molecular basis of phenotypic convergence. Annu. Rev. Ecol. Evol. Syst. 45: 203–226.

Ruxton G. D., Sherratt T. N., Speed M. P., 2004 *Avoiding Attack: The Evolutionary Ecology of Crypsis, Warning Signals and Mimicry*. Oxford University Press, New York.

Sakudoh T., Iizuka T., Narukawa J., Sezutsu H., Kobayashi I., *et al.*, 2010 A CD36-related transmembrane protein is coordinated with an intracellular lipid-binding protein in selective carotenoid transport for cocoon coloration. J. Biol. Chem. 285: 7739–7751.

Schmid-Hempel P., 2000 Mating, parasites and other trials of life in social insects. Microbes Infect. 2: 515–520.

Signor S. A., Liu Y., Rebeiz M., Kopp A., 2016 Genetic convergence in the evolution of male-specific color patterns in *Drosophila*. Curr. Biol. 26: 2423–2433.

Skinner M. E., Uzilov A. V, Stein L. D., Mungall C. J., Holmes I. H., 2009 JBrowse: a next-generation genome browser. Genome Res 19: 1630–1638.

Smith S. G., 1941 A new form of spruce sawfly identified by means of its cytology and parthenogenesis. Sci. Agric. 21: 245–305.

Sohi S. S., Ennis T. J., 1981 Chromosomal characterization of cell lines of *Neodiprion lecontei* (Hymenoptera: Diprionidae). Proc. Entomol. Soc. Ontario 112: 45–48.

Stam L. F., Laurie C. C., 1996 Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. Evolution. 1564: 1559–1564.

Stern D. L., Orgogozo V., 2008 The loci of evolution: how predictable is genetic evolution? Evolution. 62: 2155–2177.

Stern D. L., Orgogozo V., 2009 Is genetic evolution predictable? Science. 323: 746–751.

Stevens M., Párraga C. A., Cuthill I. C., Partridge J. C., Troscianko T. S., 2007 Using digital photography to study animal coloration. Biol. J. Linn. Soc. 90: 211–237.

Streisfeld M. A., Rausher M. D., 2011 Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. Evolution. 65: 629–642.

Sugumaran M., Barek H., 2016 Critical analysis of the melanogenic pathway in insects and higher animals. Int. J. Mol. Sci. 17: 1–24.

Sumner F. B., 1926 An analysis of geographic variation in mice of the *Peromyscus polionotus* group from Florida and Alabama. J. Mammal. 7: 149–184.

Sundvold H., Helgeland H., Baranski M., Omholt S. W., Våge D. I., 2011 Characterisation of a novel paralog of scavenger receptor class B member I (SCARB1) in Atlantic salmon (*Salmo salar*). BMC Genet. 12.

Tang H., Zhang X., Miao C., Zhang J., Ming R., *et al.*, 2015 ALLMAPS: robust scaffold ordering based on multiple maps. Genome Biol. 16: 3.

Taylor J., Butler D., 2017 R Package ASMap: efficient genetic linkage map construction

885 and diagnosis. J. Stat. Softw. 79: 1–29.
886 R Core Team, 2013 R: A language and environment for statistical computing.
887 Toews D. P. L., Hofmeister N. R., Taylor S. A., 2017 The evolution and genetics of
888 carotenoid processing in animals. Trends Genet. 33: 171–182.
889 True J. R., 2003 Insect melanism: The molecules matter. Trends Ecol. Evol. 18: 640–647.
890 Tsuchida K., Arai M., Tanaka Y., Ishihara R., Ryan R. O., *et al.*, 1998 Lipid transfer
891 particle catalyzes transfer of carotenoids between lipophorins of *Bombyx mori*.
892 Insect Biochem. Mol. Biol. 28: 927–934.
893 Tsuchida K., Sakudoh T., 2015 Recent progress in molecular genetic studies on the
894 carotenoid transport system using cocoon-color mutants of the silkworm. Arch.
895 Biochem. Biophys. 572: 151–157.
896 Vertacnik K. L., Geib S., Linnen C. R., 2016 *Neodiprion lecontei* genome assembly
897 Nlec1.0. NCBI/GenBank. http://dx.doi.org/10.15482/USDA.ADC/1235565.
898 Werren J. H., Richards S., Desjardins C. A., Niehuis O., Gadau J., *et al.*, 2010 Functional
899 and evolutionary insights from the genomes of three parasitoid *Nasonia* species.
900 Science. 327: 343-348.
901 Wilfert L., Gadau J., Schmid-Hempel P., 2007 Variation in genomic recombination rates
902 among animal taxa and the case of social insects. Heredity. 98: 189–197.
903 Wilson L. F., Wilkinson R. C., Averill R. C., 1992 *Redheaded Pine Sawfly: Its Ecology*
904 *and Management*. USDA, Washington DC.
905 Wittkopp P. J., Carroll S. B., Kopp A., 2003 Evolution in black and white: Genetic
906 control of pigment patterns in *Drosophila*. Trends Genet. 19: 495–504.
907 Wittkopp P. J., Beldade P., 2009 Development and evolution of insect pigmentation:
908 Genetic mechanisms and the potential consequences of pleiotropy. Semin. Cell Dev.
909 Biol. 20: 65–71.
910 Wright T. R. F., 1987 The genetics of biogenic amine metabolism, sclerotization, and
911 melanization in *Drosophila melanogaster*. Adv. Genet. 24: 127–222.
912 Yassin A., Delaney E. K., Reddiex A. J., Seher T. D., Bastide H., *et al.*, 2016 The *pdm3*
913 locus is a hotspot for recurrent evolution of female-limited color dimorphism in
914 *Drosophila*. Curr. Biol. 26: 2412–2422.
915 Yeaman S., Whitlock M. C., 2011 The genetic architecture of adaptation under
916 migration-selection balance. Evolution. 65: 1897–1911.
917 Yokoyama H., Yokoyama T., Yuasa M., Fujimoto H., Sakudoh T., *et al.*, 2013 Lipid
918 transfer particle from the silkworm, *Bombyx mori*, is a novel member of the
919 apoB/large lipid transfer protein family. J. Lipid Res. 54: 2379–90.
920
921

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.

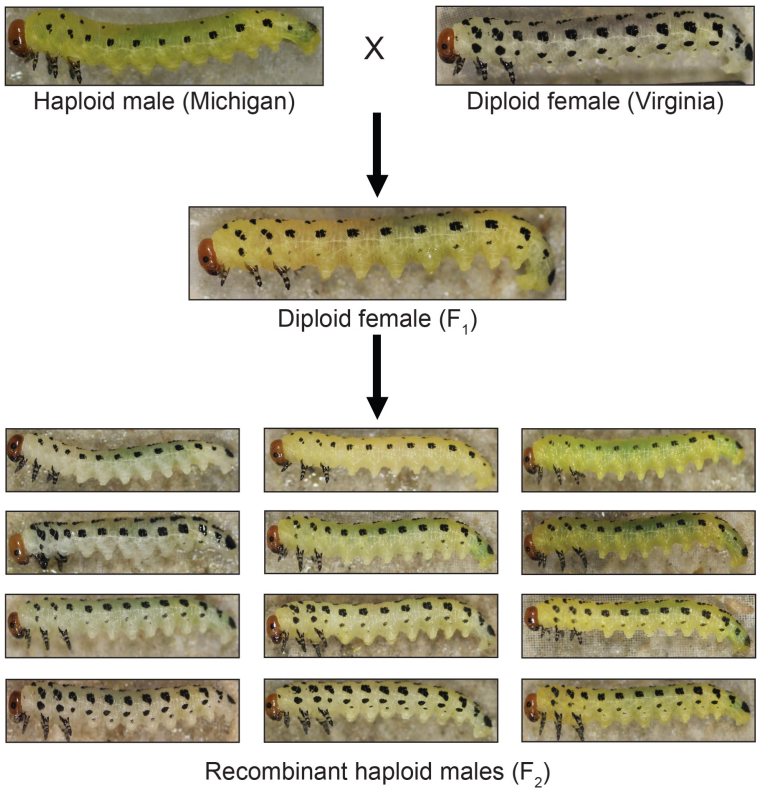

Figure 2. Intraspecific variation in *Neodiprion lecontei* larval color and cross design. We crossed white, dark-spotted diploid females from Virginia to yellow, light-spotted haploid males from Michigan. This produced haploid males with the VA genotype and phenotype (not shown) and diploid females (F_1) 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).

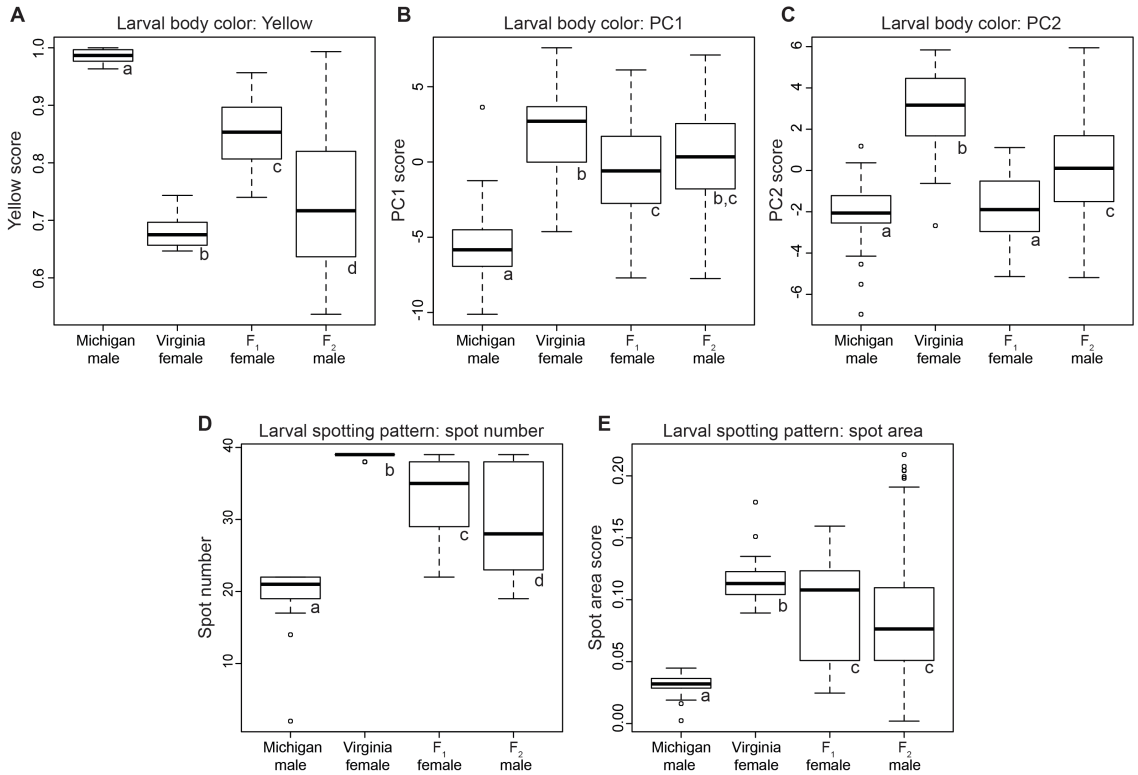

Figure 3. Larval color variation among *N. lecontei* populations and across generations. For larval body color (A-C), higher scores for yellow (A) and lower scores for PC1 (B) and PC2 (C) indicate higher levels of yellow pigment. For larval spotting pattern (D-E), higher spot numbers (D) and spot area scores (E) indicate more melanic spotting. For all traits, boxes represent interquartile ranges (median ± 2 s.d.), with outliers indicated as points, and letters denote significant differences after correction for multiple comparisons (see Table S3 for full statistical results).

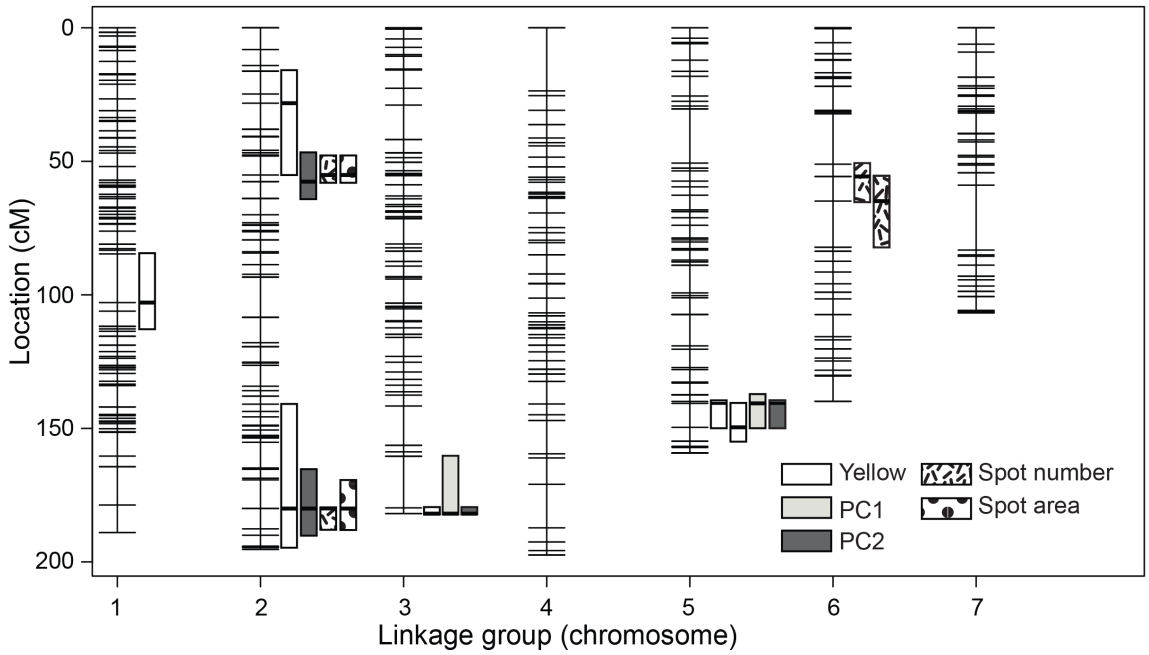

Figure 4. ***N. lecontei*** **linkage map with larval-color QTL and 1.5-LOD support intervals.**

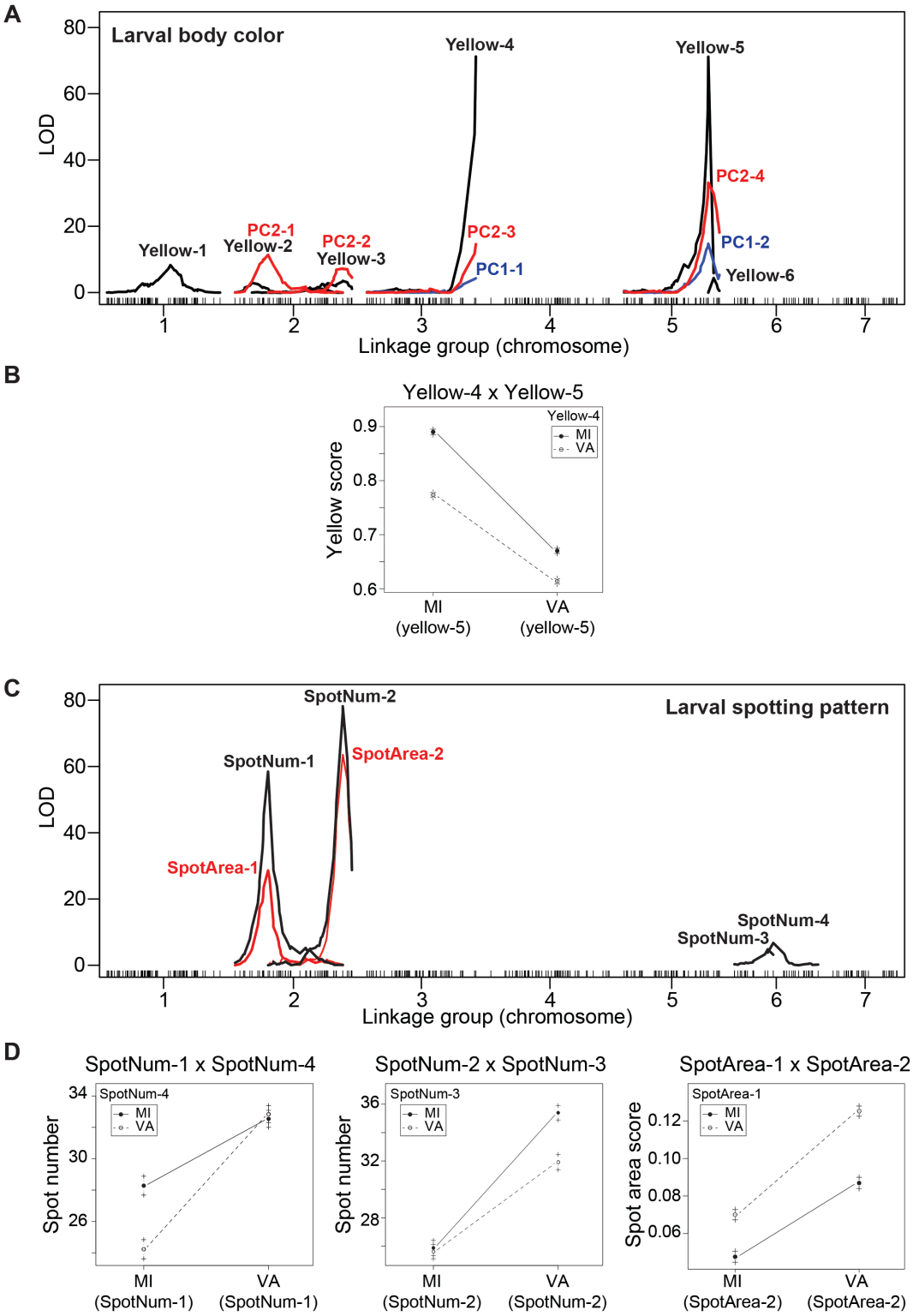

Figure 5. LOD profiles and interaction plots for larval color traits. Interval mapping analyses recovered QTL for larval body color (Yellow, PC1, PC2) on LGs 1, 2, 3, and 5 (A) and a single QTL x QTL interaction (B). QTL for larval spotting pattern (spot area and spot number) were on LGs 2 and 6 (C), with three QTL x QTL interactions (D). QTL names are as in Table 1.

Table 1. QTL locations and effect sizes for larval body color (yellow, PC1, and PC2) and larval spotting pattern (spot number and spot area).

Trait type	Trait	QTL name	LG[4]	Position (interval)[5]	Marker[6]	LOD	PVE	Effect size (SE)[7]	% diff.[8]
Body color	Yellow	Yellow-1	1	105.9 (84.4-112.5)	5744	8.29	1.32	-0.025 (0.0043)	8.06
Body color	Yellow	Yellow-2	2	28.2 (16.2-55.1)	19162	2.92	0.45	-0.014 (0.0043)	4.63
Body color	Yellow	Yellow-3	2	179.9 (140.8-194.0)	15279	3.51	0.54	-0.016 (0.0043)	5.13
Body color	Yellow	Yellow-4	3	181.9 (179.8-181.9)	1882	71.23	16.27	-0.084 (0.0043)	27.19
Body color	Yellow	Yellow-5	5	140.6 (139.9-149.6)	3222	71.17	16.25	-0.16 (0.0081)	51.64
Body color	Yellow	Yellow-6	5	149.6 (140.6-154.8)	19661	4.37	0.68	-0.035 (0.0082)	11.43
Body color	Yellow		Interaction	Yellow-3 x Yellow-5		12.45	2.03	0.061 (0.0086)	19.58
Body color	PC1	PC1-1	3	181.9 (160.4-181.9)	1882	4.31	3.56	1.15 (0.26)	15.51
Body color	PC1	PC1-2	5	140.6 (137.4-149.6)	3222	14.75	12.92	2.17 (0.26)	29.30
Body color	PC2	PC2-1	2	55.1 (46.8-63.9)	21	11.37	6.42	1.18 (0.16)	23.46
Body color	PC2	PC2-2	2	179.8 (165.2-189.9)	2143	7.24	3.99	0.91 (0.16)	17.99
Body color	PC2	PC2-3	3	181.9 (179.8-181.9)	1882	14.60	8.39	1.33 (0.16)	26.39
Body color	PC2	PC2-4	5	140.6 (139.9-149.6)	3222	33.13	21.13	2.09 (0.16)	41.45
Spotting pattern	Spot number	SpotNum-1	2	55.1 (48.0-57.6)	9282	58.46	23.52	6.31 (0.36)	33.06
Spotting pattern	Spot number	SpotNum-2	2	179.9 (179.8-187.5)	15279	78.19	35.45	7.81 (0.35)	40.96
Spotting pattern	Spot number	SpotNum-3	6	55.7 (51.1-64.9)	9040	4.84	1.43	-0.58 (0.59)	3.05
Spotting pattern	Spot number	SpotNum-4	6	64.9 (55.7-82.2)	7485	6.75	2.02	-1.00 (0.59)	5.26
Spotting pattern	Spot number		Interaction	SpotNum-1 x SpotNum-4		6.15	1.84	3.45 (0.71)	18.09
Spotting pattern	Spot number		Interaction	SpotNum-2 x SpotNum-3		4.64	1.38	-3.13 (0.71)	16.39
Spotting pattern	Spot area	SpotArea-1	2	55.1 (48.0-57.6)	9282	28.62	12.99	0.029 (0.0025)	34.86
Spotting pattern	Spot area	SpotArea-2	2	179.9 (169.1-187.5)	15279	63.73	35.47	0.048 (0.0024)	57.04
Spotting pattern	Spot area		Interaction	Spot-1 x Spot-2		1.76	0.69	0.014 (0.0049)	16.49

[4] Linkage group number
[5] Position in cM (1.5-LOD support intervals)
[6] Marker closest to QTL peak
[7] Effect sizes calculated 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).
[8] Effect sizes calculated as a percentage of the difference between average trait values for the two grandparental lines (VA and MI).

Table 2. Summary of full multiple-QTL models for five larval-color traits.

Trait type	Trait	Covariates	# QTL	# Interactions	LOD	PVE
Body color	Yellow	None	6	1	182.03	85.83
Body color	PC1	Size, Mother	2	0	27.14	25.48
Body color	PC2	Mother	4	0	65.84	50.92
Spotting pattern	Spot number	Size, Mother	4	2	122.18	73.23
Spotting pattern	Spot area	Size, Mother	2	1	94.77	63.93