Current Biology

Supplemental Information

Selection on a Genetic Polymorphism Counteracts Ecological Speciation in a Stick Insect

Aaron A. Comeault, Samuel M. Flaxman, Rüdiger Riesch, Emma Curran, Víctor Soria-Carrasco, Zachariah Gompert, Timothy E. Farkas, Moritz Muschick, Thomas L. Parchman, Tanja Schwander, Jon Slate, and Patrik Nosil

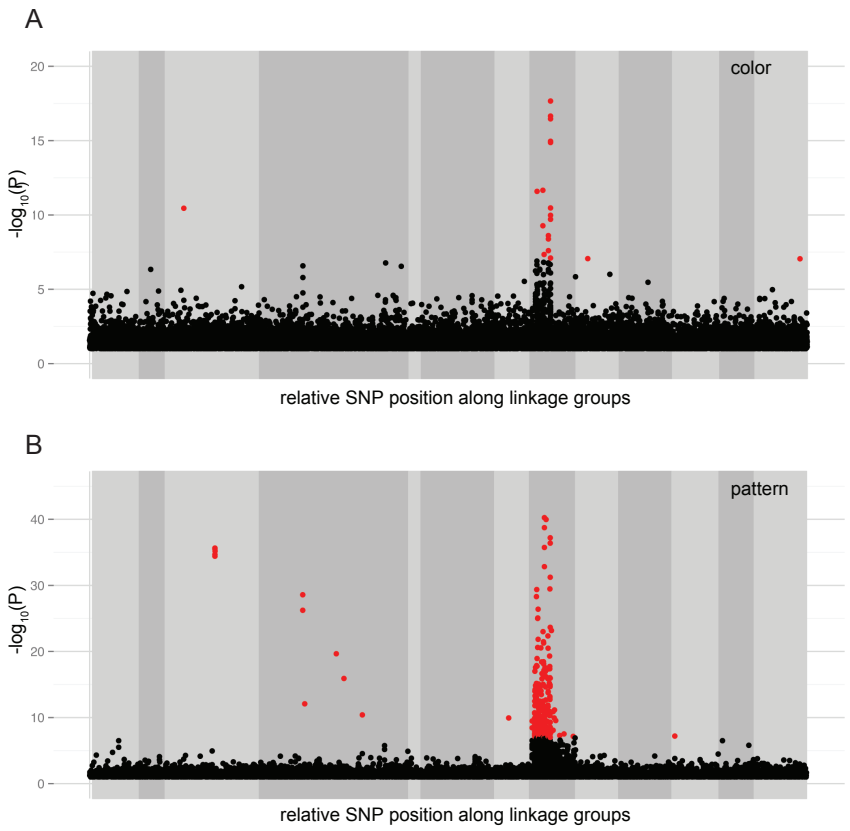

Figure S1 related to Figure 3. Manhattan plots for single-SNP GWA mapping carried out using the EIGENSTRAT method of Price et al. (2006). Results are partitioned by trait: A) color and B) dorsal pattern. Linkage groups are alternately shaded with grey rectangles and significantly associated SNPs ($P < 0.0000001$) are shown in red. See Supplemental Experimental Procedures below for details of analysis and Figure 3 of the main text for comparable results from multi-SNP GWA mapping.

Table S1 related to Figures 1 and 2. Details of populations sampled for different aspects of this study: (a) spectral color measurements, (b) frequency of color phenotypes in nature and climate effects on the proportion of melanistics.

Population code	Latitude (N)	Longitude (W)	Host plant	n-green	n-melanistic	Sample size
(a) Population samples of spectral measurements.						
FHA	34.5176	-119.8010	*A. fasciculatum*	108	5	113
HVA	34.4886	-119.7858	*A. fasciculatum*	22	6	28
RC	34.5151	-120.0710	*C. spinosus*	112	20	132
SC	34.5226	-119.8318	*C. spinosus*	10	0	10
VPC	34.5315	-119.8427	*C. spinosus*	34	2	36
(b) Population sampling for climate effects.						
BYA	34.5006	-119.8620	*A. fasciculatum*	27	1	28
ECC20A	34.5050	-119.7329	*A. fasciculatum*	32	1	33
ECC35A	34.5062	-119.7681	*A. fasciculatum*	22	2	24
ECCCampA	34.5064	-119.7616	*A. fasciculatum*	17	1	18
FHA	34.5176	-119.8010	*A. fasciculatum*	548	58	606
HVA	34.4886	-119.7858	*A. fasciculatum*	32	12	44
LA	34.5126	-119.7962	*A. fasciculatum*	20	9	29
MA	34.5151	-119.7971	*A. fasciculatum*	14	3	17
MH1978C	34.5191	-119.2710	*C. spinosus*	10	0	10
MH2559C	34.5332	-119.2431	*C. spinosus*	33	0	33
MH2919C	34.5554	-119.2632	*C. spinosus*	8	0	8
NS1A	34.4884	-119.6546	*A. fasciculatum*	19	4	23
OGA	34.5134	-119.7963	*A. fasciculatum*	15	4	19
OGC	34.5134	-119.7961	*C. spinosus*	14	2	16
OUTA	34.5317	-119.8435	*A. fasciculatum*	18	1	19
PC	34.4768	-119.7688	*C. spinosus*	81	25	106
PRC	34.5333	-119.8576	*C. spinosus*	29	4	33
R12C	34.5151	-120.0710	*C. spinosus*	35	2	37
R23A	34.5185	-120.0764	*A. fasciculatum*	29	0	29
SC	34.5226	-119.8318	*C. spinosus*	18	5	23

Table S2 related to Figure 2. Effect of color phenotype (i.e. "morph type") on the degree of crypsis when viewed by a UV sensitive avian predator in different host plant microhabitats. a) color distance and b) hue disparity. (b) and (t) indicate whether reflectance measurements were taken on the bottom or top of the leaf, respectively.

a) Color distance

Microhabitat	Term	*df*	SS	MS	*F*	*P*
A. fasciculatum - leaf	morph type	1, 309	0.7523	0.7523	2016.9	<0.0001
	sex	1, 309	0.0630	0.0630	168.8	<0.0001
	morph type * sex	7, 309	0.0059	0.0009	2.3	0.0284
C. spinosus – leaf(b)	morph type	1, 309	0.6815	0.6815	1736.0	<0.0001
	sex	1, 309	0.0465	0.0465	118.4	<0.0001
	morph type * sex	7, 309	0.0087	0.0013	3.2	0.0029
C. spinosus – leaf(t)	morph type	1, 309	0.7837	0.7837	1899.2	<0.0001
	sex	1, 309	0.0878	0.0878	212.8	<0.0001
	morph type * sex	7, 309	0.0068	0.0010	2.3	0.0243
A. fasciculatum - wood	morph type	1, 309	0.4537	0.4537	1097.2	<0.0001
	sex	1, 309	0.0812	0.0812	196.5	<0.0001
	morph type * sex	7, 309	0.0253	0.0036	8.7	<0.0001
C. spinosus - wood	morph type	1, 309	0.2173	0.2173	300.5	<0.0001
	sex	1, 309	0.0713	0.0713	98.6	<0.0001
	morph type * sex	7, 309	0.0631	0.0090	12.5	<0.0001

b) Hue disparity

Microhabitat	Term	*df*	SS	MS	*F*	*P*
A. fasciculatum - leaf	morph type	1, 309	1.5709	1.5709	2799.9	<0.0001
	sex	1, 309	0.0273	0.0273	48.6	<0.0001
	morph type * sex	7, 309	0.0378	0.0054	9.6	<0.0001
C. spinosus – leaf(b)	morph type	1, 309	1.4070	1.4070	3142.0	<0.0001
	sex	1, 309	0.0018	0.0018	4.0	0.0477
	morph type * sex	7, 309	0.0323	0.0046	10.3	<0.0001
C. spinosus –leaf(t)	morph type	1, 309	1.6782	1.6782	2416.9	<0.0001
	sex	1, 309	0.0740	0.0740	106.6	<0.0001
	morph type * sex	7, 309	0.0503	0.0072	10.4	<0.0001
A. fasciculatum - wood	morph type	1, 309	1.3937	1.3937	1967.6	<0.0001
	sex	1, 309	0.0796	0.0796	112.4	<0.0001
	morph type * sex	7, 309	0.0494	0.0071	10.0	<0.0001
C. spinosus - wood	morph type	1, 309	0.2173	0.2173	300.5	<0.0001
	sex	1, 309	0.0713	0.0713	98.6	<0.0001
	morph type * sex	7, 309	0.0631	0.0090	12.5	<0.0001

Table S3 related to Figure 2. Generalized linear models testing the effect of climate and host on the proportion of melanistic *T. cristinae* observed within a population. Each term was sequentially added to the "null" model (intercept only) and the residual deviance (res. deviance), AIC, and proportional increase in explained residual deviance (i.e. pseudo R^2) for that model was calculated. *P*-values are for the individual term added to a given model. Terms that significantly increase the explanatory power of the model are shown in bold.

Model	*df*	res. deviance	AIC	P^*	Pseudo R^2 (%)
Intercept only	-	63.698	114.67	-	-
PC1	1	**54.974**	**107.950**	**0.003**	**13.70**
PC1 + **autoco**	1	**46.103**	**101.080**	**0.003**	**27.62**
PC1 + autoco + host	1	46.060	103.030	0.835	27.69
PC1 + autoco + host + (PC1 x host)	1	45.766	104.740	0.588	28.15
PC2	1	63.135	116.110	0.453	0.90
PC2 + **autoco**	1	**57.594**	**112.27**	**0.020**	**9.58**
PC2 + **autoco** + **host**	1	**51.844**	**108.820**	**0.020**	**18.61**
PC2 + **autoco + host + (PC2 x host)**	1	**37.369**	**96.342**	**0.0001**	**41.33**
PC3	1	63.657	116.63	0.839	0.06
PC3 + **autoco**	1	**56.630**	**111.60**	**0.008**	**11.10**
PC3 + **autoco** + **host**	1	**52.025**	**109.00**	**0.030**	**18.33**
PC3 + autoco + host + (PC3 x host)	1	51.977	110.95	0.826	18.40

**P*-values calculated using Chi-square distributed test statistics with 1 degree of freedom.

Table S4 related to Figure 3. Genetic basis of the pattern (a) and color (b) phenotypes inferred from segregation of pattern phenotypes in 822 *T. cristinae* from 35 F1 families and 1625 *T. cristinae* from 61 F1 families, respectively. Families are grouped based on the expected genotype of the parents given the observed segregation of offspring phenotypes and the hypothesized dominance relationships (see Supplementary Methods for details). *P*-values are from Fisher's exact tests comparing the phenotypic ratios observed in F1 offspring to expected phenotypic ratios under the null hypothesis.

Family class	# families	F1 offspring	unstriped – green (a) or green (b) F1s	striped – green (a) or melanistic (b) F1s	expected F1s	expected F1s	*P*
(a) pattern							
UU x ss	32	717	717	0	717	0	n/a
Us x ss	3	105	46	59	52.5	52.5	0.407
(b) color							
Gm x *Gm*	8	242	203	56	194	64	0.406
GG x ??	43	1088	1088	0	1088	0	n/a
Gm x *mm*-F	7	157	77	80	78	78	0.910
Gm x *mm*-M	3	138	111	27	69	69	<0.001

Table S5 related to Figure 3. Two-locus genotype frequencies at pattern candidate locus 1 (a) or 2 (b) compared to genotypes at the color candidate locus. Genotypes at the color candidate locus were not independent of genotypes at either of the two pattern candidate loci (FET, $P < 0.01$).

(a) Color candidate and pattern candidate 1.

	Color		
pattern 1	*GG* (green)	*Gm* (green)	*mm* (melanistic)
UU (unstriped)	8	0	1
Us (unstriped)	47	23	0
ss (striped)	188	281	54

(b) Color candidate and pattern candidate 2.

	Color		
pattern 2	*GG* (green)	*Gm* (green)	*mm* (melanistic)
UU (unstriped)	1	0	0
Us (unstriped)	39	19	0
ss (striped)	203	285	55

Table S6 related to Figure 4. Parameter combinations used in simulations testing the effect of the melanistic phenotype on levels of adaptive divergence, neutral divergence, and inter-host mating. For parameters values represented by a set of values, the position of a parameter within the set was matched with the parameter values presented in the same position across all sets with multiple parameter values.

parameter	value	rationale (e.g. previous study estimating parameter)
movement:		
sd in movement distance (melanistic; σ_m)	(0.013, 0.020, 0.027, 0.040, 0.060)	[S1]
sd of movement distance (green; σ_G)	(0.010, 0.015, 0.020, 0.030, 0.040)	[S1]
mating:		
base mating fitness	0.3000	this study; [S2]
same host mating advantage	0.1200	this study; [S2]
melanistic mating advantage (male)	0.0572	this study
melanistic mating advantage (female)	0.0800	this study
spatial mating bias	0.5	biology of flightless insect
selection:		
S - maladapted pattern phenotype	0.3000	[S3]
S - green in melanistic niche	(0.1100, 0.1005, 0.1000, 0.1000, 0.0890)*	adjusted to maintain color polymorphism over the course of simulation runs
S - melanistic in green niche	(0.2200, 0.2200, 0.2225, 0.2300, 0.2270)*	adjusted to maintain color polymorphism over the course of simulation runs
other:		
total population size (N)	4000	[S4]
number of generations	1000	-
proportion of melanistic niche	0.2	biologically realistic given ratio of leaves to wood
initial frequency of the melanistic allele	0.35	this study
recombination rate among pattern and color locus	0.1	this study

* these values were used for the ‘full’ model runs (see results in Table S7(I)) and were adjusted to facilitate color being maintained as a polymorphism within simulations run under the no-difference-in-movement, no-mating-advantage, and additive fitness models (results given in Table S7[II], Table S7[III], and Table S7[IV], respectively). See Supplemental Experimental Procedures below for details of how selection parameters changed.

Table S7 related to Figure 4. Summary of the predicted evolutionary effects of melanism in *T. cristinae*. We report summary statistics for the effect of melanistic individuals on: (1) F_{ST} at the pattern locus, (2) levels of neutral F_{ST}, and (3) the frequency of inter-host mating (IHM). Simulations were run using five different levels of individual movement ('a' through 'e') under each of the four different models included in this study (one full and three reduced models; 'I' through 'IV' below). Mean (standard deviation) gross migration rates (i.e., the proportion of individuals moving between different host plant patches) are given for each of the five levels of individual movement and results are subset by model type. Comparisons between simulations excluding (n = 200 per parameter combination) and including (n = 200 per parameter combination) melanistic individuals were assessed using Mann-Whitney U tests. Below we report the test statistic '*U*' and the associated *P*-value in brackets. Spearman's rank correlation 'rho' and associated *P*-value (in brackets) were computed to assess the effect of the frequency of melanistic alleles within a given set of simulations (n = 200 per parameter combination) on relevant evolutionary parameters.

Evolutionary metric	*U*	rho
(I) full model		
a) gross migration = 0.008 (0.0005)		
F_{ST} at the pattern locus	30245 **(0.000)**	-0.8395 **(0.000)**
neutral F_{ST}	21995 (0.085)	-0.1954 **(0.006)**
IHM	14874 **(0.000)**	0.4499 **(0.000)**
b) gross migration = 0.012 (0.0007)		
F_{ST} at the pattern locus	31873 **(0.000)**	-0.8424 **(0.000)**
neutral F_{ST}	23272 **(0.005)**	-0.1146 (0.106)
IHM	12381.5 **(0.000)**	0.5905 **(0.000)**
c) gross migration = 0.016 (0.0009)		
F_{ST} at the pattern locus	29200 **(0.000)**	-0.8704 **(0.000)**
neutral F_{ST}	23093 **(0.007)**	-0.1687 **(0.011)**
IHM	12726.5 **(0.000)**	0.7268 **(0.000)**
d) gross migration = 0.023 (0.0010)		
F_{ST} at the pattern locus	22668 **(0.021)**	-0.5736 **(0.000)**
neutral F_{ST}	22686 **(0.020)**	-0.1191 (0.093)
IHM	17860.5 (0.064)	0.5625 **(0.000)**
e) gross migration = 0.033 (0.0025)		
F_{ST} at the pattern locus	32364 **(0.000)**	-0.9374 **(0.000)**
neutral F_{ST}	27300 **(0.000)**	-0.4785 **(0.000)**

IHM	8511 **(0.000)**	0.8622 **(0.000)**

(II) no-difference-in-movement model.

a) gross migration = 0.01 (0.0005)		
F_{ST} at the pattern locus	33204 **(0.000)**	-0.8869 **(0.000)**
neutral F_{ST}	20655 (0.571)	-0.1653 **(0.019)**
IHM	14301.5 **(0.000)**	0.3872 **(0.000)**
b) gross migration = 0.015 (0.0006)		
F_{ST} at the pattern locus	29846 **(0.000)**	-0.8996 **(0.000)**
neutral F_{ST}	20785 (0.497)	-0.1629 **(0.021)**
IHM	15806.5 **(0.000)**	0.6270 **(0.000)**
c) gross migration = 0.02 (0.0007)		
F_{ST} at the pattern locus	34808 **(0.000)**	-0.9306 **(0.000)**
neutral F_{ST}	21336 (0.248)	-0.0935 (0.188)
IHM	12936 **(0.000)**	0.6573 **(0.000)**
d) gross migration = 0.03 (0.0009)		
F_{ST} at the pattern locus	27866 **(0.000)**	-0.8230 **(0.000)**
neutral F_{ST}	21595 (0.168)	-0.1006 (0.156)
IHM	15316 **(0.000)**	0.6809 **(0.000)**
e) gross migration = 0.045 (0.0010)		
F_{ST} at the pattern locus	29589 **(0.000)**	-0.9124 **(0.000)**
neutral F_{ST}	20782 (0.499)	-0.1374 (0.052)
IHM	13875.5 **(0.000)**	0.7894 **(0.000)**

(III) no-mating-advantage model.

a) gross migration = 0.008 (0.0005)		
F_{ST} at the pattern locus	27341 **(0.000)**	-0.7577 **(0.000)**
neutral F_{ST}	23567 **(0.002)**	-0.3167 **(0.000)**
IHM	13836.5 **(0.000)**	0.3777 **(0.000)**
b) gross migration = 0.012 (0.0007)		
F_{ST} at the pattern locus	27307 **(0.000)**	-0.7208 **(0.000)**
neutral F_{ST}	22472 **(0.033)**	-0.2302 **(0.001)**
IHM	14023 **(0.000)**	0.5039 **(0.000)**
c) gross migration = 0.015 (0.0007)		
F_{ST} at the pattern locus	25126 **(0.000)**	-0.6604 **(0.000)**
neutral F_{ST}	22289 **(0.048)**	-0.2366 **(0.000)**
IHM	15936 **(0.000)**	0.5626 **(0.000)**
d) gross migration = 0.023 (0.0010)		
F_{ST} at the pattern locus	23419 **(0.003)**	-0.5621 **(0.000)**
neutral F_{ST}	22718 **(0.019)**	-0.2684 **(0.000)**

IHM	18386.5 (0.123)	0.5248 **(0.000)**
e) gross migration = 0.031 (0.0020)		
F_{ST} at the pattern locus	24910 **(0.000)**	-0.6466 **(0.000)**
neutral F_{ST}	24758 **(0.000)**	-0.3689 **(0.000)**
IHM	14854 **(0.000)**	0.6374 **(0.000)**

(IV) additive fitness model.

a) gross migration = 0.008 (0.0005)		
F_{ST} at the pattern locus	23888 **(0.001)**	0.0349 (0.6233)
neutral F_{ST}	20934 (0.419)	-0.0459 (0.519)
IHM	14411 **(0.000)**	0.1114 (0.1162)
b) gross migration = 0.012 (0.0006)		
F_{ST} at the pattern locus	22948 **(0.011)**	0.1349 (0.057)
neutral F_{ST}	21988 (0.086)	-0.0024 (0.973)
IHM	14317.5 **(0.000)**	-0.1275 (0.072)
c) gross migration = 0.016 (0.0007)		
F_{ST} at the pattern locus	23994 **(0.001)**	0.1644 **(0.020)**
neutral F_{ST}	21627 (0.159)	-0.1178 (0.097)
IHM	13551 **(0.000)**	-0.0389 (0.585)
d) gross migration = 0.024 (0.0009)		
F_{ST} at the pattern locus	24910 **(0.000)**	0.2719 **(0.000)**
neutral F_{ST}	23226 **(0.005)**	-0.0258 (0.717)
IHM	12689 **(0.000)**	-0.1241 (0.080)
e) gross migration = 0.033 (0.0011)		
F_{ST} at the pattern locus	28047 **(0.000)**	0.3631 **(0.000)**
neutral F_{ST}	23052 **(0.008)**	-0.0437 (0.539)
IHM	10055 **(0.000)**	-0.3044 **(0.000)**

Supplemental Experimental Procedures

Reflectance spectra and crypsis in different microhabitat types.

Reflectance spectra were recorded with a JAZ UV-VIS spectrophotometer (Ocean Optics), PX-2 pulsed Xenon light source, and QR400-7-SR reflectance probe with an integration time of 80 ms, averaging across five scans. Raw spectra were corrected for nonlinearity, stray light, and electric dark at the time of collection with Spectral Suit software (Ocean Optics). To account for lamp drift we corrected to a standard spectralon white sample between each measurement. For each individual *T. cristinae* we recorded two reflectance spectra; one from the lateral margin of the first thoracic segment and the second from the lateral margin of the third thoracic segment.

We also collected reflectance spectra from different host plant tissues (hereafter referred to as 'microhabitat' types). Specifically, we recorded reflectance spectra from *C. spinosus* leaves (both the top and the bottom of the leaf), *A. fasciculatum* leaves (these are small enough as to not have a definable top and bottom surface), *C. spinosus* wood, and *A. fasciculatum* wood. We recorded five reflectance spectra for each of the five plant microhabitat types, as described above for individual *T. cristinae*. Finally we recorded side-welling irradiance spectra (i.e. ambient light) from within four different *A. fasciculatum* plants in the field using a calibrated JAZ-UV-VIS spectrophotometer with a cosine corrector (CC-D-DA; Ocean Optics).

From raw spectra, we calculated the mean reflectance, interpolated between 300 and 700 nm, adjusted negative values to zero, and applied triangular smoothing with a triangular distance of 10 nm using the software AVICOL [S5]. Ambient light spectra were interpolated between 300 and 700 nm and units were converted to $\mu mol*s^{-1}*m^{-2}$.

Because birds are a major predator of *T. cristinae* [S3], we quantified the relative crypsis of different *T. cristinae* morphs when observed against each of the five different microhabitats using the avian tetrachromatic color space method described by Endler and Mielke [S6]. This color space takes into account the visual system of the receiver and the

ambient light environment a specific color is viewed in. We calculated both the Euclidean distance (Δ_T statistic of [S6]) and the disparity in hue (hue disparity, as described in [S7]) between each individual *T. cristinae* and the five microhabitat types described above using the software AVICOL [S5]. Both Δ_T and hue disparity were calculated by modeling the relative excitation of ultraviolet (UV), short-wavelength (SW), medium-wavelength (MW), and long-wavelength (LW) sensitive cone types using the UVS photoreceptor curves provided by [S6] given the ambient light experienced in an *A. fasciculatum* bush and with photoreceptor sensitivity normalized to a total absorption of one. We tested for statistical differences in the degree of crypsis (i.e. color matching) between color phenotypes in each microhabitat types by fitting linear models with the 'lm' function in R [S8]

Scoring the presence and prevalence of fungal infections

Individuals were collected as juveniles and reared in net bags on the same host plant they were collected from (see main text for details) until 10 - 12 days after their final molt. Each individual was then preserved in 95% ethanol and subsequently scored for the presence of necrotic tissue indicating fungal infection. These black marks, generally located on sutures joining adjacent sclerites, are easily visible in adults, grow over time, and are more visible once individuals are preserved in ethanol (due to bleaching that occurs to the cuticle colors following preservation in ethanol).

Climatic variation among populations of *T. cristinae*

We downloaded 19 bioclimatic variables from the WoldClim database at a resolution of 1 km^2 (hereafter referred to as "bioclim" variables; http://www.bioclim.org) for each of the 20 populations shown in Table S1b. One of the climatic variables, layer 14: precipitation in the driest month, was invariable among our sampled populations and was excluded from all analyses. Because many of the remaining 18 bioclim variables were highly correlated we carried out a principal components analysis on their centered and scaled matrix using the 'prcomp' function in R [S8]. We retained the first three principal component axes (PCs) for analyses described below as they accounted for a cumulative 97.92 % of the variation in climate across our sites (see table below for PC loadings).

Table of loadings for 18 WorldClim variables ("bioclim layer") following principal components analyses carried out on variation in climate among 20 populations of *T. cristinae*.

bioclim layer	PC1	PC2	PC3
1: Annual Mean Temperature	0.3073	0.0598	-0.0960
2: Mean Diurnal Range (Mean of monthly (max temp - min temp))	0.1816	0.1382	-0.4923
3: Isothermality (BIO2/BIO7) (* 100)	0.2578	-0.2362	-0.0103
4: Temperature Seasonality (standard deviation *100)	-0.2187	0.3038	-0.0559
5: Max Temperature of Warmest Month	0.0501	0.3855	-0.2796
6: Min Temperature of Coldest Month	0.2969	-0.1337	0.0438
7: Temperature Annual Range (5-6)	-0.1583	0.3447	-0.2156
8: Mean Temperature of Wettest Quarter	0.3078	-0.0777	-0.0415
9: Mean Temperature of Driest Quarter	0.1623	0.3512	-0.1603
10: Mean Temperature of Warmest Quarter	0.1573	0.3563	-0.1600
11: Mean Temperature of Coldest Quarter	0.3029	-0.1079	-0.0332
12: Annual Precipitation	-0.2507	-0.2094	-0.2455
13: Precipitation of Wettest Month	-0.1854	-0.2340	-0.4133
15: Precipitation Seasonality (Coefficient of Variation)	0.2495	-0.1470	-0.3122
16: Precipitation of Wettest Quarter	-0.2056	-0.2452	-0.3432
17: Precipitation of Driest Quarter	-0.2359	0.2023	0.1846
18: Precipitation of Warmest Quarter	-0.3095	0.0392	0.0099
19: Precipitation of Coldest Quarter	-0.2278	-0.2308	-0.2898

Because population samples are not randomly distributed through space, we tested for spatial autocorrelation among populations in the proportion of melanistic individuals prior to constructing generalized linear models (GLMs) by computing Moran's I [S9]. These analyses indicate that there is spatial autocorrelation in the proportion of melanistic individuals observed among the populations we sampled (Moran's I, $P < 0.01$), therefore we fit generalized linear models (GLMs) that included a spatial autocovariate that accounted for spatial structure in the proportion of melanistic individuals observed within populations [S10]. Specifically, the autocovariate was computed as the weighted average

of the presence of melanistic morphs in a spatial neighborhood of k populations [S11]; therefore each population receives a value that is a function of the proportion of melanistic morphs within their neighborhood. We computed this spatial autocovariate with the 'autocov_dist' function in the R package *spdep*, version 0.5-68, and a neighborhood radius size of eight km. Models were then fit using the 'glm' function in R [S8] with a binomial error distribution and logit link function. We specifically modeled the proportion of melanistic individuals observed within the sampled populations as a function of climate (i.e., each of the first three climate PC axes), the host plant a given population was found on, the spatial autocovariate, and the interaction between climate and host plant. Each of these three terms (i.e., explanatory variables) was sequentially added to the null model and the residual deviance was computed. We then determined whether any of the explanatory variables had a significant effect on the proportion of melanistic *T. cristinae* by comparing AIC scores as well as Chi-square distributed *p*-values for models containing each of the terms against the null model.

Phenotype-dependent dispersal of color phenotypes.

To test whether there was phenotype-dependent dispersal between color phenotypes of *T. cristinae* we set up mesocosms that consisted of two *A. fasciculatum* plants each, planted 90 cm apart, and separated by soil obtained from the localities plants and insects were collected from. Mesocosms were covered in cages constructed of chicken wire to ensure that predation by birds could not occur during the experimental trials. We ran a total of eight experimental replicates (four replicates per mesocosm) by releasing four melanistic and four green *T. cristinae* onto one of the two *A. fasciculatum* plants of each mesocosm. Fresh plants were planted in each of the mesocosms after two replicates (i.e., eight experimental plants were used total). To minimize any effects that moving into a novel environment might have on the likelihood of dispersal we collected *T. cristinae*, *A. fasciculatum* plants, and soil all from the same area (population code FHA; 34.5176N, 119.8010W). Each experimental replicate was set up in the late afternoon (1600h) and dispersal was recorded the following day (0900h). Each morning we sampled experimental plants with visual surveys and by tapping each branch individually to knock any *T. cristinae* that did not disperse off the plant. We recorded the phenotype of each

individual that was recaptured (green versus melanistic) from the same plant it was released on the following day and those individuals that were not recaptured were scored as dispersing. We compared generalized linear mixed models (GLMMs) with binomial error distributions that predicted dispersal and either included or excluded color phenotype as a fixed effect to determine whether the color phenotype of individuals influenced their likelihood of dispersing. While the results of this experiment show that melanistic individuals are more likely to disperse than green individuals, we are unable to determine the mechanism underlying this difference. A number of mechanistic explanations are possible including less available woody microhabitat for melanistic individuals or simply higher levels of activity in melanistics. Future work is warranted to determine the mechanism underlying the patterns we observe here.

No-choice mating trials and mating preferences among color phenotypes.
3554 no-choice mating trials were carried out by placing a single male and single female *T. cristinae* together in a petri dish. 3320 of these trials were from previously published work [S2, S1]. The goal of these previous studies was to test for sexual isolation between populations irrespective of their phenotypic composition, and as such, these studies did not examine the role of the melanistic phenotype in any way. Moreover, the ecological and sexual relevance of the melanistic phenotype was not previously known. The 234 trials new to this study were conducted in the exact same manner, in terms of sampling and mating test protocol, as the previously published ones. In all trials, individuals were collected from the wild and paired for the mating tests randomly with respect to phenotype such that the mating trials were conducted using natural morph frequencies. 3053 of these trials were between two green individuals, 473 were between one green and one melanistic individual, and 28 were between two melanistic individuals. Each trial was scored for the presence versus absence of copulation after a one-hour period. Further details of how the trials were conducted have been previously published [S2]. To test whether the likelihood of mating success differed depending on the phenotypes included in the trials we modeled mating success as a function of the color of the two individuals included in the trial using binomial generalized linear mixed models that included population of origin as a random effect.

Quantifying cuticular hydrocarbon (CHC) variation among color phenotypes.
To obtain CHC data, live individuals were anaesthetized by freezing for one hour, and submerged in individual vials containing 1 ml of HPLC-grade hexane for 10 minutes to extract the cuticular hydrocarbons from their body surface. Samples were subsequently analyzed with a 6890 Hewlett Packard (now Agilent) gas chromatograph (GC) equipped with a DB-5 MS column (50 m ×0.25 mm i.d.), using the following temperature program: 100°C for 1 min, then 20°C per min to 280°C. This final temperature of 280°C was held for 40 min. The injector and detector temperatures were also set to 280°C.

In total we quantified 26 different mono- and dimethylated CHCs for each individual *T. cristinae*; specifically, we quantified eight pentacosanes, eight heptacosanes (including the six monomethyl-heptacosanes previously described by [S12]), and ten nonacosanes. Following standard approaches in CHC analysis [S12], we analyzed proportional rather than absolute abundances of components; allowing us to control for individual differences in body, thereby reducing experimental error (see also [S13, S14]). The total amount of each target cuticular hydrocarbon (see below) was determined by multiplying the area count of the respective chromatographic peak with 200 ng of the internal standard and by dividing the product by the area count of the internal standard. Proportional CHC components were calculated by dividing the amount of each component in a given sample by the sum of all components in that sample. These hydrocarbon proportions were then transformed using log-contrasts [S13, S15] to remove the non-independence among analyzed variables. Log-contrasts were calculated by dividing the value for each hydrocarbon by the component 5Me27 (5-methylheptacosane), and then taking the $\log_{10}$ of these new variables, resulting in 25 log-contrast transformed values for every individual. Results obtained by dividing by other components were qualitatively similar.

To reduce data dimensionality, we conducted a principal components analysis (on the correlation matrix) to reduce the number of log-contrast CHCs. Principal component (PC) axes with an eigenvalue over one were retained as variables for subsequent analyses

(resulting in three axes retained which accounted for 78.6% of the total variation; see Table below for loading of PC axes). To test for color-morph specific differences (factor with two levels: green vs. melanistic) in CHC profiles we then performed a multivariate analysis of variance (MANOVA) on individual's PC axis scores while including sex and the interaction between sex and color as additional factors.

Table of PC loadings for 25 log-contrast transformed cuticular hydrocarbons (CHCs) following principal components analyses carried out on variation in CHC profiles among 600 individual *T. cristinae*.

CHC compound	PC1	PC2	PC3
25C-compound 1	0.888	0.317	0.048
25C-compound 2	0.781	0.055	0.190
25C-compound 3	0.936	0.112	0.115
25C-compound 4	0.478	0.481	-0.345
25C-compound 5	0.827	0.122	-0.136
25C-compound 6	0.737	0.472	-0.036
25C-compound 7	0.871	0.233	-0.001
25C-compound 8	0.808	0.038	0.078
27C-compound 1	0.424	0.630	0.304
27C-compound 2	0.726	0.558	-0.167
27C-compound 4	-0.126	0.822	-0.302
27C-compound 5	0.150	0.869	-0.295
27C-compound 6	0.830	0.433	0.191
27C-compound 7	-0.325	0.798	0.294
27C-compound 8	0.808	0.278	0.142
29C-compound 1	-0.776	0.531	0.189
29C-compound 2	-0.595	0.735	-0.089
29C-compound 3	-0.388	0.820	-0.269
29C-compound 4	-0.692	0.266	-0.197
29C-compound 5	-0.671	0.660	0.077
29C-compound 6	-0.492	0.718	-0.130
29C-compound 7	0.079	0.829	-0.043
29C-compound 8	0.548	0.473	0.333

29C-compound 9	-0.804	0.311	0.395
29C-compound 10	-0.554	0.525	0.463

Quantitative divergence of the color polymorphism among populations of *T. cristinae*.

SNPs that we identify as being associated with pattern and color phenotypes in this study [see genome-wide association mapping sections] were not present it population genetic data sets we have from previous work in *T. cristinae* [S16, S17], precluding a direct comparison between differentiation at these candidate SNPs and putatively 'neutral' levels of genomic differentiation. As such, we adapted the method implemented in [S18] to estimate F_{ST} at the locus underlying the color polymorphism as:

$$F_{ST-color} = \frac{V_B}{V_B + 2V_A}$$

where V_B is the variance in genotypes between populations and V_A is the additive variance within each population. We estimated V_B and V_A using equations (5) and (6) from [S18] for each of 28 pairwise comparisons between the 8 populations. Namely, variance components were estimated as:

$$V_B = 0.5 \times 2(p_x - p_y)^2(1 - d(1 - p_x - p_y))^2$$

and

$$V_A = 0.5 \times \sum_{i=X,Y} 2\ p_i(1 - p_i)(1 - d(1 - 2p_i))^2$$

where p_i is the frequency of the recessive allele in population *i* and *d* is the dominance coefficient. We estimated p_i from the frequency of melanistic individuals in each of the 8 populations that we have estimates of genome-wide F_{ST}, assuming Hardy-Weinberg equilibrium, as $\sqrt{P_i}$. All estimates of F_{ST} were calculated assuming complete dominance ($d = 1$). For each of the 28 population pairs for which we estimated $F_{ST\text{-}color}$ we also have previously published estimates of F_{ST} at 86,130 SNPs distributed across the genome (Table S2 of [S19]). Therefore, we examined patterns of divergence in the proportion of melanistic individuals within each population ($F_{ST\text{-}color}$) by comparing $F_{ST\text{-}color}$ with genome-wide estimates of F_{ST}.

Segregation of pattern and color polymorphisms in F1 families.

To estimate inheritance patterns and dominance relationships at the locus (or loci) controlling pattern and color, we analyzed segregation of these two phenotypes in 61 F1 families generated by crossing *T. cristinae* collected from the wild. We captured individuals as sub-adults (1 – 3 molts from the penultimate molt) and raised them in captivity until adulthood. Sub-adults were kept in 300 ml containers, grouped by sex (5 – 30 individuals per container), and fed *ad lib*. with fresh clippings of the host plant *C. spinosus* that were replaced every two days. Upon reaching sexual maturity (through April and May) crosses were set up using one of two methods. For approximately half the crosses we placed a single male and single female *T. cristinae* into a bag made of mosquito netting, recorded their phenotypes, and placed the bag over an approximately 30 cm long piece of *C. spinosus* in the wild. The bags were placed in a location where *T. cristinae* are known to occur (34.51753 N, 119.80125 W). We added ~ 200g of soil into each net bag and sealed the bags. Because *T. cristinae* is univoltine, the adults mate and the female lays eggs until she dies in the bag. The eggs then go through a period of diapause from July to January, and typically hatch during the month of February. The breeding bags were thus left in the field for 9 months. We returned to the breeding bags in early March of the following year to collect any F1 offspring present. The F1 offspring were then brought into the lab and raised in the same fashion as their parents. At approximately the 5^{th} instar we recorded the phenotype of each F1 offspring.

For the second half of crosses, a single adult male and single adult female were placed together in deli containers that contained a 1 cm layer of soil and a 10 cm clipping of the *C. spinosus* host plant. Fresh host plant was supplied every other day and the eggs produced were collected over the course of two months. The eggs of individual pairs were then placed on the dirt of a planted *C. spinosus* host plant in the greenhouse. These plants were surrounded by a cage constructed of mosquito netting and left in a green house at temperatures between 10 and 25 degrees Celsius until the eggs hatched four to six months later. Upon emergence, F1 offspring were collected and their phenotypes were recorded.

We first explored the likelihood that variation in pattern and color phenotypes were controlled by three alleles segregating at a single locus by comparing segregation of phenotypes in seven F1 families (see below for justification of use of these seven). Based on dominance relationships at SNPs associated with pattern and color variation (see following paragraphs) we specifically tested the likelihood of a three allele mechanism of inheritance with the green phenotype dominant to the green-striped phenotype and both green phenotypes dominant to the melanistic phenotype (i.e. green > green-stripe > melanistic). The seven families used to test this hypothesis were the result of a cross between a green-striped and a green parent and all produced some melanistic F1 offspring. Under the three-allele hypothesis the parents of these families would need to be heterozygous and carrying a melanistic allele; therefore, the second allele at the single locus would need to be the green-striped allele (in the green-striped parent) or the green allele (in the green parent). Under the dominance scenario stated above segregation of phenotypes in the F1 families are expected to be 2 green: 1 green-striped: 1 melanistic. We tested for departures from the expected ratio of F1 offspring phenotypes using Fisher's exact tests (FETs) in R [S8].

Given the result that variation in pattern and color phenotypes were not controlled by a single locus with three alleles (see results of three-allele analysis), we estimated inheritance and dominance patterns at the locus (or loci) underlying variation in pattern phenotypes using the frequency of green-unstriped and green-striped F1 offspring observed in the 35 families generated by a cross between an unstriped-green and striped-green individual. Specifically, we tested whether inheritance of dorsal stripe patterning was consistent with inheritance of a single Mendelian locus with the *u*nstriped phenotype and allele (*U*) dominant to the *s*triped phenotype and allele (*s*). We chose this model based on both the results of our genome-wide association (GWA) mapping analyses and the fact that the dorsal striped phenotype is frequently scored as a discrete trait with high repeatability [S2, S4, S20]. While we were unable to explicitly determine the genotypes of the parents of these families, we were able to objectively predict phenotypic ratios of the F1 offspring that are expected under the null hypothesis stated above. For example, if

a cross was between one parent that was homozygous *UU* and another homozygous *ss* we expect all offspring to be heterozygous *Us* and show an unstriped phenotype. However, this pattern does not allow us to differentiate between scenarios where a single locus controls pattern versus one where patterning is controlled by more than one locus. The other possible genotypic combination of an unstriped and striped parent is *Us* x *ss*. If pattern phenotypes are controlled by a single locus with *U* dominant to *s* we expect a segregation ratio of pattern phenotypes of 1 unstriped-green: 1 striped-green in F1 offspring. We tested for departures from these expectations in the observed phenotypic ratios of F1 offspring with Fisher's exact tests in R [S8]. We note that although this was a somewhat tautological approach to classifying F1 families these results from crosses were supported by GWA mapping analyses (see main text), and are thus not intended as the sole test of inheritance of the pattern and color phenotypes.

We carried out a parallel analysis to determine patterns of inheritance and dominance at alleles controlling color phenotypes within 61 F1 families. As for the pattern phenotype, we were unable to determine genotypes of the parents in these crosses; therefore, we objectively classified families based on the observed phenotypes of both parents and the segregation of phenotypes in the F1 offspring. We tested the null hypothesis that the color polymorphism is controlled by a single locus with the green allele (*G*) being dominant to the melanistic allele (*m*; see main text for details). We chose to test this model of inheritance because the color phenotypes segregate as a discrete polymorphism in the wild. Specifically, we classified families as "heterozygous x heterozygous" (*Gm* x *Gm*; $n = 8$) if they contained at least 1 melanistic offspring in the F1 generation and were the result of a cross between two green parents, "homozygous x unknown" (*GG* x ??; $n = 43$) if they contained no melanistic F1s and were the result of a cross between two green parents, "heterozygous x homozygous female" (*Gm* x *mm*-F; $n = 7$) if they contained at least one melanistic F1 and were produced by a cross between a green male and melanistic female parent, and "heterozygous x homozygous male" (*Gm* x *mm*-M; $n = 3$; see Table S5 for details) if they contained at least one melanistic F1 and were produced by a cross between a melanistic male and a green female. No families were produced as

the result of a cross between two melanistic parents. We tested for departures from the expected phenotypic ratios observed within the F1 offspring with FETs in R [S8].

Digital images of individuals used for genome-wide association (GWA) mapping. To characterize the genetic basis of pattern and color polymorphisms we carried out genome-wide association (GWA) mapping analyses using 602 adult *T. cristinae* sampled from a phenotypically variable population found on the host plant *Adenostoma fasciculatum* (population code: FHA, N34°30.958, W119°48.050). Individuals were sampled as sub-adults using insect sweep nets. We then raised individuals in 500 ml plastic containers fed *ad libitum* on freshly collected clippings of *A. fasciculatum* until adulthood. Upon reaching adulthood we recorded digital images of each individual. We captured digital images of each individual using in GWA mapping in RAW format with a Canon 600d equipped with an EF-S 60mm F2.8 macro lens (Canon (UK) Ltd., Surry, UK) and two wireless external flashes (Yongnuo YN560-II speedlight, Yongnuo Digital, www.yongnuo.eu). Each image was captured with the camera set on manual, an aperture of f/14, a shutter speed of 1/250s, and flashes set to 1/8 power in manual mode. Each image also included a standard color chip (Colorgauge Micro, Image Science Associates LLC, Williamson, NY, USA). Following image capture each individual was placed in an individually labeled 1.5 ml microcentrifuge tube and preserved in 95% ethanol. Prior to any pattern or color analyses the red, green, and blue color channels of each image were linearized and corrected to 80% reflectance using a neutral gray color target (target #10) of the standard color chip with Adobe's Photoshop Lightroom 4 software (Adobe Systems Software Ireland Ltd).

To estimate the proportion of the dorsal body area that was striped we measured the size of the dorsal stripe and the size of total dorsal body area in pixels from the digital photographs described above using the polygon selection tool in the software ImageJ, version 1.46 [S21]. We then divided the area of the dorsal stripe by the total dorsal body area to obtain an estimate of the size of the dorsal stripe. We limited measuring the area of the dorsal stripe to green-colored (i.e., non-melanistic 'striped' or 'unstriped') individuals because the striped phenotype is not expressed in melanistic *T. cristinae* (n

green individuals = 546). While the presence or absence of the dorsal stripe can be reliably scored as a discrete character [S4, S22] it can also be more accurately quantified via the approach taken here and in previous studies [S3, S23]. Prior to GWA mapping analyses we corrected dorsal pattern measurements for differences between the sexes by extracting residuals from linear models with sex as the independent variable. Sex-corrected residuals were then centered and standardized by normal quantile transformation as suggested by [S24]. For each individual we also recorded color as a binary phenotype, with green individuals being scored as "0" and melanistic individuals scores as "1".

We extracted genomic DNA from three to five legs of each of the 602 individuals used for phenotypic sampling using DNeasy Blood and Tissue Kits (Qiagen). We then generated individually barcoded restriction-site associated DNA libraries for each individual using previously published protocols [S25] which have been applied to *T. cristinae* specifically [S16]. DNA sequencing of seven pooled and size-selected libraries was accomplished on seven lanes (one library per lane) of the Illumina HiSeq2000 platform with V3 reagents at the National Center for Genome Research (Santa Fe, New Mexico, USA).

SNP calling and genotyping

After removing barcodes and the following six base pairs of the EcoRI cut site using a custom Perl script developed and implemented in a previous study [S16] we obtained a total of 881,432,319 DNA sequences with an average length of ~83.5 bp (95% interval=63-86 bp). This script also identified barcodes that were away by 1 bp due to synthesizing or sequencing errors and removed adapters at the 3' end when present. The mean number of reads per individual was 1,464,173 (95% interval=706,596-2,235,066). We aligned 94.01% of the reads (828,709,553) to the reference genome [S17] using bowtie2 version 2.1.0 [S26] with the local model and the '--very-sensitive-local' preset (-D 20 -R 3 -N 0 -L 20 -i S,1,0.50). We used samtools version 0.1.19 [S27] to sort and index alignments. Variants were called using samtools mpileup and bcftools using the full prior and requiring the probability of the data to be less than 0.5 under the null

hypothesis that all samples were homozygous for the reference allele to call a variant. Insertion and deletion polymorphisms were discarded. We identified 2,124,595 single nucleotide polymorphisms (SNPs) that were reduced to 524,832 SNPs after discarding SNPs for which there were sequence data for less than 40% of the individuals, low confidence calls with a phred-scale quality score lower than 20, and SNPs with more than two alleles. Average depth of the retained SNPs across all individuals was ~3537x (mean coverage per SNP per individual ~ 6x). We used a custom Perl script to calculate empirical Bayesian posterior probabilities for the genotypes of each individual and locus using the genotype likelihoods and allele frequencies estimated by bcftools along with Hardy-Weinberg priors (i.e. $p(A)=p^2$; $p(a)=(1-p)^2$; $p(Aa)= 2p(1-p)$). Finally, we calculated the posterior mean genotype for each individual, at each locus, as two times the probability of the homozygous minor allele genotype plus the probability of the heterozygous genotype. These imputed genotypes were used for all multi-locus GWA mapping analyses.

Multi-SNP GWA mapping

GWA mapping was carried out using Bayesian sparse linear mixed models (BSLMMs) as implemented in the software *gemma* [S28]. BSLMMs take a multiple-SNP Bayesian approach to model the genetic architecture of traits while controlling for relatedness of individuals within the sample. A major advantage of BSLMMs is that instead of modeling the phenotypic effects of SNPs individually, effects of SNPs are modeled as coming from a mixture of two normal distributions [S28]. As such, BSLMMs run in *gemma* provide estimates of the proportion of phenotypic variation that can be explained by the combined effects of polygenic (i.e., infinitesimal) and 'large' (i.e., detectable) effect SNPs (proportion of total variance explained; PVE), the proportion of the explained phenotypic variation that is due to the effects of large effect SNPs alone (PGE), and the number of large effect SNPs underlying the phenotype of interest (n-SNP).

We ran *gemma* with the default SNP quality control options and excluded SNPs with a minor allele frequency < 0.01 (346,788 SNPs retained for analyses carried out for the colour trait and 346,660 SNPs retained for analyses carried out on pattern traits). As

recommend for quantitative versus discrete traits, we ran linear BSLMMs for the dorsal pattern trait and probit BSLMMs for color. All analyses were carried out with 10 independent Markov-chain Monte Carlo (MCMC) chains that were run for 20 million steps following a burn-in period of 5 million steps. Parameter states were recorded every 10 steps and written every 1000 steps. During these analyses, all other options in *gemma* were set to default values.

In addition to estimating hyperparameters that describe the genetic architecture of traits, *gemma* provides posterior inclusion probabilities (PIPs; γ parameter in the BSLMMs) for individual SNPs that reflect the fraction of MCMC iterations of the BSLMM for which a given SNP was estimated to have a non-zero large effect on phenotypic variation [S28].

Single-SNP GWA mapping

In addition to multi-locus BSLMM GWA mapping, we conducted a single-SNP GWA mapping analysis using the EIGENSRAT method of Price et al. [S29] in the R package *GenABEL* [S30]. Prior to single-SNP GWA mapping analyses, we generated discrete genotypes for each individual, at each SNP, from the imputed genotype values used for multi-SNP GWA mapping. Specifically, we converted imputed genotypes ranging from 0 to 0.6 (inclusive) to the homozygous reference genotype (coded as '1'), imputed genotypes ranging from 0.6 to 1.4 (exclusive) to the heterozygous genotype (coded as '2'), and imputed genotypes ranging from 1.4 (exlusive) to 2 (inclusive) to the non-reference genotype (coded as '3'). These discrete genotypes were used for all single-SNP analyses described below.

We carried out three rounds of quality control (QC) on our data prior to single-SNP GWA mapping. First we used the *check.marker* function in *GenABEL* (note that we allowed for violation of Hardy-Weinberg equilibrium by setting the *p-level* option to zero during this first step) to remove SNPs with minor allele frequencies less than 0.005, individuals with call rates < 0.95, individuals with the proportion of alleles identical-by-state (IBS) > 0.95, and individuals with abnormally high levels of heterozygosity (false discovery rate < 0.01). Following the first round of QC we removed individuals with

abnormal relatedness values by performing multidimensional scaling (MDS) analysis on a distance matrix generated from a matrix of genomic kinship coefficients estimated by the *ibs* function in *GenABEL*. We plotted individual's scores for the first two axes of variation identified by the MDS analysis and removed those individuals with abnormal relatedness values (i.e., those that clustered separately from the main cluster of individuals). Finally, we removed SNPs that were out of Hardy-Weinberg using the *check.marker* function setting the *p-level* option to 0.0001. Following QC we retained 357,385 SNPs for GWA mapping using the EIGENSTRAT method of [S29], implemented using the *egscore* function in *GenABEL*. For these analyses we used the first 5 axes of genetic variation, generated from principle components analysis of the genomic kinship matrix, to correct for population structure.

Dominance relationships at SNPs associated with pattern and color

We calculated dominance relationships between alleles at each candidate locus using the *d/a* statistic [S31, S32]. Here *d* is the dominance effect, calculated as the mean phenotype of the heterozygous genotype class subtracted from half the distance between the mean phenotype of the two homozygous genotype classes. *a* represents the additive effect of an allele at a given locus, calculated as half the phenotypic difference between homozygous genotypes. Phenotypic measurements of dorsal pattern and color traits are as described above. *d/a* values for strictly dominant or recessive alleles are expected to be 1 or -1, respectively, while for additive gene effects, *d/a* values are expected to be 0. Here we adhere to previously followed conventions when interpreting *d*/*a* and classify alleles as dominant when *d*/*a* is greater than 0.75, recessive when *d*/*a* is less than -0.75, partly recessive when between -0.25 and -0.75, partly dominant when between 0.25 and 0.75, and additive when *d/a* is between -0.25 and 0.25 [S31, S32].

Patterns of linkage disequilibrium among trait-associated SNPs

We computed pair-wise r^2 among all SNPs in four different groups of SNP: (1) the top two candidate SNPs associated with dorsal patterning (i.e. those with PIP > 0.5), (2) the top two SNPs associated with dorsal patterning and the top one associated with coloration, (3) 1000 randomly selected SNPs mapping to linkage group 8 (LG 8), and (4)

1000 randomly selected SNPs from across the genome (genome-wide). Pair-wise r^2 was computed using the *r2fast* function of the *GenABEL* package in R [S8]. In addition to r^2, we computed D' using the *dprfast* function of the *GenABEL* package. Finally, we assess patterns of LD by testing whether genotypes at color and pattern candidate SNPs segregated independently in our sample of 602 *T. cristinae* used for GWA mapping using Fisher's Exact Tests (Table S6).

Modeling the effects of the melanistic polymorphism on evolutionary dynamics

We developed a spatially explicit, individual-based, diploid, separate-sex model around our empirical results. This model simulates the evolution of *T. cristinae* populations living on two host plant species and monitors the movement of individuals and their alleles between patches of host plants. In nature, both *A. fasciculatum* and *C. spinosus* are widespread and long-lived plants growing in the chaparral of California. Notably, patches of these plants are spatially arranged as a mosaic across the landscape. As such, we adopted a two-dimensional patch structure in our model with patches of host plant species randomly distributed across the landscape. The rationale behind developing this model was to quantitatively predict if, and to what extent, the presence of the melanistic phenotype could: (1) facilitate gene flow at the adaptive locus controlling pattern phenotypes, (2) influence levels of neutral genetic divergence among populations of *T. cristinae* adapted to either *A. fasciculatum* or *C. spinosus* hosts, and (3) affect levels of inter-host mating, a measure of the degree of reproductive isolation between populations on different host species. Although it might be considered intuitive that the combined effects of higher dispersal tendency and mating advantage of melanistic individuals would reduce adaptive divergence, it is not clear the extent to which each of these factors on their own could prevent adaptive differentiation. We addressed this by constraining various model parameters to be equal (details below). Additionally, previous theoretical work has shown that the effects of selection on neutral differentiation can be nuanced and context-dependent, especially for neutral loci unlinked to those under selection [S33, S34]. Thus, it is more difficult to predict *a priori* how the existence of melanistic individuals will affect neutral divergence, an issue our model addresses. Simulation source code was written by SMF in the C programming language and will be publicly

archived with documentation at http://sourceforge.net/projects/TimemaMelanisticModel/files/ upon acceptance for publication. Random number generation in the model was accomplished with the Mersenne Twister [S35].

Life cycle: The model considered non-overlapping generations (*T. cristinae* are univoltine insects with 1 generation per year and non-overlapping generations). Each generation consisted of migration, followed by viability selection, followed by mating and replacement of adults by offspring. Population size was held constant as follows: (1) there was no mortality during migration, and (2) if *N* individuals were present on a plant prior to viability selection, *N* offspring were produced on that plant. Drift was included due to the probabilistic nature of migration, mortality, mating, and gamete formation. Mating was non-random with respect to phenotype as detailed below. Hence, there were two steps at which selection could act: viability selection and sexual selection due to "soft" male mate choice (see "*Migration"* below).

Genomes of individuals: Each individual possessed a diploid genome modeled as a sequence of zeros and ones at each genomic position (representing alleles at a locus). In the results shown in the main text, we focused on three loci: the color locus, the pattern locus, and an unlinked, neutral locus. Each of these had two alleles (color locus: *G*reen and *m*elanistic alleles; pattern locus: *U*nstriped and *s*triped alleles; capital letter denotes dominant allele). A fourth "locus" was also considered to represent sex chromosomes and thus enabled an XX-XO chromosomal sex determination system (as occurs in *Timema* [S36]) to be included in the model.

Spatial arrangement of environments and selection: The simulation environments were set up such that equal frequencies of host plant patches (i.e. 50% of host plant patches were *A. fasciculatum* and 50% were *C. spinosus*) were randomly arranged in a two-dimensional grid measuring three patches wide and six patches deep. Each patch (i.e. host plant) was further divided into two niches representing the wood (20% of each patch) and leaves (80% of each patch) of a host plant. Patches were considered to be one unit wide

and one unit long. The woody niche within each patch occupied the left-most 20% of this square.

Migration: Each generation, levels of movement of each individual was determined with two random number draws. Direction (expressed in radians) was sampled from a uniform distribution on the interval [0, 2π). Distance was drawn from a normal distribution with a mean of zero and standard deviation (SD) of σ_m for melanistic individuals or σ_G for green individuals (to reflect our empirically documented differences in dispersal between color morphs). Hence, the gross migration rates for each color were directly proportional to σ_m and σ_G (respectively), with a standard deviation of x units ($x << 1$) leading to a gross migration rate (i.e., probability an individual changed patches) on the order of ~0.5x. We present movement parameter values σ_m and σ_G used in each simulation in Table S7 and observed gross migration rates under different simulation scenarios in Table S8. Borders of the environment (i.e., exterior edges) were treated as hard boundaries; hence, if an individual's random movement took it outside the habitat, it was placed on the edge instead.

Viability selection: After movement, individuals that were mismatched to their host plant or niche survived with probability $1 - S$ (not to be confused with the abbreviation for the striped pattern allele *s*) according to values of S detailed in Table S7. Viability selection acted against green individuals (i.e. *GG* or *Gm* genotypes at the color locus) in the wood niche of both host plant types, against melanistic individuals (i.e. *mm* genotype at the color locus) in the green niche of both host plant types, and against green individuals with the maladaptive pattern phenotype in the green niche of either host plant (i.e. *UU* and *Us* on *A. fasciculatum* and *ss* on *C. spinosus*).

Reproduction: Each generation individuals reproduced and produced a number of offspring equal to the number of offspring in a patch before viability selection. Each offspring's genotype was determined from parents sampled from that same patch by generating a haploid gamete from each parent assuming independent assortment of alleles at different loci. To most accurately reflect the empirical knowledge of the system [S1,

S2], the formation of mating pairs worked as follows (i.e., all of the following choices and assumptions are based upon our current best empirical understanding of system). For each offspring produced we randomly sampled a male from within the patch. If the randomly sampled male was melanistic he was guaranteed to successfully produce an offspring. To reflect differences in mating success among green and melanistic males we observed in our no-choice mating trials, randomly chosen males that were green successfully produced offspring with probability 0.84. We then paired the male with a randomly selected female. To reflect the spatial component of successful mating (*T. cristinae* are flightless and are more likely to mate with males in close proximity), males were 50% more likely to mate with a female if that female originated from within the same niche as the male. Mating success was then determined based on the female's phenotype. To reflect levels of reproductive isolation that exist between populations of *T. cristinae* adapted to different host plant environments [S2], successful mating was 25% less likely to occur if the sampled female originated from the different host plant environment as the male (this imposes a fitness cost on females immigrating from the alternate host environment). To reflect the mating advantage we identified for melanistic females in this study, successful mating was 16% more likely if the female was melanistic. The likelihood of a pair of individuals successfully producing an offspring given the conditions outlined above was scaled to lie between 0 and 1 (melanistic females originating from the same niche, on the host plant as the male had a likelihood of mating of 1). If a randomly selected male did not successfully produce an offspring with a randomly selected female, we iterated the 'female choice' procedure for a maximum of 5 trials after which the male produced an offspring with the next randomly selected female.

Simulation initialization and model parameterization: Simulations were initiated with maximal divergence among host plant environments at all loci other than the color locus (i.e. fixed allele frequency differences and $F_{ST} = 1$ for loci other than the color locus) and were carried out for 1000 generations. We do not incorporate novel mutation in our model because our goal here is to determine the influence of melanistic individuals, which exist at moderate frequencies in most wild populations, on facilitating gene flow

and breaking down adaptive divergence among populations adapted to different host plant environments.

Parameterization of our model was determined using empirical estimates of parameter values wherever possible. For example, the mating scheme outlined above was determined based on results from mating trials carried out both for this study and in a previous study [S2]. We carried out simulations under four different variants of the model described above. We first ran the most biologically realistic model given our empirical understanding of the biology and ecology of *T. cristinae* (herein referred to as the 'full' model). Simulations under the full model were carried out as described in the previous sections and using the starting parameter combinations described in Table S7. We then carried out simulations in modified versions of the full model described above (hereafter referred to as 'reduced' models). In reduced models we modified key features of the full model describing differences between melanistic and green individuals (e.g., degree of movement, mating advantage, or dominance relationships at candidate loci). The motivation here was to test how individual features of the full model that differed between phenotypes, rather than their combination, affect divergence.

The first reduced model lacked a difference in movement between green and melanistic individuals (herein referred to as the "no-difference-in-movement" model). The second reduce model lacked a difference in an individual's likelihood to successfully mate based on their color phenotype (herein referred to as the "no-mating-advantage" model), and the third reduced model had a modified fitness scheme such that the dominance relationships of alleles at pattern and color phenotypes were removed and phenotypic effects of alleles were strictly additive (herein referred to as the "additive fitness" model). We ran simulations under the no-difference-in-movement, no-mating-advantage, and additive fitness models to determine whether the effects of melanistic individuals we observe under the full model are driven (or influenced strongly) by either the ability of melanistic individuals to move more readily between environments, their higher mating success, or dominance and epistasis relationships among alleles at pattern and color loci, respectively. We explored how general our model findings were for all four models

considered here by varying the amount of movement individuals underwent each generation (i.e., parameters controlling movement of green and melanistic individuals; simulations carried out for five different levels of movement).

The only other parameters in the reduced models that we varied relative to the full model were the strength of selection against the melanistic phenotype in the green niche and the strength of selection against the green phenotypes in the melanistic niche (see Table S7). Changes in these selection parameters were adjusted between simulations carried out using different parameter combinations to ensure that color was maintained as a polymorphism for at least 1000 generations in the majority of simulations (see Table S7). We report the selection parameters used for simulations carried out under the full model in Table S7. For simulations carried out under the no-difference-in-movement model selection against the green phenotype in the melanistic niche was 0.1100, 0.1005, 0.1000, 0.1000, or 0.0880 and selection against the melanistic phenotype in the green niche was 0.2200, 0.2220, 0.2228, 0.2320, or 0.2370. Each of the five values in these sets represent the selection coefficient used at a given rate of movement (e.g., simulations carried out at the highest rate of movement [$m = M = 0.06$] would have been carried out using the selection parameters 0.0880 and 0.2370). For simulations carried out under the no-mating-advantage model, selection against the green phenotype in the melanistic niche was 0.3850, 0.3750, 0.3600, 0.3300, or 0.3000 and selection against the melanistic phenotype in the green niche was 0.0900, 0.0900, 0.0900, 0.0900, or 0.1200. For simulations carried out under the additive fitness model, selection against the green phenotype in the melanistic niche was 0.1500, 0.1500, 0.1500, 0.1500, or 0.1500 and selection against the melanistic phenotype in the green niche was 0.2180, 0.2180, 0.2180, 0.2160, or 0.2150. We ran 200 independent simulations for each parameter combination with no melanistic alleles present and 200 independent simulations with an initial frequency of melanistic alleles of 35% (reflecting the mean frequency of the melanistic phenotype in nature).

Supplemental Results

Segregation of pattern and color polymorphisms in F1 families

Segregation of phenotypes was highly inconsistent with the scenario of a single locus with three alleles controlling variation in color and patterning. Specifically, phenotypes of 205 F1 offspring, produced from seven crossing families, segregated in a ratio of 153 green: 3 green-striped: 49 melanistic, a significant departure from the expected ratio of 2 green: 1 green-stripe: 1 melanistic (FET; $P < 0.001$).

We analyzed phenotypic segregation in 35 F1 families to test whether pattern phenotypes segregated according to a model of simple Mendelian genetic inheritance with the unstriped allele (*U*) dominant to the striped allele (*s*). Segregation of pattern phenotypes in 32 of the 35 F1 families that were the result of a cross between an unstriped and a striped parent resulted in all F1 offspring lacking a dorsal stripe (n-offspring = 717; expected ratio if parents were genotypes *UU* and *ss*; Table S5). Segregation ratios in the remaining three families (n-offspring = 105) did not differ from a ratio of 3:1 (FET, $P > 0.05$; Table S5), supporting a simple genetic control of patterning by a single locus with the *U* allele dominant to the *s* allele.

Segregation ratios of color phenotypes in the vast majority (58 of 61) of F1 mapping families did not differ from expectations generated by a model of inheritance with color being inherited as a simple Mendelian trait with the green phenotype dominant to the melanistic phenotype (FET, $P > 0.05$; Table S5). However, segregation of color in three families that were the result of a cross between a green female and a melanistic male parent resulted in an excess of green F1 offspring (111 observed versus 69 predicted; FET, $P < 0.001$; Table S5). This deviation from expectations was not observed for crosses between green males and melanistic females (FET > 0.1; Table S5). Combined, these results suggest that in crosses between a green female and melanistic male a sex-linked modifier locus may decrease expression of the recessive melanistic phenotype, possibly do to sex-linked incomplete dominance; however, given the low sample size (three families) and modest departure from the expected ratio of 1 green: 1 melanistic (see Table S5), further work, such as gene expression and functional approaches, are warranted.

Multi-SNP GWA mapping

GWA mapping shows that both color and pattern polymorphisms are controlled by a simple genetic architecture. For color (analyzed with probit BSLMMs), our genotypic data can explain 95.25% of total phenotypic variation (95% equal tail posterior-probability interval [ETPPI] for PVE = 81.01% - 99.94%). Of this explained variation, 94.73% (ETPPI for PGE = 80.79% - 99.83%) can be explained by 7 SNPs (ETPPI for n-SNP = 2 - 20) with large phenotypic effects. For pattern (analyzed with linear BSLMMs), our genotypic data can explain 49.89% (ETPPI = 33.62% - 75.53%) of total phenotypic variation. Of this explained variation, 70.80% (ETPPI = 42.24% - 97.92%) can be explained by 4 (ETPPI = 1 - 18) SNPs with large phenotypic effects. The top candidate SNPs identified by GWA mapping of color (defined as those with a posterior inclusion probability [PIP] > 0.5), mapped to an intron of a gene encoding a cysteinyl-tRNA synthetase, a protein implicated in forming precursor proteins required for the production of melanin pigments [S37]. The two candidate SNPs associated with pattern (with PIP > 0.5) were located within an exon of a predicted gene lacking functional annotation and a non-coding genomic region, respectively.

Patterns of linkage disequilibrium among trait-associated SNPs

Analyses of pair-wise LD between different groups of SNP show that candidate SNPs associated with color and pattern phenotypes are in elevated LD relative to genome-wide expectations. LD between the two top pattern candidate SNPs was high (r^2 = 0.591; D' = 0.897), suggesting they tag the same functional variant. By contrast, LD between the pattern candidate SNP and color candidate SNPs was lower (r^2 = 0.038 and 0.041; D' = 1.00 and 0.837), but elevated when compared to genome-wide expectations (mean [90% empirical equal-tail probability interval]: r^2 for 1000 randomly selected SNPs on LG 8 = 0.007976867 [0.00004009348, 0.03348607] and D' = 0.451967605 [0.02961785, 1.00000000]; r^2 for 1000 randomly selected SNPs from across the genome = 0.006647494 [0.00004009348, 0.0307725] and D' = 0.503476863 [3.355100e-02, 1.0000000]). We note that in terms of comparing the relative LD of different sets of SNPs (e.g., candidate SNPs to genome-wide expectation) using D' as a measure of LD

results in qualitatively similar patterns as those obtained using r^2. However, estimates of D' do not account for differences in allele frequencies between the two SNPs being compared and provide elevated estimates of LD when low-frequency variants exist [S38]. Importantly, the inflated D' estimates (relative to r^2) observed between pattern and color candidate SNPs is likely driven by the almost complete absence of *UU* genotypes in our sample (see Table S6). As such, we focus on estimates of LD derived from r^2 in the main text.

Supplemental References

S1. Nosil, P., Crespi, B. J., and Sandoval, C. P. (2003). Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. Proc. R. Soc. B Biol. Sci. *270*, 1911–1918. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1691465&tool=pmcentrez&rendertype=abstract [Accessed March 12, 2012].

S2. Nosil, P., Crespi, B. J., and Sandoval, C. P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature *417*, 440–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12024213.

S3. Nosil, P., and Crespi, B. J. (2006). Experimental evidence that predation promotes divergence in adaptive radiation. Proc. Natl. Acad. Sci. U. S. A. *103*, 9090–9095.

S4. Sandoval, C. P. (1994). The effects of the relative geographic scales of gene flow and selection on morph frequencies in the walking-stick *Timema cristinae*. Evolution (N. Y). *48*, 1866–1879.

S5. Gomez, D. (2006). AVICOL, a program to analyse spectrometric data. Last Updat. Oct. 2011 *Free execu*.

S6. Endler, J. A., and Mielke, P. W. (2005). Comparing entire colour patterns as birds see them. Biol. J. Linn. Soc. *86*, 405–431. Available at: http://doi.wiley.com/10.1111/j.1095-8312.2005.00540.x.

S7. Stoddard, M. C., and Prum, R. O. (2008). Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. Am. Nat. *171*, 755–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18419340 [Accessed March 8, 2012].

S8. Core Team, R. (2013). R: A language and environment for statistical computing. R Found. Stat. Comput. *1*. Available at: http://www.r-project.org.

S9. Gittleman, J. L., and Kot, M. (1990). Adaptation: Statistics and a null model for estimating phylogenetic effects. Syst. Zool. *39*, 227–241. Available at: http://sysbio.oxfordjournals.org/cgi/doi/10.2307/2992183.

S10. Dormann, C. F., McPherson, J. F., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. D., Hirzel, A., Jetz, W., Daniel Kissling, W., et al. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography (Cop.). *30*, 609–628. Available at: http://doi.wiley.com/10.1111/j.2007.0906-7590.05171.x [Accessed January 20, 2014].

S11. Augustin, N. H., Mugglestone, M. A., and Buckland, S. T. (1996). An autologistic model for the spatial distribution of wildlife. J. Appl. Ecol. *33*, 339–347.

S12. Schwander, T., Arbuthnott, D., Gries, R., Gries, G., Nosil, P., and Crespi, B. J. (2013). Hydrocarbon divergence and reproductive isolation in *Timema* stick insects. BMC Evol. Biol. *13*, 151. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3728149&tool=pmcentrez&rendertype=abstract [Accessed May 23, 2014].

S13. Blows, M. W., and Allan, R. a (1998). Levels of mate recognition within and between two Drosophila species and their hybrids. Am. Nat. *152*, 826–837. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18811430.

S14. Rundle, H. D., Chenoweth, S. F., Doughty, P., and Blows, M. W. (2005). Divergent selection and the evolution of signal traits and mating preferences. PLoS Biol. *3*, e368. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1262626&tool=pmcentrez&rendertype=abstract [Accessed May 28, 2014].

S15. Aitchison, J. (1986). The statistical analysis of compositional data. 12th Editi. (London - New York: Chapman and Hall).

S16. Nosil, P., Gompert, Z., Farkas, T. E., Comeault, A. a, Feder, J. L., Buerkle, C. A., and Parchman, T. L. (2012). SI: Genomic consequences of multiple speciation processes in a stick insect. Proc. R. Soc. B Biol. Sci. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22696527 [Accessed July 21, 2012].

S17. Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J. S., Buerkle, C. A., Feder, J. L., Bast, J., Schwander, T., et al. (2014). Stick insect genomes reveal natural selection's role in parallel speciation. Science (80-.). *344*, 738–742.

S18. Santure, A. W., and Wang, J. (2009). The joint effects of selection and dominance on the QST - FST contrast. Genetics *181*, 259–276. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2621174&tool=pmcentrez&rendertype=abstract [Accessed December 20, 2013].

S19. Nosil, P., Gompert, Z., Farkas, T. E., Comeault, A. A., Feder, J. L., Buerkle, C. A., and Parchman, T. L. (2012). Genomic consequences of multiple speciation processes in a stick insect. Proc. R. Soc. B Biol. Sci. *279*, 5058–5065. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.0813 [Accessed January 29, 2013].

S20. Sandoval, C. P., and Nosil, P. (2005). Counteracting selective regimes and host preference evolution in ecotypes of two species of walking-sticks. Evolution (N. Y). *59*, 2405–2413. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16396181.

S21. Abràmoff, M. D., Magalhães, P. J., and Ram, S. J. (2004). Image Processing with ImageJ. Biophotonics Int. *11*, 36–42.

S22. Nosil, P. (2009). Adaptive population divergence in cryptic color-pattern following a reduction in gene flow. Evolution (N. Y). *63*, 1902–1912. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19245395 [Accessed August 1, 2012].

S23. Comeault, A. A., Soria-Carrasco, V., Gompert, Z., Farkas, T. E., Buerkle, C. A., Parchman, T. L., and Nosil, P. (2014). Genome-wide association mapping of phenotypic traits subject to a range of intensities of natural selection in *Timema cristinae*. Am. Nat. *183*, 711–727. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24739202 [Accessed May 3, 2014].

S24. Guan, Y., and Stephens, M. (2011). Bayesan variable selection regression for genome-wide association studies and other large-scale problems. Ann. Appl. Stat. *5*, 1780–1815.

S25. Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., and Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. Mol. Ecol. *21*, 2991–3005. Available at: http://doi.wiley.com/10.1111/j.1365-294X.2012.05513.x [Accessed March 9, 2012].

S26. Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

S27. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2723002&tool=pmcentrez&rendertype=abstract [Accessed May 23, 2014].

S28. Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with Bayesian sparse linear mixed models. PLoS Genet. *9*, e1003264. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3567190&tool=pmcentrez&rendertype=abstract [Accessed November 7, 2013].

S29. Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. a, and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16862161 [Accessed January 20, 2014].

S30. Aulchenko, Y. S., Ripke, S., Isaacs, A., and van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. Bioinformatics *23*, 1294–1296. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17384015 [Accessed March 24, 2014].

S31. Burke, J. M., Tang, S., Knapp, S. J., and Rieseberg, L. H. (2002). Genetic analysis of sunflower domestication. Genetics *161*, 1257–67. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1462183&tool=pmcentrez&rendertype=abstract.

S32. Miller, C. T., Glazer, A. M., Summers, B. R., Blackman, B. K., Norman, A. R., Shapiro, M. D., Cole, B. L., Peichel, C. L., Schluter, D., and Kingsley, D. M. (2014). Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. Genetics *197*, 405–420. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24652999 [Accessed September 1, 2014].

S33. Barton, N., and Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. Heredity (Edinb). *56*, 357–376.

S34. Gavrilets, S. (2004). Fitness landscapes and the origin of species (Princeton, NJ: Princeton University Press).

S35. Saito, M., and Matsumoto, M. (2006). SIMD-oriented Fast Mersenne Twister: a 128-bit pseudorandom number generator. In Monte Carlo and Quasi-Monte Carlo Methods, A. Keller, S. Heinrich, and H. Niederreiter, eds. (Heidelberg, Berlin: Springer-Berlin), pp. 607–622.

S36. Schwander, T., and Crespi, B. J. (2009). Multiple direct transitions from sexual reproduction to apomictic parthenogenesis in *Timema* stick insects. Evolution (N. Y). *63*, 84–103. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18803687 [Accessed June 13, 2011].

S37. Solano, F. (2014). Melanins: skin pigments and much more — types, structural models, biological functions, and formation routes. New J. Sci. *2014*, 1–28.

S38. Slatkin, M. (2008). Linkage disequilibrium-understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. *9*, 477–485.