nature research

https://doi.org/10.1038/s41559-020-01305-y

Supplementary information

Ecology shapes epistasis in a genotype–phenotype–fitness map for stick insect colour

In the format provided by the authors and unedited

Supplementary Information for:

Ecology shapes epistasis in a genotype-phenotype-fitness map

Patrik Nosil, Romain Villoutreix, Clarissa F. de Carvalho, Jeffrey L. Feder, Thomas L. Parchman, and Zach Gompert

Supplementary Methods

Reflectance data. We focused the core analyses in the main text on two traits (red-green, RG and green-blue, GB) that represent colour variation within the human visible spectrum. This was done under the assumption that these variables accurately represent colour variation in *T. chumash*, and the perception of it by their avian predators. We thus directly tested this assumption, including quantification of ultra-violet reflectance, by recording the spectral reflectance of 10 *T. chumash* specimens with different body colour (including green, dark, and intermediate colourations). These samples were collected on mountain Mahogany (*Cercocarpus*) branches at Glendora Ridge Mile 8.06 (population code GR8.06; latitude 34.22, longitude -117.71) in the spring of 2015. Our results, described below in detail, largely validate our assumption, justifying the use of RG and GB values quantified from photographs.

We collected spectral data using a USB2000 Fibre Optic Spectrometer (Ocean Optics Inc.) equipped with a 400-micron reflection probe (R400-7-SR) and a pulsed Xenon lamp (PX-2) with an output spectrum of 220-750nm. The spectra were measured at a 45^{o} angle with an integration time of 50 milliseconds, the boxcar width adjusted to 5, and averaging across 20 scans. These measurements then were corrected for nonlinearity, stray light, and electric dark using the OceanView software (Ocean Optics). The reflectance was measured relative to a Spectralon >99% white reflectance standard provided by the manufacturer (WS-1). For each individual, we recorded two reflectance spectra: one from the dorsal anterior part of the body (comprising thorax and head) and the second from the dorsal posterior part (abdomen). We interpolated the raw reflectance in the light spectrum between 300-700nm, corrected the negative values to zero and applied triangular smoothing with a distance of 10 nm using the software AVICOL [1]. Finally, we estimated the mean reflectance as that averaged between the dorsal and abdominal measurements in R.

We estimated the mean reflectance across the measured spectrum to obtain an average spectrum curve for *T. chumash.* All specimens presented low reflectance at ultraviolet wavelengths (between 300-400 nm), with average reflectance below 6% (Figure). In the literature, authors tend to consider ultraviolet spectra with less than 10% reflectance not significant [2-5]. Thus, *T. chumash* reflect mainly in the visible spectrum, as do leaves and bark generally [6], rather than in the ultraviolet spectrum.

Luminous stimuli are apprehended by different photoreceptors based on their different sensitivities (termed 'quantum catch'). Here, the avian visual system was used because birds are a major predator of *Timema*, and likely the main predator of *Timema* during late life-history stages [7,8]. Birds have four types of photoreceptors: cones that capture long-wave stimuli (LW, red); mediumwave (MW, green); shortwave (SW, blue); and ultraviolet (UV) [9]. Therefore, colours seen by birds are a result of the interaction of three primary colours plus ultraviolet, stimulating different types of cells in the retina.

Quantum catch for the four photoreceptors was calculated using the file provided by Endler & Mielke [10], with average photoreceptors' sensitivities of birds with ultraviolet-sensitive type of photoreceptor (UVS), given most passerines exhibit this system [11,12]. We used the mean reflectance for each nanometer on the same *T. chumash* specimens analysed above. The analyses were conducted using the *pavo* R package [13]. The results showed MW and LW are the most stimulated cones in general. The UV cone was very marginally stimulated in comparison, suggesting that the UV wavelength range does not contribute strongly to the image perceived by birds.

In summary, our results show that *T. chumash* mostly reflect colours in the visible spectrum, and that this range is largely what is seen by birds, their predators. There is marginal reflectance in the ultraviolet range, but it is low enough to likely be weakly significant for the final colour perceived. Hence, our work focused on analyses of digital photographs is justified and biologically relevant.

Additional genetic results. Additional details of the genetic mapping results are contained below as figures.

Additional experimental results. Some details of the experimental results that were not included in the main text are contained below.

Bayesian lasso and ridge regression. We implemented two alternatives to Bayesian multiple regression with model averaging to estimate additive and epistatic effects of the five colour-associated SNPs on expected fitness (as predicted by the Bayesian phenotypic selection gradient analyses). These alternatives place semi-informative, shrinkage inducing priors on the regression coefficients to improve the robustness and stability of coefficient estimates in models where the number of predictors is large relative to the number of parameters [14-16]. For example, in our case we are estimating five additive SNP effects, 10 two-way epistatic interactions, 10 three-way epistatic interactions, five four-way epistatic interactions, and one five-way epistatic interaction (as well as block-specific intercepts) for each host-plant treatment based on observations of ~200 individuals per treatment ($N = 437$ total).

The two approaches we used were Bayesian lasso regression and Bayesian ridge regression. These methods differ in the prior assumptions they make about the regression coefficients. With Bayesian lasso, a double exponential (Laplace) prior is placed on the regression coefficients, whereas for Bayesian ridge regression a Gaussian prior is used. The Gaussian prior induces more uniform shrinkage of all coefficients towards zero, whereas the Laplace prior shrinks coefficients with smaller parameter estimates (as would be obtained with less informative priors) to zero more strongly than larger parameter estimates (as such lasso is more commonly used for variable selection). In non-Bayesian analysis, a tuning parameter is included in both lasso and ridge regression that must be inferred through cross-validation. However, in Bayesian lasso and Bayesian ridge regression, this parameter (the standard deviation for the Gaussian prior and rate parameter for the Laplace prior) is included in the model and thus uncertainty in the parameter can be integrated over.

We fit the Bayesian lasso and Bayesian ridge regression models using Hamiltonian Monte Carlo (HMC) as implemented in the *rstan* (version 2.19.2) interface with *Stan* via *R* (version 3.5.2) [17-19]. In both analyses, Gaussian priors with mean 0 and standard deviation of 1 were placed on the standard deviation for the (normal) sampling distribution, and Cauchy priors with location 0 and scale 1 were placed on the shrinkage inducing parameter (the standard deviation for the Gaussian or rate for the Laplace across regression coefficients). As with other analyses, the two host plant treatments were analyzed separately. In each case, we used the NUTS algorithm (No U-Turn Sampler) and ran three HMC chains with a burn-in of 500 steps and 2000 sampling steps. The burn-in steps were also used for adaptively tuning the NUTS algorithm.

Bayesian multiple regression and model averaging with dominance effects of SNPs on expected fitness. Results connecting genotype to expected fitness for the five colour-associated SNPs presented in the main text account for genetic interactions between loci (epistasis), but not interactions between alleles at the same locus (dominance). We extended the Bayesian multiple regression models by including dominance terms to assess the contribution of dominance to the genotype-fitness map. We only considered dominance at two of the five loci, SNPs 3 and 4, as these were the only loci with intermediate allele frequencies (minor allele frequency > 0.05) and thus were all three genotypes (both homozygotes and the heterozygote) occurred at appreciable frequencies, which is necessary to parse additive versus dominance effects.

We fit the Bayesian multiple regression/model-averaging models as described in the main text, except that we included dominance effects for SNPs 3 and 4. Thus, these models had five additive effects, 10 two-way epistatic effects, 10 three-way epistatic effects, five four-way epistatic effects, one five-way epistatic effect, two dominance effects, and three block-specific intercepts. Dominance covariates were defined for each SNP as $d_{ij} = 1 - \left| g_{ij} - 1 \right|$ where $g_{ij}$is the genotype (count of non-reference reads) for locus *i* and individual *j*, and $d_{ij}$is the corresponding dominance covariate, which equals 1 if the individual is a heterozygote and 0 otherwise. Note, that we work with non-integer valued estimates of genotypes, and thus this coefficient can take on any value between 0 and 1. We fit the model using the *bms* package in *R* (version 0.3.4; [20]) and with the same priors and MCMC conditions described in the main text.

Overall, we found little evidence for dominance of the colour-associated SNPs with respect to fitness. For example, PIPs for dominance coefficients were low (~0.2 or less) and our estimate for the total number of dominance effects was less than one in both host treatments. Moreover, leave-one-out cross-validation (as described in the main text) indicated little to no improvement in the predictive power of the model if dominance was included. For example, after removing block effects, the predictive power of the models (as measured by the Pearson correlation between observed and predicted expected fitness) with versus without dominance was 0.53 (95% CI = 0.42-0.62) versus 0.56 (95% CI = 0.46-0.65) for the AC treatment and 0.55 (95% CI = 0.44-0.64) versus 0.46 (95% CI = 0.34-0.56) for the MM treatment. Thus, we focus the main text on the effects of epistasis rather than dominance.

Supporting analyses and pleiotropic effects of colour-associated loci. Thus far, our fitness analyses have focused on colour-associated SNPs, and we have implicitly assumed that the main phenotypic effect of these SNPs is on colour, not other traits. Here, we report additional analyses that explore and relax these assumptions, including a test for pleiotropic effects of colour-associated SNPs on other (unmeasured) traits affecting survival. These additional analyses comprise three main approaches, which we applied only to the AC treatment because this is where fitness epistasis was detected and a null model analysing random SNPs was rejected.

First, we conducted additional Bayesian model averaging analyses that no longer factor colour into the response variable. Here, we use survival directly as our response variable, rather than using expected fitness based on phenotypic selection and colour scores. Our independent variables are the same five SNPs used in the main analyses based on expected fitness. We considered models with only main effects of these SNPs, and models that also included interactions between and among SNPs. These represent pure additive models versus models with both additive and epistatic effects, respectively. The results show substantial predictive power when considering only genotype and survival, including a marked increase in predictive power when epistasis is considered (Table 1).

Second, we have conducted a truly 'phenotype-free' analysis of fitness epistasis in MAPIT. Here, we quantify the epistatic effect of every SNP in the *Mel-Stripe* region on survival (response variable), rather than only considering the most strongly colour-associated SNPs. This revealed that the five specific SNPs exhibiting epistasis for colour show collective evidence of epistasis for survival as well (individual *P*-values are 0.11. 0.18, 0.15. 0.09 0.10, Fisher's combined probability test, $X^2 = 21.05$, d.f. = 10, $P = 0.007$). A similar trend, albeit weaker, was detected for the five SNPs with additive effects on colour that were used in our Bayesian model-averaging survival analyses (individual *P*-values are 0.25, 0.52. 0.11. 0.15, 0.52, Fisher's combined probability test, $X^2 = 13.47$, d.f. = 10, $P = 0.051$). Our finding of fitness epistasis thus appears robust to different analytical approaches, but stronger results are obtained when considering selection acting through colour phenotype (i.e., expected fitness from the phenotypic selection analysis as the response variable) than when considering genotype alone.

Third, we have employed an approach for testing pleiotropy. The idea is to include both genotype (in our case, the five colour-associated SNPs and interactions among them) and measured phenotype (in our case, RG and GB colour scores) as independent covariates in a model that predicts survival (response variable).

If effects of genotype on survival are observed even when controlling for phenotype, this implies pleiotropy, i.e., effects of genotype on survival independent from colour, via unmeasured traits. Our results reveal a modest (~5%) increase in cross-validation predictive power in the model with both covariates, relative to a model including only colour (Table 1). Thus, the results maintain clear evidence for selection on colour phenotype, but also suggest some pleiotropic effects of colour loci on other (unmeasured) traits influencing fitness.

Supplemental Table 1. Details of the experimental bushes. Note that a total of 76 individuals were released onto each experimental plant.

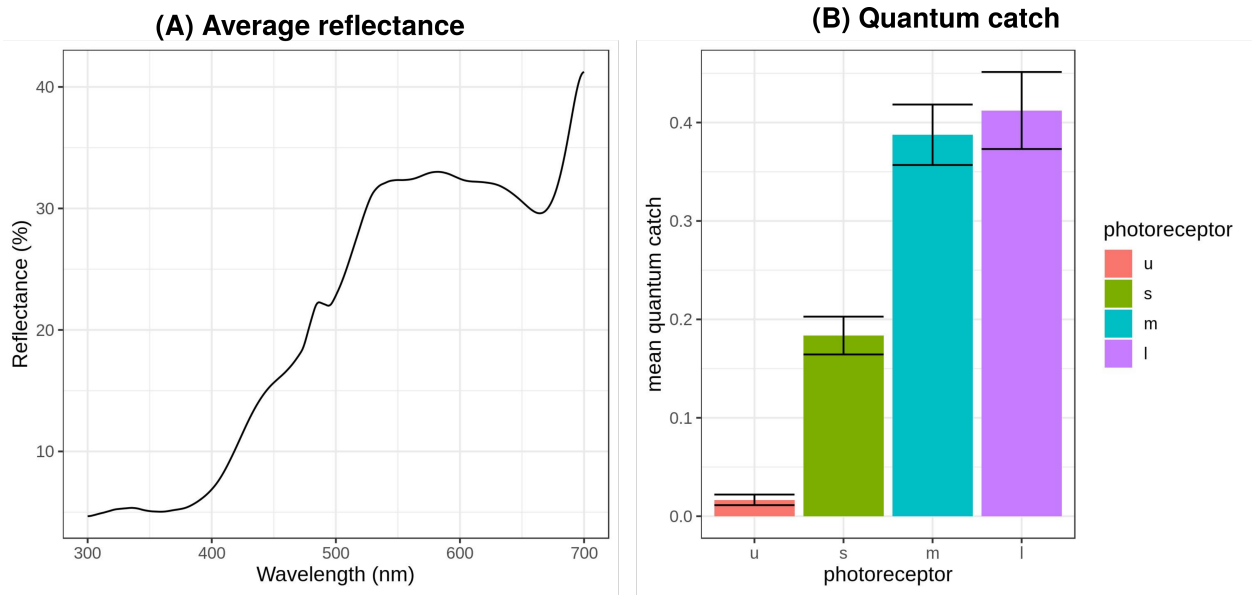
Host and treatment	Latitude	Longitude	Elevation	Number recaptured
AC1	34 15.626	-118 6.220	4731 feet	23
MM1	34 15.626	-118 6.256	4728 feet	12
AC2	34 15.688	-118 6.058	4795 feet	22
MM2	34 15.683	-118 6.124	4776 feet	14
MM3	34 15.742	-118 6.042	4811 feet	9
AC3	34 15.770	-118 6.014	4825 feet	18

Supplemental Table 2. Approximate phenotypic selection gradients. Posterior means and 95% credible intervals are presented.

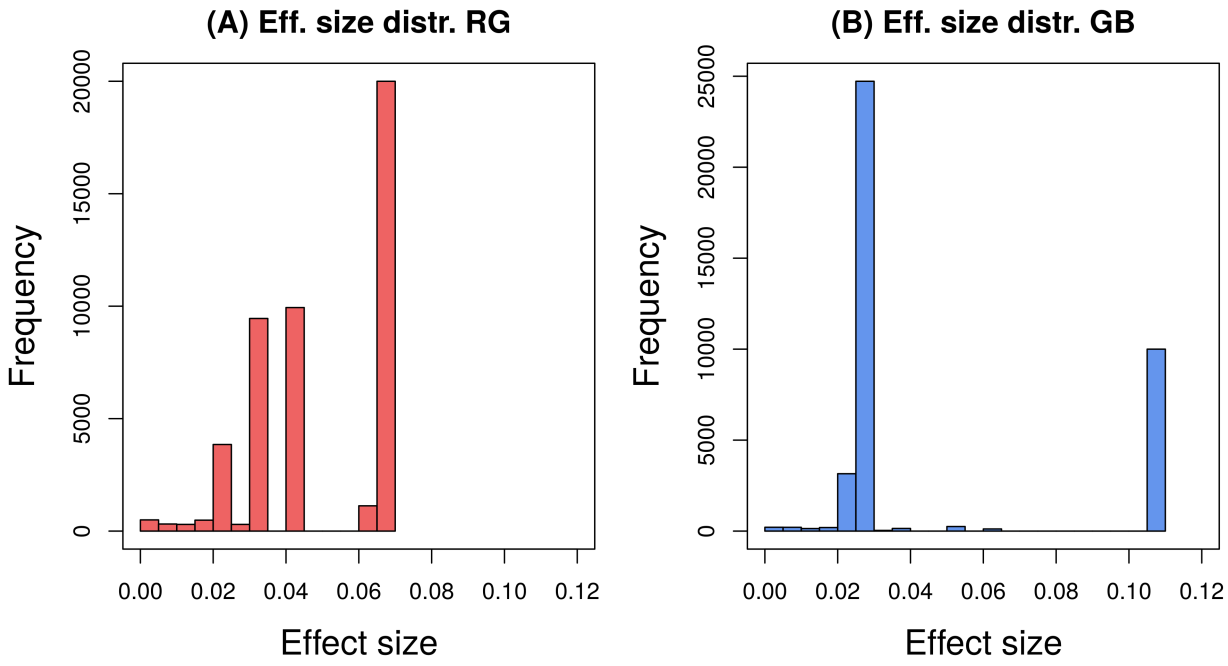
Form	**Treatment**	**Block 1**	**Block 2**	**Block 3**
Linear RG	AC	-0.124 (-0.241,-0.002)	-0.145 (-0.28, -0.025)	-0.104 (-0.224, -0.019)
Linear GB		0.079 (-0.008, 0.195)	0.018 (-0.096, 0.119)	0.057 (-0.024, 0.280)
Quadratic RG		0.120 (-0.030, 0.272)	0.178 (0.026, 0.370)	0.03 (-0.353, 0.175)
Quadratic GB		0.012 (-0.154, 0.181)	0.024 (-0.141, 0.199)	-0.105 (-0.491, 0.061)
Correlational		-0.144 (-0.295, -0.021)	-0.144 (-0.303, -0.019)	-0.052 (-0.174, 0.134)
Linear RG	MM	0.038 (-0.055, 0.167)	0.024 (-0.068, 0.139)	-0.028 (-0.12, 0.053)
Linear GB		-0.003 (-0.069, 0.060)	0.006 (-0.057, 0.085)	-0.003 (-0.063, 0.057)
Quadratic RG		0.070 (-0.065, 0.188)	0.088 (-0.054, 0.221)	0.087 (-0.016, 0.202)
Quadratic GB		0.060 (-0.054, 0.183)	0.072 (-0.046, 0.195)	0.042 (-0.074, 0.143)
Correlational		0.040 (-0.058, 0.128)	0.045 (-0.040, 0.137)	0.059 (-0.016, 0.150)

Supplemental Table 3. Summary information for the five colour-associated SNPs, including their linkage group (LG), scaffold, position (in base pairs, bp), the reference and alternative alleles (Ref/Alt, respectively), the RG and GB posterior inclusion probabilities (RG PIP and GB PIP), the RG and GB model-averaged effect estimates, and the odds ratio. The latter , that is the relative likelihood of an effect compared to no effect. Scaff. = Scaffold.

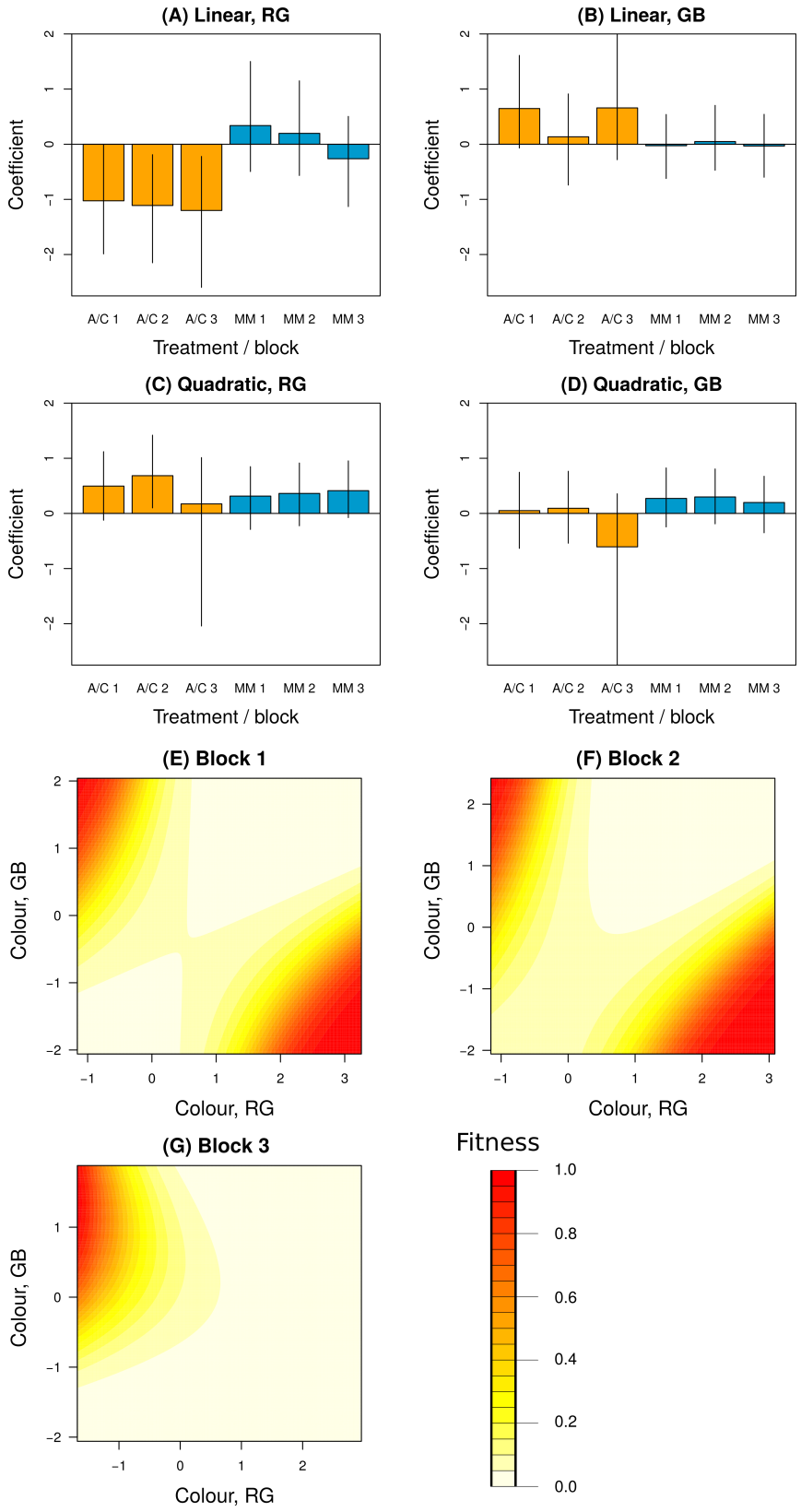
LG	Scaff.	Position	Ref/ Alt	MAF	RG PIP	GB PIP	RG effect	GB effect	RG odds ratio	GB odds ratio
8	128	5359458	T/C	0.128	<0.001	0.712	1.02E-05	2.14E-02	<1	478
8	128	5359483	C/T	0.073	0.945	0.014	-2.98E-02	4.24E-04	635	9
8	128	5359496	G/A	0.401	1	0.878	-6.84E-02	2.6E-02	672	589
8	128	5690741	C/T	0.335	0.153	0.73	-3.2E-03	1.89E-02	103	490
8	128	5708354	T/C	0.09	0.994	<0.001	4.03E-02	-7.16E-05	668	<1



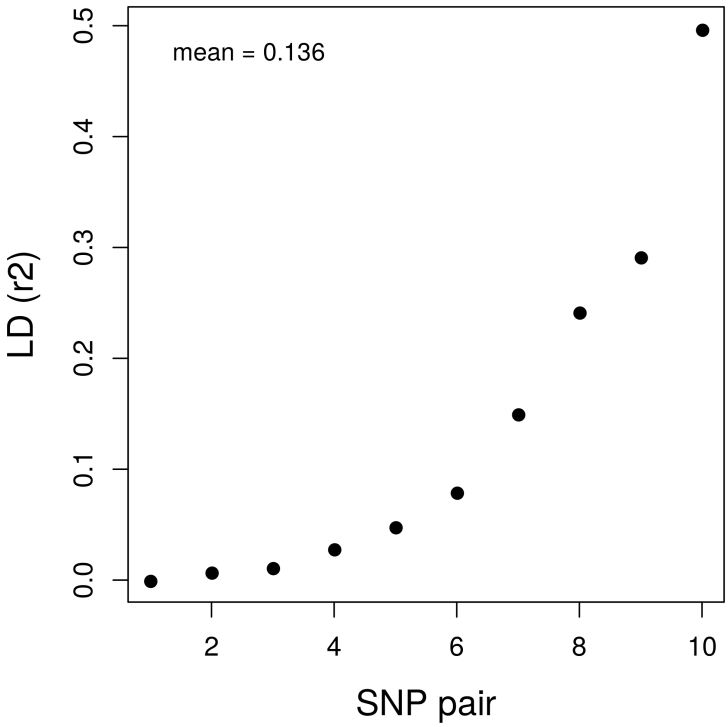
Supplementary Figure 1. (A) Average reflectance curve in *T. chumash.* (B) Quantum catch of colours at each photoreceptor of average avian ultraviolet-sensitivity system (UVS). Reflectance for all *T. chumash* specimens were averaged at each nanometer. u – ultraviolet photoreceptor, s – small wavelength photoreceptor (blue), m – medium wavelength photoreceptor (green), and l – long wavelength photoreceptor (red).



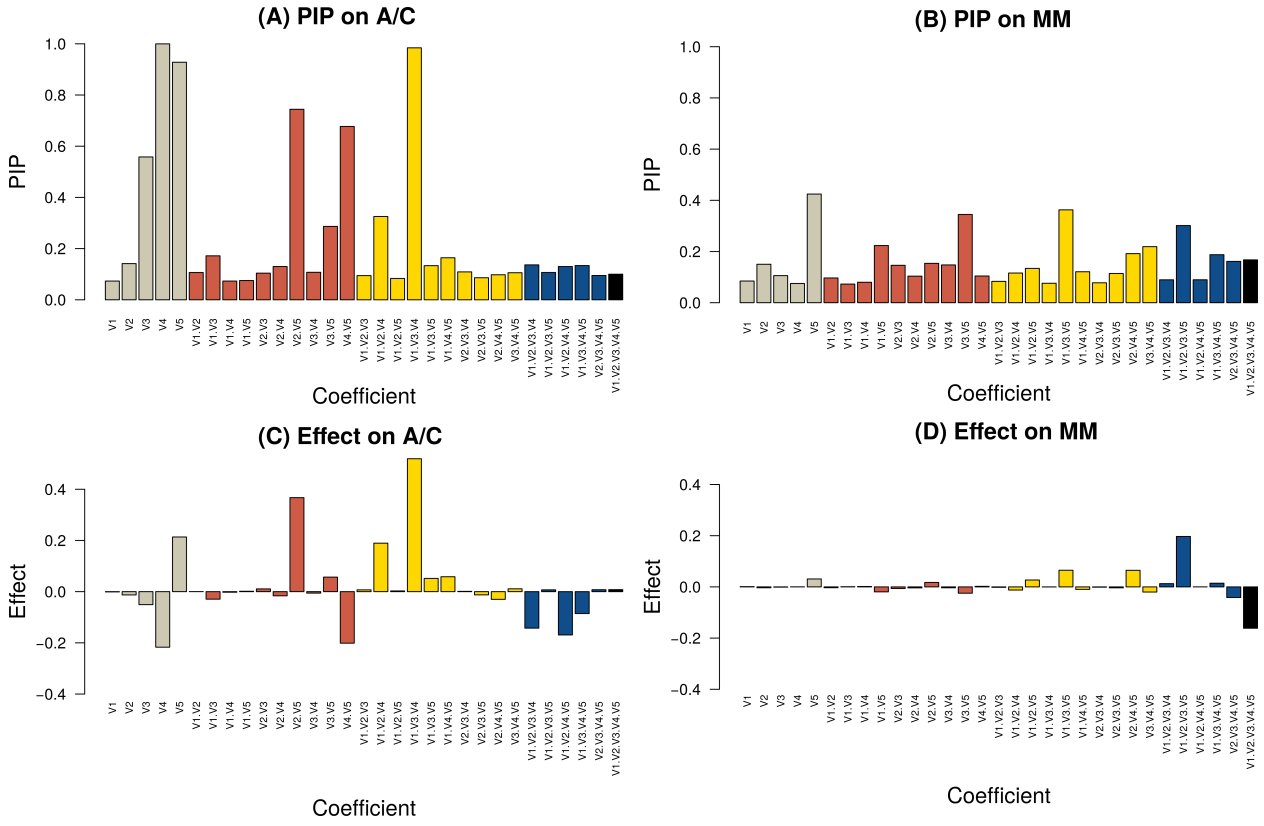
Supplementary Figure 2. Histograms give the distribution of effect (Eff.) sizes of single nucleotide polymorphisms (SNPs) on (A) red-green (RG) and (B) green-blue (GB) colour variables. This was computed over all 157 SNPs within the *Mel-Stripe* locus and was taken as an expectation over the posterior distribution (by sampling SNPs and their associated conditional effect sizes according to their posterior inclusion probabilities). 10,000 samples from the posterior were used to generate these plots.



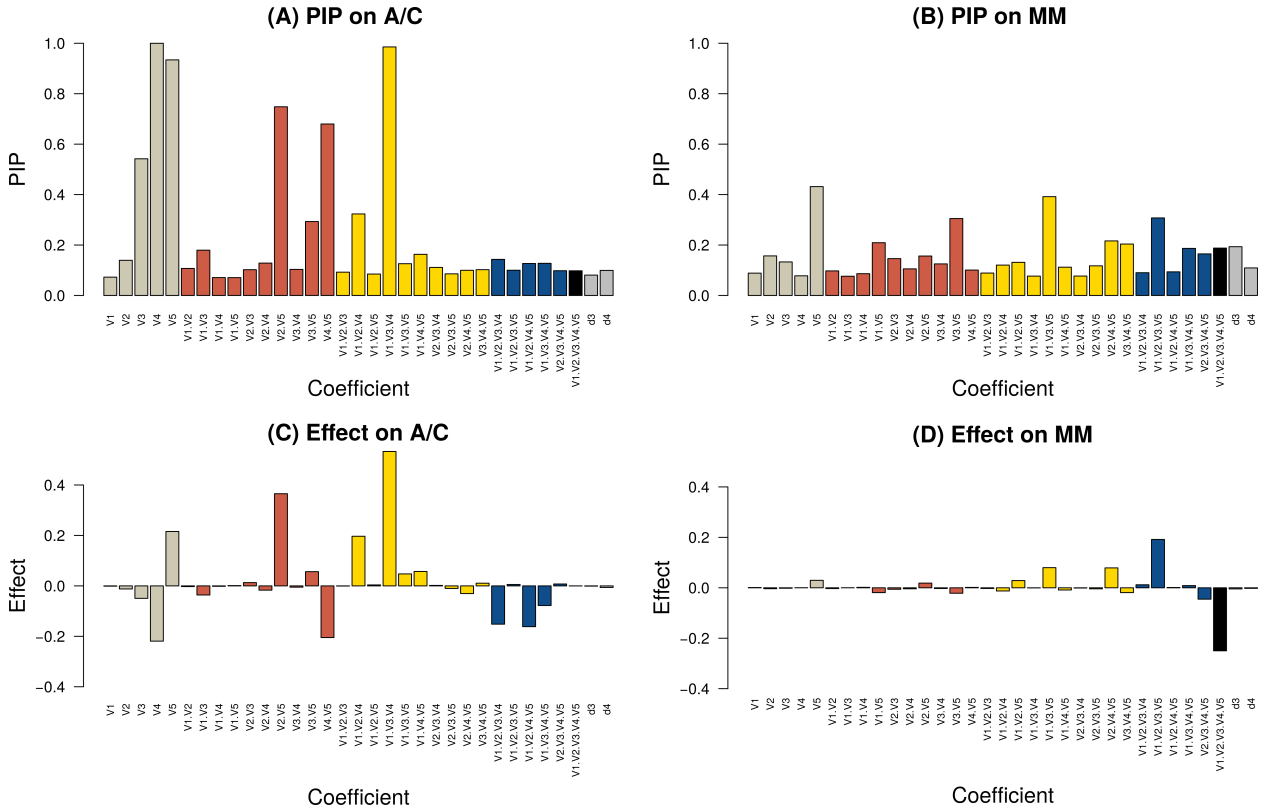
Supplementary Figure 3. Plots show Bayesian estimates of linear (A, B) and quadratic selection (C, D) on red-green (RG) and green-blue (GB) colour variables. Bars denote posterior medians and vertical lines delimit the 95% equal-tail probability intervals. Negative and positive values of quadratic selection correspond with disruptive and stabilizing selection, respectively. Estimates of correlational selection are provided in the main text. A/C is the *Adenostoma* and *Ceanothus* treatment and MM is Mountain Mohagany (*Cercocarpus*). (E-G) Heat maps of survival probabilities from the selection gradient analysis in the A/C treatment. Results are shown for each block. Darker colours reflect higher survival probability.



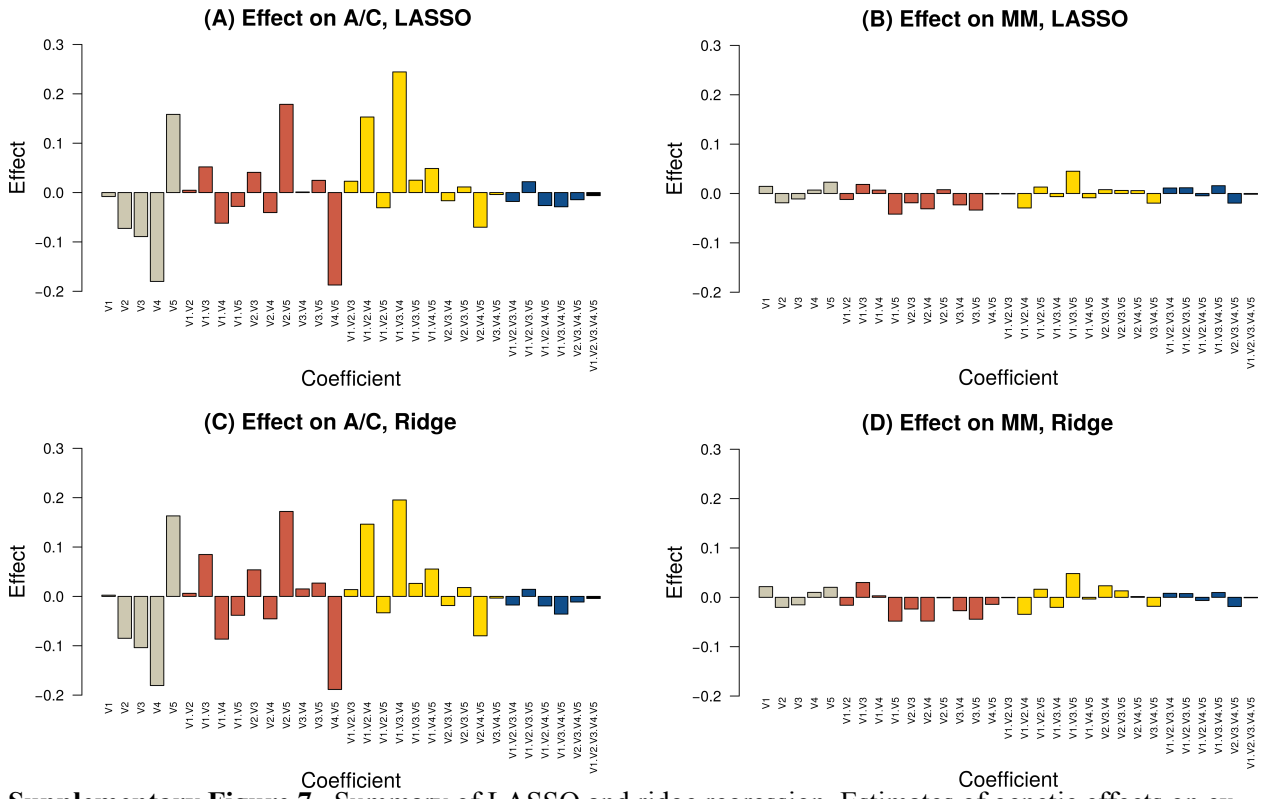
Supplementary Figure 4. Plots shows estimates of pairwise linkage disequilibrium (LD) for the five colour-associated single nucleotide polymorphisms, SNPs (10 pairs total). LD was measured by the coefficient of determination (r^2) between genotypes for pairs of SNPs. SNP pairs are sorted by LD (smallest to largest) and the mean LD across the 10 pairs is all given.



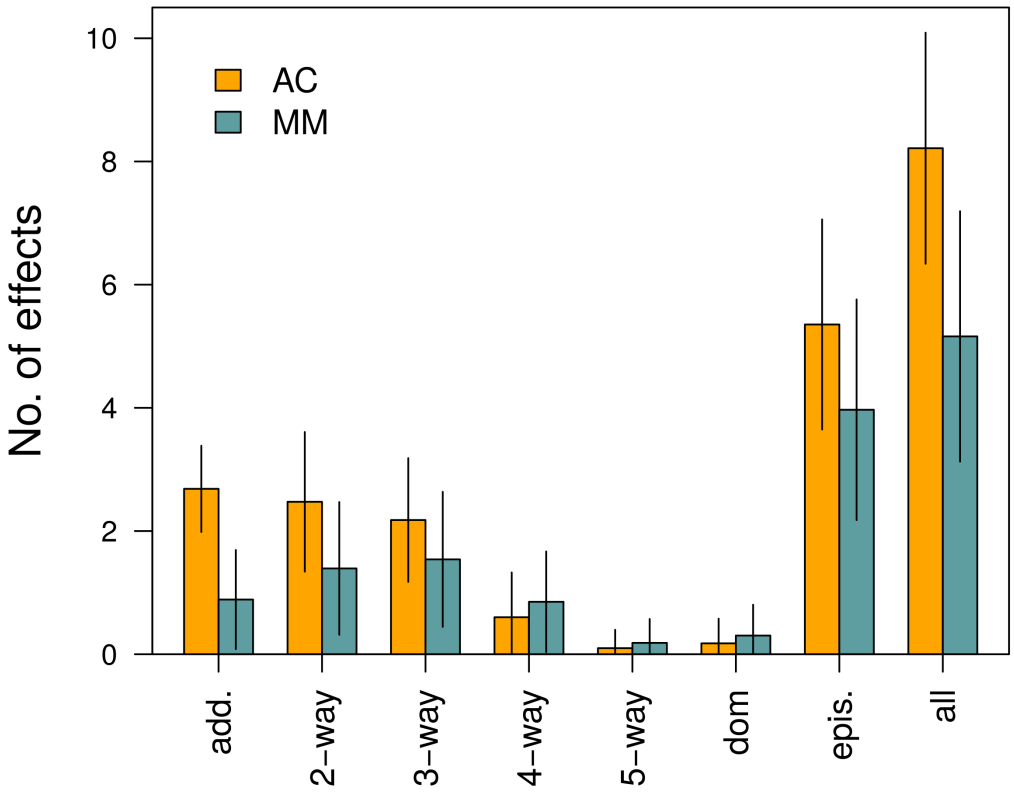
Supplementary Figure 5. Summary of Bayesian multiple regression results. Covariate posterior inclusion probability (PIPs) from Bayesian multiple regression of expected fitness on colour-associated single nucleotide polymorphisms (SNPs) in the A/C (A) and MM (B) host plant treatments. A/C is the *Adenostoma* and *Ceanothus* treatment and MM is Mountain Mohagany (*Cercocarpus*). Shown are the main effects and two-way, three-way, four-way and five-way interactions between/among SNPs, which represent additive versus epistatic effects, respectively. Model-averaged effects from Bayesian multiple regression of expected fitness on colour-associated SNPs in the AC (C) and MM (D) host plant treatments. Shown are the main effects and two-way, three-way, four-way and five-way interactions between/among SNPs, which represent additive versus epistatic effects, respectively. SNPs, SNP pairs, SNP triplets, etc. are labelled on the x-axis.



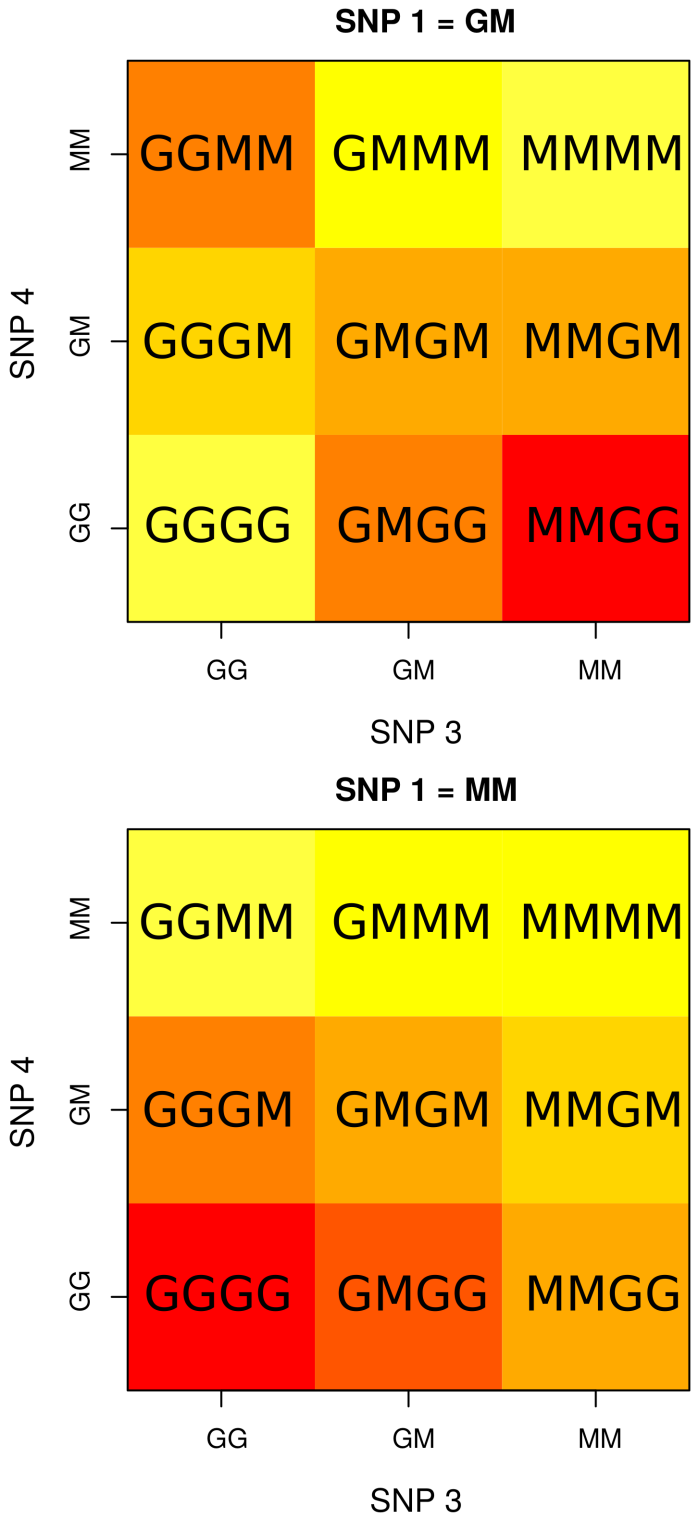
Supplementary Figure 6. Summary of Bayesian multiple regression results that include consideration of potential effects of genetic dominance. Covariate posterior inclusion probability (PIPs) from Bayesian regression with dominance in the AC (A) and MM (B) host plant treatments. A/C is the *Adenostoma* and *Ceanothus* treatment and MM is Mountain Mohagany (*Cercocarpus*). Shown are the main effects, interactions between single nucleotide polymorphisms (SNPs), and interactions between alleles within loci, which represent additive, epistatic, and dominance effects, respectively. Dominance effects are shown in gray (denoted 'd' in the x-axis labels) and are given for only SNPs 3 and 4. Analogous results without dominance are described in the main text. Model-averaged effects from Bayesian regression with dominance in the AC (C) and MM (D) host plant treatments. Shown are main effects, interactions between SNPs, and interactions between alleles within loci, which represent additive, epistatic, and dominance effects, respectively. Dominance effects are shown in gray (denoted 'd' in the x-axis labels) and are given for only SNPs 3 and 4. SNPs, SNP pairs, SNP triplets, etc. are labelled on the x-axis. Analogous results without dominance are described in the main text.



Supplementary Figure 7. Summary of LASSO and ridge regression. Estimates of genetic effects on expected fitness from Bayesian LASSO for AC (A) and MM (B) host plant treatments. A/C is the *Adenostoma* and *Ceanothus* treatment and MM is Mountain Mohagany (*Cercocarpus*). Shown are main effects and interactions between single nucleotide polymorphisms (SNPs), which represent additive versus epistatic effects, respectively. Estimates of genetic effects on expected fitness from Bayesian ridge regression for AC (C) and MM (D) host-plant treatments. Shown are main effects and interactions between SNPs, which represent additive versus epistatic effects, respectively. SNPs, SNP pairs, SNP triplets, etc. are labelled on the x-axis.

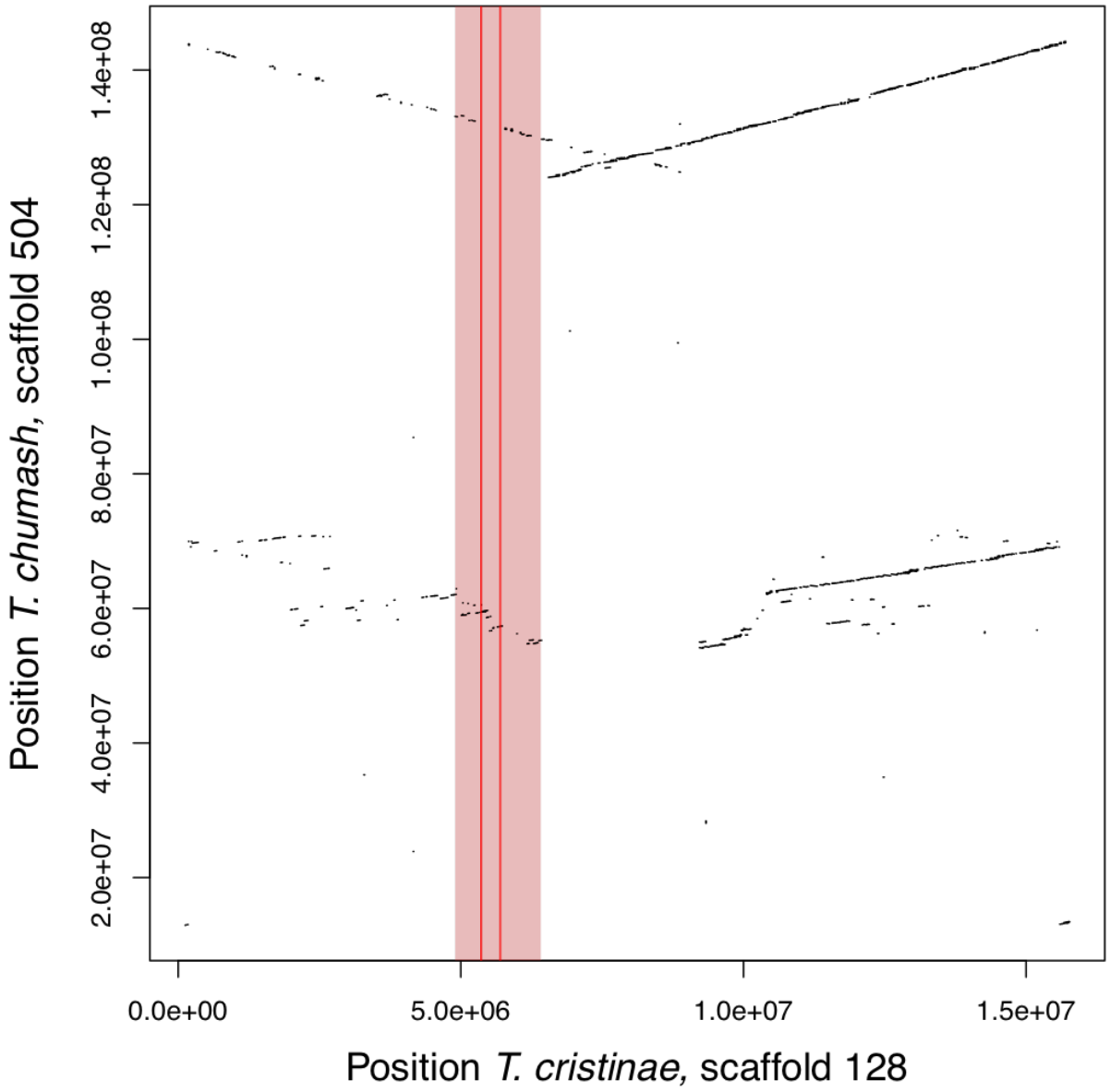


Supplementary Figure 8. The number of estimated effects in the Bayesian multiple regression with dominance. Error bars represent posterior standard deviations (analogous to standard errors). Analogous results without dominance are described in the main text. No. = number, add. = additive, epis. = epistatic (all epistatic effects). Analogous results without dominance are given in the main text.



Supplementary Figure 9. An example of epistasis among three loci (i.e., single nucleotide polymorphisms, SNPs). Alleles at each locus have been coded in terms of whether they increase green or melanistic colouration (increases green, G; increases melanism, M). Darker colours represent higher survival probability. Only the GM and MM genotypes are shown for SNP 1, as the GG homozygote was not ob-

served in our data (i.e., the G allele is rare). Analagous results for a simpler pairwise interaction are given in the main text.



Supplementary Figure 10. Dot plot showing the alignment between scaffold 128 from the *T. cristinae* melanic genome and scaffold 504 from the *T. chumash* genome. The latter contains nearly all alignments with scaffold 128 from *T. cristinae* (including all alignments spanning the five focal SNPs). The pink shaded region denotes the indel identified by past work. Vertical red lines denote the five focal colour-associated SNPs (these cluster in two groups of three and two SNPs; not all five lines are visible due to the scale depicted). The alignment suggests these five loci exist in a single copy and are near each other in the *T. chumash* genome (although other parts of scaffold 128 appear to be duplicated). The crossing dots to the right of the indel on the x-axis support the existence of a chromosomal inversion.

Supplementary References

1 AVICOL, a program to analyse spectrometric data. Last update October 2011. Free executable available. (2006).

2 Endler, J. A. THE COLOR OF LIGHT IN FORESTS AND ITS IMPLICATIONS. Ecological Monographs 63, 1-27, doi:10.2307/2937121 (1993).

3 Cuthill, I. C., Bennett, A. T. D., Partridge, J. C. & Maier, E. J. Plumage reflectance and the objective assessment of avian sexual dichromatism. Am. Nat. 153, 183-200, doi:10.1086/303160 (1999).

4 Finger, E. & Burkhardt, D. BIOLOGICAL ASPECTS OF BIRD COLORATION AND AVIAN COLOR-VISION INCLUDING ULTRAVIOLET RANGE. Vision Research 34, 1509-1514, doi:10.1016/0042-6989(94)90152-x (1994).

5 Schmidt, V., Schaefer, H. M. & Winkler, H. Conspicuousness, not colour as foraging cue in plant-animal signalling. Oikos 106, 551-557, doi:10.1111/j.0030-1299.2004.12769.x (2004).

6 Endler, J. A. Some general comments on the evolution and design of animal communication systems. Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci. 340, 215-225 (1993).

7 Nosil, P. & Crespi, B. J. Experimental evidence that predation promotes divergence in adaptive radiation. Proc. Natl. Acad. Sci. U. S. A. 103, 9090-9095, doi:10.1073/pnas.0601575103 (2006).

8 Sandoval, C. P. Differential Visual Predation on Morphs of Timema-Cristinae (Phasmatodeae, Timemidae) and Its Consequences for Host-Range. Biol. J. Linnean Soc. 52, 341-356 (1994).

9 Bennett, A. T. D. & Cuthill, I. C. Ultraviolet Vision in Birds: What Is Its Function? . Vision Research 34, 1471–1478. (1994).

10 Endler, J. A. & Mielke, P. W. Comparing entire colour patterns as birds see them. Biol. J. Linnean Soc. 86, 405-431, doi:10.1111/j.1095-8312.2005.00540.x (2005).

11 Odeen, A. & Hastad, O. Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. Mol. Biol. Evol. 20, 855-861, doi:10.1093/molbev/msg108 (2003).

12 Odeen, A., Hastad, O. & Alstrom, P. Evolution of ultraviolet vision in the largest avian radiation - the passerines. BMC Evol. Biol. 11, doi:10.1186/1471-2148-11-313 (2011).

13 Maia, R., Eliason, C. M., Bitton, P. P., Doucet, S. M. & Shawkey, M. D. pavo: an R package for the analysis, visualization and organization of spectral data. Methods in Ecology and Evolution 4, 906-913, doi:10.1111/2041-210x.12069 (2013).

14 Goldstein, M. BAYESIAN-ANALYSIS OF REGRESSION PROBLEMS. Biometrika 63, 51-58, doi:10.1093/biomet/63.1.51 (1976).

15 Park, T. & Casella, G. The Bayesian Lasso. Journal of the American Statistical Association 103, 681-686, doi:10.1198/016214508000000337 (2008).

16 Hans, C. Bayesian lasso regression. Biometrika 96, 835-845, doi:10.1093/biomet/asp047 (2009).

17 Hoffman, M. D. & Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15, 1593-1623 (2014).

18 Carpenter, B. et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software 76, 1-29, doi:10.18637/jss.v076.i01 (2017).

19 RStan: the R interface to Stan. R package version 2.19.2. http://mc-stan.org/. (2019).

20 Zeugner, S. & Feldkircher, M. Bayesian Model Averaging Employing Fixed and Flexible Priors: The BMS Package for R. Journal of Statistical Software 68, 1-37 (2015).