Additive genetic effects in interacting species jointly determine the outcome of caterpillar herbivory

Zachariah Gompert[1,a,b], **Tara Saley**[a,b], **Casey Philbin**[c], **Su'ad A. Yoon**[d], **Eva Perry**[d], **Michelle E. Sneck**[e], **Joshua G. Harrison**[f], **C. Alex Buerkle**[f], **James A. Fordyce**[g], **Chris C. Nice**[h], **Craig Dodson**[c], **Sarah L. Lebeis**[i,j], **Lauren K. Lucas**[a], and **Matthew L. Forister**[d]

[a]Department of Biology, Utah State University, Logan, UT 84322, USA; [b]Ecology Center, Utah State University, Logan, UT 84322, USA; [c]Hitchcock Center for Chemical Ecology, University of Nevada, Reno, NV 89557, USA; [d]Department of Biology, University of Nevada, Reno, NV 89557, USA; [e]Department of BioSciences, Rice University, Houston, TX 77005, USA; [f]Department of Botany, University of Wyoming, Laramie, WY 82071, USA; [g]Department of Ecology & Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA; [h]Department of Biology, Texas State University, San Marcos, TX 78666, USA; [i]Department of Microbiology, University of Tennessee, Knoxville, TN, USA; [j]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA

This manuscript was compiled on December 19, 2022

Plant-insect interactions are common and important in basic and applied biology. Trait and genetic variation can affect the outcome and evolution of these interactions, but the relative contributions of plant and insect genetic variation and how these interact remain unclear and are rarely subject to assessment in the same experimental context. Here we address this knowledge gap using a recent host range expansion onto alfalfa by the Melissa blue butterfly. Common garden rearing experiments and genomic data show that caterpillar performance depends on plant and insect genetic variation, with insect genetics contributing to performance earlier in development and plant genetics later. Our models of performance based on caterpillar genetics retained predictive power when applied to a second common garden. Much of the plant genetic effect could be explained by heritable variation in plant phytochemicals, especially saponins, peptides, and phosphatidyl cholines, providing a possible mechanistic understanding of variation in the species interaction. We find evidence of polygenic, mostly additive effects within and between species, with consistent effects of plant genotype on growth and development across multiple butterfly species. Our results inform theories of plant-insect coevolution and the evolution of diet breadth in herbivorous insects and other host-specific parasites.

Plant-insect interaction | Genomic prediction | Polygenic | Phytochemicals | Coevolution

A central challenge for the biological sciences is to understand the causes and consequences of trait variation within and among species. Experimental manipulations aimed at understanding the molecular basis of organismal variation have most often been done in settings stripped of all or most ecological context. This approach can be fruitful for simple traits, including some aspects of morphology (e.g., (1–4)), but is lacking when it comes to interspecific interactions that include the evolution of crop pests, emerging infectious diseases, and other host-parasite associations (5, 6).

Plants and herbivorous insects have contributed much to our understanding of the formation and persistence of interactions between hosts and parasites, in part because they are experimentally tractable but also because insects are the most diverse macroscopic organisms on the planet and their specialized feeding habits play a role in their diversification (7–11). Yet classic studies of the molecular basis of plant-insect interactions have relied on candidate genes or targeted classes of phytochemical compounds (e.g., (12–14)). More recently, evolutionary geneticists have taken advantage of new technologies to explore the genetic basis of herbivory in a genomic context. With very few exceptions, these studies have focused on genetic variation in either herbivores or plants (15–19)(but see (20)), but rarely both in the same study and never to our knowledge paired with modern metabolomic approaches that allow for untargeted discovery of influential compounds (21). This leaves us with considerable uncertainty concerning the relative importance of heritable traits in herbivores and in plants for determining the outcome of plant-insect interactions. For example, particular genetic variants in an herbivore might be associated with increased feeding efficiency, but only when challenged with particular plant variants such as specific defensive metabolites or combinations of physical defenses (22). However, without an understanding of the genetic architecture of both the herbivore physiology and the plant traits, the evolutionary trajectory of the system cannot be understood in the context of available theoretical models or forecast with respect to the evolution of defense in the plant or increased performance in the herbivore. We address this need using a recent host range expansion onto alfalfa by the Melissa blue

Significance Statement

Studies of ecological interactions often ignore genetic variation and studies of coevolution have rarely assayed the genetics of hosts and parasites at the same time. We show that genetic differences among Melissa blue caterpillars and alfalfa plants account for 17-49% of the variability in caterpillar growth and survival. The genetic contribution includes heritable variation in defensive compounds including saponins. Our results suggest that the outcome of this plant-insect interaction is affected by many genes with mostly independent (additive) effects. Moreover, genetic differences among alfalfa plants have consistent effects on caterpillar growth in multiple butterfly populations and species. Our results thus advance understanding of the evolution of ecological interactions, including host-parasite interactions beyond herbivorous insects and plants.

ZG, CAB, JAF, CCN, SLL, LKL, and MLF designed the study. ZG, TS, CP, SAY, EP, MES, JGH, CAB, LKL and MLF conducted the experiments and generated the data. ZG, TS, CP, and MLF analyzed the data. ZG, CP and MLF drafted the manuscript. All authors helped edit and revise the manuscript.

The authors declare no competing interests.

---

[1] To whom correspondence should be addressed. E-mail: zach.gompert@usu.edu

Summary of hypotheses and effects

Hypothesis	Plant genetics	Insect genetics	Combination
(i)	No	Yes	N/A
(ii)	Yes	No	N/A
(iii)	Yes	Yes	Additive
(iv)	Yes	Yes	Epistatic
Null	No	No	N/A

Fig. 1. Main hypotheses tested about the contribution of plant and insect genetics to caterpillar performance: (i) caterpillar performance is primarily affected by insect (*L. melissa*) genetics, (ii) caterpillar performance is primarily affected by plant (*M. sativa*) genetics, (iii) the genetics of the interacting species have similar effects on caterpillar performance and combine additively, (iv) the genetics of the interacting species have similar effects on caterpillar performance and combine epistatically, and (v) the null hypothesis that neither insect or plant genetic variation have an appreciable effect on caterpillar performance. The illustration (by R. Ribas) shows a *L. melissa* caterpillar feeding on alfalfa, while being tended by ants; additional biotic or abiotic factors, such as the presence of mutualistic ants, also affect caterpillar performance in the wild (25) but are not a component of this study.

butterfly, emphasizing the role of prediction when building an understanding of the functional genetic basis of a novel plant-insect interaction.

The Melissa blue butterfly (*Lycaeides melissa*) is widespread in western North America (23). It exists in isolated populations associated with larval host plants in the legume family, including many species of *Astragalus* and *Lupinus* (24, 25). The Melissa blue colonized alfalfa (*Medicago sativa*) after the plant was introduced to the western USA as a forage crop in the mid 1800s, and is now commonly found on naturalized (i.e., feral) alfalfa along roadsides and trails (24). Melissa blue butterflies show evidence of adaptation to alfalfa, but this host plant remains inferior to known native hosts in terms of caterpillar development with cascading life history effects (26–28). As insect growth and survival is often reduced on novel hosts, the lower quality of alfalfa for Melissa blue butterflies is likely typical of a general phenomenon (29). Alfalfa is phenotypically variable (30), and thus is not a homogeneous resource for Melissa blue butterflies. In particular, phenotypic variation among naturalized alfalfa populations, including phytochemical variation, affects Melissa blue caterpillar growth and host patch occupancy (25, 31, 32). However, it is unclear how much of this phenotypic variation has a genetic basis. Moreover, as is true for other plant-insect interactions, the relative contributions of plant (alfalfa) and insect (Melissa blue) genetic variation to the outcome of the interaction is unexplored, including whether growth and successful development from caterpillar to adult is influenced by additive or epistatic genetic variation in the interacting species.

Here, we use multiple common garden rearing experiments combined with multilocus genetic mapping and genomic prediction to build and test models that quantify the relative effects and interactions of alfalfa and Melissa blue genetic variation on caterpillar performance (i.e., growth and survival). We specifically test the following alternative hypotheses: (i) caterpillar performance is primarily affected by Melissa blue genetic variation and architecture, (ii) caterpillar performance is primarily affected by genetic variation and architecture in the host plant, (iii) the genetics of the interacting species have similar effects on caterpillar performance and combine additively, (iv) the genetics of the interacting species have similar effects on caterpillar performance and combine epistatically,

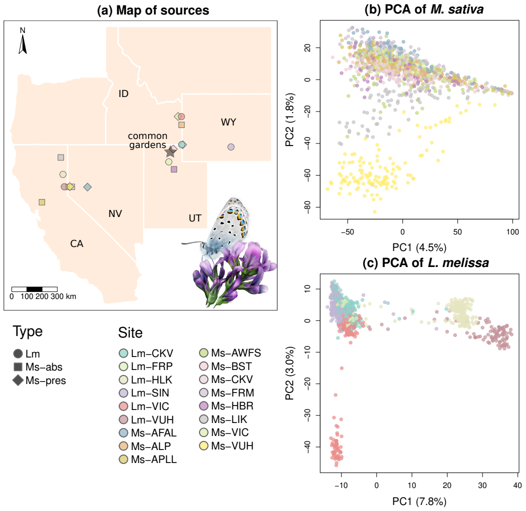


Fig. 2. (a) Map of plant (*M. sativa*) and insect (*L. melissa*) common garden source populations. Symbol shapes denote source type–Lm = *L. melissa*, Ms-abs = *M. sativa* site without *L. melissa* butterflies, and Ms-pres = *M. sativa* site with *L. melissa* butterflies, and are colored to indicate different populations within taxa. The inset illustration shows an adult *L. melissa* perched on *M. sativa* (illustration by R. Ribas). (b) Ordination of genetic variation via principal component analysis (PCA) for the *M. sativa* common garden plants. (c) Ordination of genetic variation via PCA for the *L. melissa* caterpillars from the rearing experiment. Points in (b) and (c) denote individual plants or caterpillars and are colored to match the map (a).

and (v) the null hypothesis that neither Melissa blue nor alfalfa genetic variation has an appreciable effect on caterpillar performance (Fig. 1). Genetic mapping of 1760 plant traits, including 1750 phytochemical metabolites, contributes to testing these hypotheses and also allows us to probe the functional basis of plant-genetic effects on caterpillar performance. Finally, we conduct complementary rearing experiments to test the consistency of plant genetic effects (i.e., their lack of interaction with herbivore genetics) across butterfly populations and species.

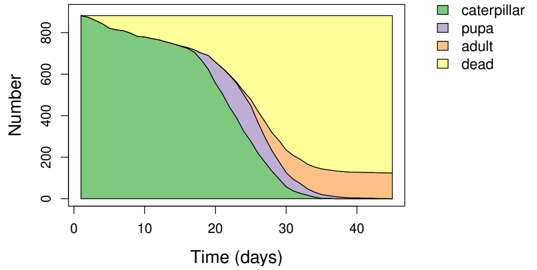
Results

Overview of the primary common garden rearing experiment.

We planted a common garden comprising 1080 alfalfa (*M. sativa*) plants at the Greenville Experimental Farm near Logan, UT (41.765° N, 111.814° W) in 2018 (Fig. S1a). Seeds for this garden were collected from 11 naturalized (i.e., feral) *M. sativa* sites in the western USA, including five sites where *L. melissa* butterflies are found (Table S1, Fig. 2a). Caterpillars for the experiment were sourced from six sites by obtaining eggs from gravid *L. melissa* females in 2019. We detected substantial genetic variation and only subtle genetic differentiation among the source locations for alfalfa (161,008 SNPs, mean expected heterozygosity = 0.168, F_{ST} = 0.029) and for *L. melissa* (63,194 SNPs, mean expected heterozygosity = 0.065, F_{ST} = 0.045) (Figs. 2b, S2). Nearby SNPs (<100 bps) exhibited appreciable linkage disequilibrium in *M. sativa* (median r^2 = 0.050, 95th percentile = 0.862) and *L. melissa* (median r^2 = 0.002, 95th percentile = 0.052), but this decayed rapidly with physical distance with especially low levels of LD beyond 100 bps in *L. melissa* (Fig. S3).

The main rearing experiment was conducted in summer 2019. For this experiment, caterpillars were reared individually on each of the 1080 alfalfa plants. Rearing was done in a growth

(a) Survival and development



(b) Weight and survival correlations

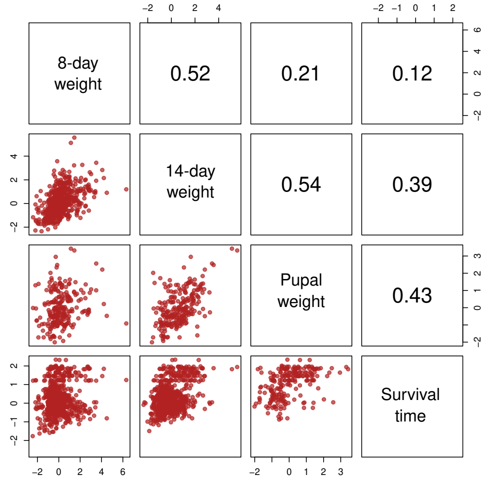


Fig. 3. (a) Plot shows survival and development of *L. melissa* over the course of the rearing experiment. Colored regions denote the number of individuals that were living caterpillar, pupa, adults or dead at each day post hatching. (b) Plots show pairwise correlations between *L. melissa* performance traits. Scatterplots are shown in the lower-triangle panels–each point denotes one individual–and Pearson correlations are reported in the corresponding upper triangle panels. Traits are given along the diagonal panels: 8-day weight, 14-day weight, pupal weight, and truncated survival time. Scatterplots and Pearson correlations are based on residuals after controlling for confounding environmental effects (see Methods for details).

chamber, with caterpillars fed fresh leaf tissue as needed. In this experiment, 26.1% of the caterpillars survived to pupation and 14.1% survived to eclose as adults (mean survival time = 21.8 days) (Fig. 3a). Mean *L. melissa* weights were 2.94 mg (SD = 2.13) at 8 days, 12.7 mg (SD = 7.71) at 14 days, and 20.0 mg (SD = 7.21) at pupation. Weight and survival were variable within and among groups of caterpillars from different source populations and within and among groups that consumed plants grown from different *M. sativa* source populations (Figs. S4 and S5). Weight and survival metrics of performance were positively correlated, including, 8-day weight vs. 14-day survival (Pearson $r = 0.0916$, 95% confidence interval [CI] = 0.0237–0.159), 14-day weight vs. survival to pupation ($r = 0.472$, 95% CI = 0.416–0.525), and pupal weight vs. survival to eclosion ($r = 0.449$, 95% CI = 0.342–0.545) (Fig. 3b). Past work has shown that weight and lifetime fecundity are highly correlated in *L. melissa* (26).

Plant and caterpillar genetic variation affect performance. Using multilocus genome-wide association methods (see Figs. S6, S7 and S8 for evidence of adequate Markov chain Monte Carlo performance), we found evidence that both *M. sativa* and *L. melissa* genetic variation contributed to caterpillar performance in the common garden rearing experiment (Fig. 4a), consistent with our hypotheses (iii) and (iv) (Fig. 1). Specifically, *M. sativa* genetics (161,008 SNPs) explained between 2% (survival to 8 days) and 36% (14-day weight) of the variation in performance (mean across traits = 17%), and *L. melissa* genetics (63,194 SNPs) explained 5% (weight at pupation and survival to pupation) to 29% (8-day weight) of the variation in the same nine caterpillar performance measures (mean = 15%) (values denote point estimates of the percent variance explained, PVE; see Table S2 for credible intervals; cross-validation results are shown in the next section). Caterpillar genetics contributed more to performance metrics from early development (e.g., 8-day weight and survival to 8 and 14 days), whereas plant genetics mattered more for later development (e.g., 14-day weight, pupal weight, and survival to pupation and adult), resulting in a trend towards a negative relationship between caterpillar and plant genetic contributions across traits (Pearson r = -0.52, 95% CI = -0.88 to 0.22, P = 0.15). We detected mostly positive genetic correlations among performance traits (Fig. 4b), with similar but not identical genetic correlations calculated from *M. sativa* and *L. melissa* polygenic scores (Pearson correlation between *M. sativa* and *L. melissa* genetic correlations, $r = 0.80$, 95% CI = 0.63 to 0.89, $P = 5.923e^{-9}$). Polygenic scores in this context quantify the estimated effect of many plant or caterpillar genetic variants on a performance trait.

Mapping results suggested mostly a polygenic basis for the performance traits, with point estimates of > 10 loci affecting most traits (Tables S2, S3 and S4; Fig. 4c,d), but with more evidence of specific SNPs strongly associated with performance in *L. melissa*. This included ten SNPs with posterior probabilities of association (i.e., posterior inclusion probabilities) > 0.5 with at least one performance trait (Fig. 4d, Table S5). Some of these SNPs were in or near (<20 kbps) genes with biologically plausible functions for affecting performance, such as *MSP-300*, *Lipase member H* and *Juvenile hormone acid O-methyltransferase*, all of which were associated with 8-day weight. For example, *MSP-300* affects muscle development and muscle-ectoderm attachment in *Drosophila* (33). Insect lipases metabolize fats, are expressed in gut tissue, and can affect survival and reproductive capacity in insects; *Lipase member H* in particular has further been associated with viral resistance in the moth *Bombyx mori* (34, 35). *Juvenile hormone acid O-methyltransferase* is involved in juvenile hormone biosynthesis and thus in the regulation of insect growth and development, especially metamorphosis (36, 37). A single *M. sativa* SNP was strongly associated with survival to pupation (posterior inclusion probability > 0.5; chromosome 1, position = 12,930,966 bp). This SNP was found in a gene encoding *TOM1-like protein 9* and was within 30 kbps of six additional genes, including two genes with known links to plant-insect interactions: *dentin sialophosphoprotein*, which is associated with soybean compensatory growth after cutworm herbivory (38), and *photosystem I reaction center subunit psaK*, which has been mechanistically linked to tolerance to aphids and aphid feeding preference in *Arabidopsis* (39) (Table S6). We obtained similar results with complementary genetic mapping analyses that included 20 genetic PCs as additional controls for population structure when estimating SNP-performance associations; this was true both in terms of the percentage of variation in performance explained (Pearson correlations >

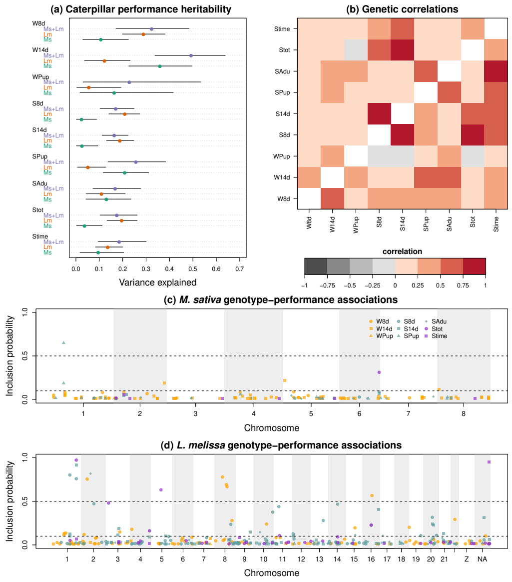


Fig. 4. Genetic mapping of caterpillar performance. (a) Dotchart shows Bayesian estimates of the proportion of trait variation explained by *M. sativa* genetics (Ms), *L. melissa* genetics (Lm), or both combined (Ms+Lm) for each caterpillar performance trait; W8d = 8-day weight, W14d = 14-day weight, Wpup = pupal weight, S8d = 8-day survival, S14d = 14-day survival, SPup = survival to pupation, SAdu = survival to adult, Stot = total survival time, and Stime = (truncated) survival time. Points and horizontal lines denote point estimates (posterior medians) and 95% equal-tail probability intervals, respectively. (b) Heatmap shows genetic correlations between pairs of caterpillar performance traits based on *M. sativa* genetics (lower triangle) or *L. melissa* genetics (upper triangle). Manhattan plots in (c) and (d) shown posterior inclusion probabilities (PIPs) for genotype-performance associations based on *M. sativa* and *L. melissa* SNPs, respectively. Points denote SNPs with different colors and symbols for different performance traits. Only SNPs with PIPs $\geq$ 0.01 are depicted. Horizontal lines at PIPs of 0.1 and 0.5 are included for reference.

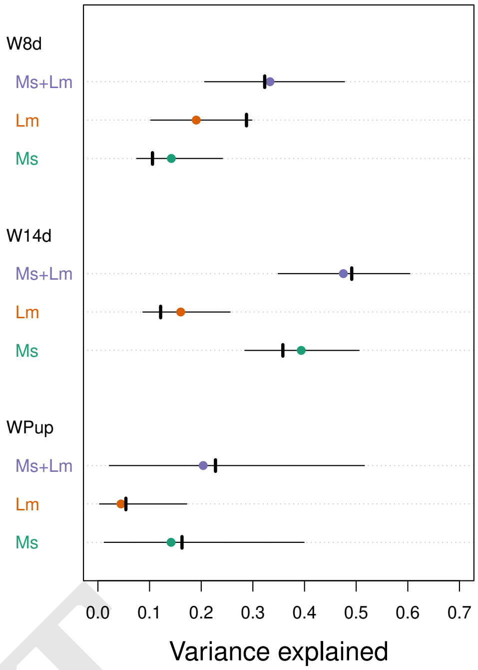


Fig. 5. Genetic mapping of caterpillar performance with epistasis. The dotchart shows Bayesian estimates of the proportion of trait variation explained by *M. sativa* genetics (Ms), *L. melissa* genetics (Lm), or both combined (Ms+Lm) for 8-day weight (W8d), 14-day weight (W14d), and pupal weight (Wpup). Points and horizontal lines denote point estimates (posterior medians) and 95% equal-tail probability intervals, respectively, for the proportion of trait variation explained by additive effects and pairwise epistatic effects. Vertical black lines denote point estimates (posterior medians) for the proportion of variation explained by additive genetic effects alone (as presented in Fig. 4a).

0.99, $P < 0.001$ for caterpillar and plant genetics) (Tables S7 and S8, Fig. S9), and in terms of specific SNP-performance associations (Pearson correlations for posterior inclusions probabilities, *M. sativa* $r = 0.76$, $P < 0.001$, *L. melissa* $r = 0.98$, $P < 0.001$) (Fig. S10).

We repeated the genetic mapping approach using a combined data set of both *M. sativa* and *L. melissa* genetic loci (i.e., the combined 224,202 SNPs) (genetic PCs were not included here or in subsequent analyses). The combined data set generally explained more of the variation in caterpillar performance, 17% to 49% (mean = 24%), than either *M. sativa* or *L. melissa* genetic loci alone. Moreover, the combined variation explained for each performance trait was well described by a model where the variance explained separately by plant and caterpillar genetics combined additively. Specifically, in a linear regression model, the percent variance in performance traits explained by plant and caterpillar genetics separately explained 97% of the variation in the estimates of the variance explained by the combined genetic data sets (linear regression, $\beta_{plant} = 1.17$, $P = 6.6e^{-6}$, $\beta_{caterpillar} = 0.80$, $P = 0.00037$, $r^2 = 0.97$) (Fig. S11), consistent with our hypothesis (iii) (Fig. 1).

Given the evidence of additivity of genetic effects between species presented above, we next turned to more direct tests of the hypothesis that epistatic interactions contribute to caterpillar performance, with a specific focus on caterpillar and pupal weight (see Methods for details and justification). To minimize the low power associated with testing all SNP-SNP interactions, we tested for marginal epistasis, that is for evidence of an epistatic interaction between each SNP and any of the other SNPs. We failed to find significant evidence of marginal epistasis among *M. sativa* SNPs, among *L. melissa* SNPs, or between *M. sativa* and *L. melissa* SNPs (i.e., no SNPs achieved genome-wide significance) (Figs. S12 and S13). Our failure to find epistasis in this manner could be driven in part by limited power to detect it. Thus, we next re-fit the multilocus genome-wide association models described above, but with additional terms for epistatic interactions. This allowed us to directly ask where including epistasis increases our ability to explain caterpillar performance. To do this in a statistically tractable way, we added pairwise interactions between the 150 SNPs with the most evidence of marginal epistasis (i.e., lowest P-values); this added an additional 11,175 terms to each model. Models including these epistatic effects failed to explain more of the variation in caterpillar performance than our purely additive models (Fig. 5). Thus, these direct tests of epistasis provide additional evidence against hypothesis (iv) and thus in favor of hypothesis (iii) (i.e., additivity within and between species) (Fig. 1). Consequently, we focus on the additive models in tests of predictive power below, before presenting additional tests of additivity versus epistasis in subsequent sections of this paper.

Predicting caterpillar performance from plant and caterpillar genotype. We next showed that our genotype-phenotype models were moderately successful at predicting caterpillar performance. This is relevant both for validating these models and for demonstrating their potential utility and limitations in making predictions about effects and evolutionary trajectories in nature. Specifically, genomic predictions of performance from 10-fold cross validation exhibited statistically significant positive correlations with observed performance values for three out of ten performance traits for *M. sativa* genetics, five out of ten traits for *L. melissa* genetics, and six out of ten traits for *M. sativa* and *L. melissa* genetics combined (Fig. 6a). Especially pronounced positive correlations between observed and predicted performance were detected for 14-day weight based on *M. sativa* genetics and 8-day and 14-day survival based on *L. melissa* genetics. More generally, our ability to predict performance traits was well explained by our estimates of heritability (i.e., PVE): we calculated Pearson correlations of 0.89 (95% CI = 0.55 to 0.98, $P = 0.0013$) and 0.62 (95% CI = -0.073 to 0.91, $P = 0.074$) between PVE estimates and the correlation between observed and predicted traits for *M. sativa* and *L. melissa* genetics, respectively (Fig. 6b). In other words, we better predicted caterpillar performance for the performance traits that were more heritable.

Having demonstrated moderate predictive power within the main common garden, we next asked whether genotype-phenotype models estimated from this garden could successfully predict *L. melissa* performance for additional caterpillars fed *M. sativa* from a second, smaller common garden (the Gene Miller Life Science Garden; $N = 180$ plants) (Fig. S1). This second garden, planted in 2018 on the Utah State University campus $\sim$2.5 km from the Greenville Experimental Farm garden, included plants from six of the 11 *M. sativa* source sites and caterpillars from each of the sites used in the main experiment. Survival rates for caterpillars reared on plants from this garden were similar to those reared on plants from the main garden (Fig. S14). Predictive performance for the second garden differed notably for *M. sativa* versus *L. melissa* genotype-phenotype models, with statistically significant positive correlations between observed and predicted trait values in the new garden for only one trait for *M. sativa* genetics versus six of the ten performance traits for *L. melissa* genetics (Fig. 6c). Predictions for the combined data set were similar to those based on *L. melissa* genetics alone. Consistent with these patterns, estimates of PVE from the main garden explained predictive power for *L. melissa* genetics (Pearson $r = 0.93$, 95% CI = 0.68 to 0.98), but not *M. sativa* genetics ($r = 0.17$, 95% CI = -0.56 to 0.75). Thus, unmeasured environmental differences likely limit our ability to predict performance from plant genetics across gardens to a much greater extent than for caterpillar genetics (plant growth environments differed, but caterpillar rearing environments did not) despite these gardens being separated by only $\sim$2.5 km. Differences in the exact genetic composition of the two gardens could add to this effect.

Genetic associations with plant traits explain the plant genetic contribution to caterpillar performance. Having shown that plant genetic variation affects caterpillar performance, we now focus on the Greenville Experimental Farm (see Fig. S1a) to identify possible components of the functional basis of the documented plant-genetic effects. This also allowed us to further test for additive versus epistatic interactions between plant and caterpillar genotypes on caterpillar performance (see our hypotheses (iii) versus (iv) in Fig. 1). We first determined the extent to which genetic loci associated with caterpillar performance were also associated with other plant traits, including potential plant vigor or defense traits (17). Such genetic correlations could arise from pleiotropy, but also from linkage disequilibrium between distinct loci affecting the plant traits and caterpillar performance (i.e., genetic correlations do not demonstrate a causal genetic link between traits). Still, such an association would be consistent with the hypothesis that these traits, or other genetically correlated traits, constitute possible mechanisms by which plant genotype affects caterpillar performance. To do this, we measured and mapped 1760 plant traits in the Greenville Experimental Farm garden using the same multilocus mapping approach and *M. sativa* SNP data set described above. The traits included plant height, leaf length, leaf width, leaf area, leaf shape, leaf weight, specific leaf area, leaf toughness, trichome density, levels of herbivory on the plants in the field, and 1750 plant chemistry metabolites, which were quantified and characterized using liquid-chromatography combined with mass spectrometry (LC-MS; similar to (25, 32)).

We documented genetic variation affecting most of the plant traits, with mean PVEs of 20.5% for the non-chemical traits (minimum = 5.6%, maximum = 38.7%) and 10.9% (310 traits > 20% and 20 > 50%) for the 1750 chemical traits (Table S9). Additionally, in the main Greenville Experimental Farm common garden, the distribution of PVE for the 1750 chemical traits differed markedly from that for 1750 matched, randomized traits, consistent with a clear genetic contribution to this variation in leaf metabolites (Fig. S15).

Multiple plant traits, including chemical and non-chemical traits, exhibited genetic correlations with each caterpillar performance trait; in other words, plant trait polygenic scores were correlated with caterpillar performance polygenic scores when inferred from plant genetics (Figs. 7a,b, S16). However, because of the large number of measured traits and genetic correlations among the plant traits (Fig. S17), many of the genetic correlations between plant traits and caterpillar performance were likely redundant. Thus, to identify the combined subset of traits most strongly predictive of caterpillar performance (and thus the best candidates for a mechanistic link to performance), we next fit a LASSO penalized regression model for the polygenic scores of each caterpillar performance trait (based on plant genetics) as a function of the polygenic scores for the 1760 plant traits. These models explained 41 to 80% of the variation in the caterpillar performance scores (mean = 69.2%, cross-validation predictive r^2 ranged from 0.39 to 0.76) (Table S10, Fig. 7c). On average 260 of the 1760 traits were retained in these models (i.e., given non-zero regression coefficients), with a range of 117 (survival time) to 347 (8-day survival) traits (Figs. 7d,e and S18). Both chemical and non-chemical traits were retained in the models. Non-chemical traits with the biggest effects included a positive effect of plant height on 14-day weight ($\beta = 0.037$), positive effects of trichome density ($\beta = 0.036$) and specific leaf area ($\beta = 0.031$) on survival to adulthood, and a negative effect of leaf toughness on survival to adulthood ($\beta =$ -0.34). Consistent with a previous phenotypic assay of caterpillar performance and plant metabolomic variation in this system (32), top chemical traits included several saponins, including saponins (two distinct Medicagenic acids) associ-

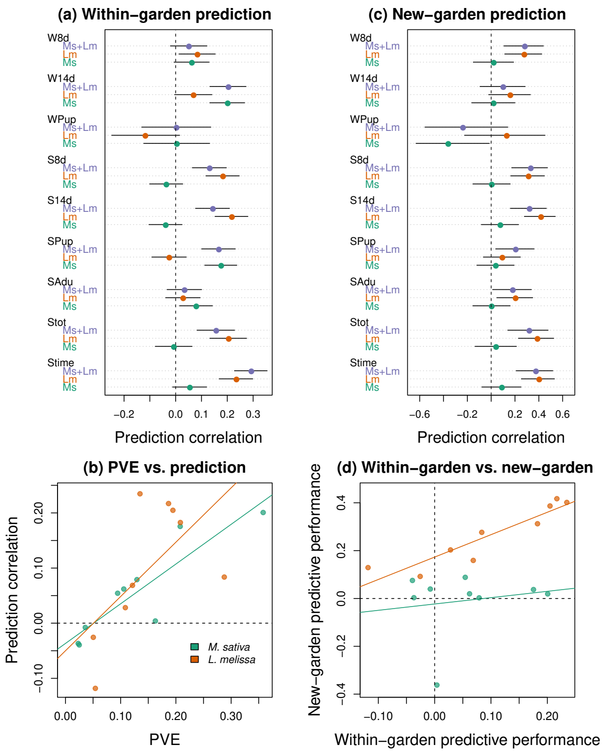


Fig. 6. Genomic prediction of caterpillar performance. (a) Dotchart shows Pearson correlations between cross-validation genomic predictions of phenotypes and the observed values based on *M. sativa* genetics (Ms), *L. melissa* genetics (Lm), or both combined (Ms+Lm) for each caterpillar performance trait; W8d = 8-day weight, W14d = 14-day weight, Wpup = pupal weight, S8d = 8-day survival, S14d = 14-day survival, SPup = survival to pupation, SAdu = survival to adult, Stot = total survival time, and Stime = (truncated) survival time. Points and horizontal lines denote point estimates (posterior medians) and 95% equal-tail probability intervals, respectively. For example, a large value on the x axis indicates a high correlation between observed performance values and predictions from genotype based on cross validation. (b) Scatterplot of the proportion of variation explained by genetics (PVE) versus the Pearson correlation of genomic predictions from (a). Each point denotes a trait and is colored to indicate values from *M. sativa* or *L. melissa* genetics. Colored lines are best fits from ordinary linear regression, and a dashed line denotes the 0 value on the y-axis. (c) Dotchart similar to (a), but for genomic predictions of phenotypes in a second common garden (the Gene Miller Life Science Garden) based on the models fit from the main garden. (d) Scatterplot of correlations between observed caterpillar performance trait values and genomic predictions of these values using cross-validation within the main garden versus prediction for samples in the Gene Miller Life Science Garden based on the models fit for the main garden.

ated with effects on caterpillar weight and survival (Tables S12 and S13). The flavonoid glycoside Tricin 7-glucoside was associated with reduced survival, whereas several peptides (e.g., MESA.583 = $C_{13}H_{18}O$, a fragment of a N-acyl amine; MESA.615 = $C_{23}H_{43}N_7O_7$; MESA.849 = $C_{14}H_{19}NO_3$, a N-acyl amine) were associated with reduced weight or survival (Tables S12 and S13). Lastly, we fit LASSO regression models on the 1064 principal components (PCs) from an ordination of the plant trait and chemistry polygenic scores, which represent 1064 independent (orthogonal) variables. Our goal here was to provide additional evidence that multiple, distinct genetic factors contributed to explaining caterpillar performance polygenic scores. Models based on these predictors explained 27 to 76% of the variation in the caterpillar performance scores (mean = 56.6%, cross-validation predictive r^2 ranged from 0.25 to 0.72), with an average 180 of the 1064 PCs retained in the LASSO models (range = 52 to 337) (Fig. S19).

Compared to predicting polygenic scores for caterpillar performance, our ability to predict caterpillar performance at the phenotypic level from plant-trait polygenic scores was notably reduced (Table S10, Figs. S20 and S21). This was expected as plant genetics only explained a modest proportion of the variation in performance and thus the ability to explain variation in these traits (not just polygenic scores) was necessarily capped by performance-trait heritabilities. Still, when considering all performance traits together, plant trait polygenic scores explained more of the trait variation than expected by chance (Fisher combined test, $\chi^2 = 34.42$, df = 18, $P = 0.011$). This signal was driven primarily by association of plant traits with 8 and 14 day weight and survival to pupation and eclosion.

Lastly, we determined the extent to which the association of plant trait polygenic scores with caterpillar performance polygenic scores (both inferred from plant genetics) was affected by *L. melissa* genotype. Such an interaction would suggest caterpillar performance is affected by epistatic interactions between *M. sativa* and *L. melissa* genotypes, as predicted by our hypothesis (iv) (Fig. 1). We used principal component (PC) scores from the first four principal components of the *L. melissa* genotype matrix, which together accounted for ∼15% of the *L. melissa* genetic variation, as summaries of *L. melissa* genotype. We then fit LASSO penalized regression models for caterpillar performance polygenic scores as a function of these PC scores, plant trait polygenic scores, and interactions between each plant trait polygenic score and each of the four PCs. This allowed us to test for epistasis at the level of plant morphology and phytochemistry polygenic scores from *M. sativa* and four axes of *L. melissa* genetic background and thereby avoid the lack of power that would be associated with exhaustively testing SNP-SNP interactions (nonetheless, these models still included $4 \times 1760 = 7040$ possible interaction terms). We found no evidence of epistasis between *M. sativa* and *L. melissa* affecting caterpillar performance. Specifically, including these interaction terms in the models actually reduced the variance explained by the LASSO models (Table S11) and the interaction terms were retained less frequently in the models than the non-interaction terms (Figs. S22 and S23). We obtained similar results when fitting models for caterpillar performance trait values rather than polygenic scores, with a smaller proportion of interaction terms retained in the model for most traits (Figs. S24 and S25) and no overall increase in variance explained by models with versus without interactions (i.e., the variance explained in 14-day weight doubled, but the variance explained in 8-day weight was halved, and there was no detectable general increase in variance explained across traits) (Fig. S26). Thus, these results support our hypothesis (iii) with additive contributions of plant and caterpillar genetics (Fig. 1).

Plant and caterpillar genetics have consistent effects on performance. We conducted two additional experiments to determine the extent to which genetic differences among *M. sativa* plants or populations had consistent affects on caterpillar performance for different butterfly populations and species. This constitutes another test of additivity versus epistasis for plant and insect genotypes (our hypotheses (iii) versus (iv) in Fig. 1) and of the potential for our findings to provide general predictions beyond our main study populations. In the first of these experiments, *L. melissa* (Lycaenidae) caterpillars from four populations were reared on greenhouse-grown *M. sativa* sourced from six sites (Table S14). Two additional butterfly species, *Colias eurytheme* (a legume specialist) (Pieridae) and *Vanessa cardui* (a generalist that rarely feeds on alfalfa) (Nymphalidae), were reared on these same plants. Whereas only modest genetic differences exist among the *L. melissa* populations (Fig. 2) (23, 24), these three butterfly species are deeply divergent ($\sim$ 100 million years) creating substantial opportunities for the effect of *M. sativa* genotype and phenotype to interact with genetic differences among the butterfly taxa (40). Caterpillars were fed leaf tissue from multiple individual plants, but each caterpillar was given plants from a single source population and leaves from each plant were fed to all three butterfly species. Survival rates were highest for *C. eurytheme*, followed by *L. melissa* and lastly *V. cardui* (Fig. S27). Plant population (here used as a proxy for plant genotype) explained $\sim$3-10% of the variation in 8-day weight for each butterfly species, and 9-14% of the variation in 14-day weight, with larger effects in the butterfly species less-well adapted to *M. sativa* (Table S15). Caterpillar population explained a small but non-zero proportion of the variation in 8-day weight in *L. melissa* (this could not be assessed in the other species), but not a significant amount of variation in 14-day weight. Thus, consistent with our main results above, genetic differences among plant and caterpillar populations (caterpillar populations for *L. melissa* only) explained variation in caterpillar performance, with plant genetics mattering more for 14-day weight than 8-day weight and caterpillar genetics mattering more for 8-day weight than 14-day weight. Plant population and plant maternal family also explained variation in plant growth and development traits, consistent with our common garden results above (Table S9). Importantly, the effect of each plant population on caterpillar performance was remarkably consistent across *L. melissa* populations and even across different species, with moderate to large positive correlations (though not always significantly so) in the effect of each plant population on 8 and 14-day weight across all pairs of population and species (Fig. S28).

The final complementary experiment used the same three butterfly species: *L. melissa*, *C. eurytheme*, and *V. cardui*, but instead involved feeding each caterpillar leaf tissue from a single *M. sativa* plant from a third common garden near the University of Nevada (UNR Main Station in Reno, NV; Fig. S1). We used these data to ask whether the effect of plant genotype (here, individual plant) on caterpillar weight was consistent across species. We detected modest, positive pairwise correlations between the three species of caterpillars, suggesting a degree of similarity of plant genotypes that affect performance of these different herbivorous species (Fig. S29). Specifically, the correlations were as follows: *V. cardui* vs. *L. melissa* $r = 0.33$ ($P = 0.015$, t $= 2.52$, df $= 52$); *C. eurytheme* vs. *L. melissa* $r = 0.43$ ($P = 0.0010$, t $= 3.48$, df $= 52$); *C. eurytheme* vs. *V. cardui* $r = 0.15$ ($P = 0.28$, t $= 1.08$, df $= 52$). Thus, these two experiments combined with our main results show that genetic variation within *M. sativa* affects caterpillar performance across populations and species of butterflies in a remarkably consistent manner, consistent with the additivity hypothesis (hypothesis (iii) in Fig. 1).

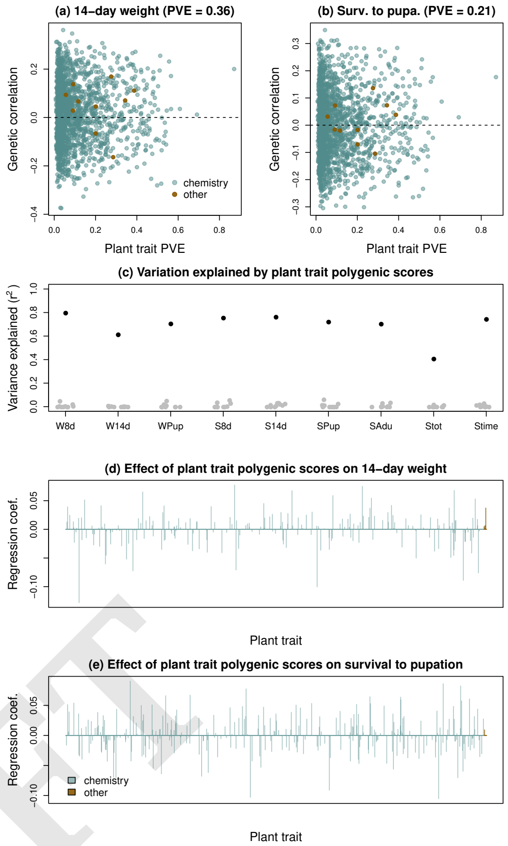


Fig. 7. Associations between plant trait polygenic scores and caterpillar performance polygenic scores. Scatterplots show genetic correlations between plant chemistry and other plant traits and 14-day caterpillar weight (a) and survival to pupation (b) inferred from plant genetics as a function of the proportion of plant trait variation explained by genetics (PVE). A dashed horizontal line denotes a genetic correlation of zero. Panel (c) shows the variance explained by lasso regression models of caterpillar performance polygenic scores estimated from plant genetics as a function of polygenic scores for 1750 plant chemistry traits and 10 non-chemistry traits. Black dots denote inferred values of r^2 and gray dots show similar estimates using randomized plant trait polygenic scores (10 random data sets each). Panels (d) and (e) show standardized regression coefficients from the lasso models for 14-day weight (d) and survival to pupation (e).

Discussion. From an ecological perspective, the greatest diversity of life is not counted in the number of species or other taxonomic units, but in the diversity of inter-specific interactions (41). The ubiquity of plant-feeding insects has made them a focal point for understanding the evolution, persistence, and variability of interactions (9, 42, 43). The outcomes of plant-insect interactions (e.g., caterpillar survival) might depend on genetic variation within each species and these genetic effects could compound additively or non-additively. Taken all together, our results support the hypothesis that both plant (alfalfa) and insect (Melissa blue butterfly) genotype matter for caterpillar growth and survival, and that these contributions are mostly additive (our hypothesis (iii) in Fig. 1). These results are qualitatively similar to those reported in another study (20), which identified individual plant (*Arabidopsis thaliana*) and caterpillar (*Pieris rapae*) genes affecting caterpillar performance. The advance over previous work that we offer here is in quantitative, genomic prediction of caterpillar performance, which, in contrast to the identification of specific genes, provides a formal connection from trait genetics to models of evolution for quantitative traits (44). We specifically demonstrated that the combined effects of plant and insect genotype explain a substantial proportion of variation in caterpillar growth and survival (17-49%), and that these mostly-additive effects can predict performance from genotypes in cross-validation analyses. Moreover, models that included pairwise epistatic effects failed to explain caterpillar performance better than the additive-only models. We were able to identify specific traits and phytochemicals associated with the plant contribution to performance, most notably plant size, and several saponins, peptides, and phosphatidyl cholines. Whereas some of these classes of chemicals (e.g., saponins) are best known as insect toxins or feeding-deterrents (e.g., (45–47)), our results suggest these classes include molecules with positive and negative effects on performance, consistent with other recent metabolomic work (25, 32). We also found evidence that plant genotype had consistent effects on performance in multiple butterfly populations and distantly related species, including a second legume specialist (*C. eurytheme*) and a generalist (*V. cardui*). This too is consistent with results from the only other similar study (20), which documented conserved changes in gene expression in response to herbivores across multiple plant and butterfly species. This consistency is relevant to the predictability and nature of the evolution of plant-insect interactions, as we discuss more below.

Our results have clear implications for the study of coevolution, which takes many forms and pertains to the formation of new species and new interactions (43). Quantitative theories of coevolution have historically been dominated by gene-for-gene models, in which the fitness of a particular genetic variant in (for example) a parasite is conditioned on the presence of a specific gene in the host (22). Evidence in support of gene-for-gene models has come mostly from plant-pathogen systems (22) (but see (48)). In contrast, diffuse models of coevolution relax some of the expectations for gene-by-gene interactions, and have been favored by researchers working with more macroscopic parasites, including herbivorous insects (49). However, relevant investigations in plants and insects have mostly relied on experiments that contrast categories of individuals (strains or biotypes) rather than more comprehensive or continuous variation in genetically variable populations (reviewed in (12)), which has left the field with uncertainty regarding the most relevant theoretical context for the diversity of evolving plant-insect interactions. The results that we report are not consistent with the gene-for-gene model of coevolution, as the performance of our focal herbivore was both highly polygenic and successfully predicted without interactions between caterpillar and plant genotypes. Instead, our results suggest that genetic differences in plant quality and defense have similar effects regardless of insect genotype or even species.

Our results also shed light on the evolution of diet breadth and host use in herbivorous insects. Specifically, the finding of substantial heritable variation in the Melissa blue butterfly for growth and survival suggests that ongoing adaptation to alfalfa, which at present is a marginal host (26), is not constrained by a lack of genetic variation. This is consistent with earlier work on this system (28). Likewise, alfalfa appears to harbor genetic variation to evolve traits that reduce the success of the Melissa blue even further, and this inference likely extends to other herbivores given the consistent effects of plant variation on other butterfly species reported here and on other herbivores in an observational study (25). While the persistence of plant genetic variation affecting herbivores might be attributable to the age of these interactions (since most herbivores of alfalfa in North America are recent colonists), we suspect other factors are more important. First, the asymmetry in our predictions, with consistent caterpillar-genetic effects but not plant-genetic effects on performance between common gardens, suggests a major role for plasticity in the effect of plant genotype on caterpillar performance. This is not surprising given considerable evidence that biotic and abiotic environmental factors affect plant quality and plant defenses in alfalfa (31) and other plants (50), but does mean genetic variation in performance measured in the lab and common garden might not strongly predict effects in specific natural populations (51). Moreover, other biotic and abiotic factors could contribute more to caterpillar growth and survival in the wild, and some of these could interact with plant genotype. For example, recent work has shown that the abundance of ants, which tend Melissa blue caterpillars and thereby reduce the threat from enemies (see image in Fig. 1), greatly increases caterpillar survival and population persistence on alfalfa, with ant abundance indirectly affected by alfalfa phytochemistry (25, 52). In contrast to the complexity of plant effects, the more consistent effects of caterpillar genetic variants raises the possibility that the ability of herbivores to successfully utilize plants might more readily evolve, while the ability of plants to evolve defenses will be more contingent (on local environments, etc.). This again supports a diffuse model of coevolution (49) and could eventually help us understand the accumulation of host-specific herbivores on plants through evolutionary time.

Beyond issues specific to herbivorous insects and their host plants, genetic variation within species is important for host-parasite interactions (53), including for example susceptibility to parasitic diseases in humans and other animals being a function of both genetic variation in the hosts and among pathogen strains (54). However, as is the case for plant-insect interactions, genomic investigations of other pairwise interactions have rarely considered both species simultaneously, but have focused on either the host or parasite. If epistatic, among-species interactions were common (as assumed by the gene-for-gene model of coevolution), the piecewise approach (focusing on one interacting species rather than the pair) might be a major roadblock to progress in understanding the evolution of these systems. However, if additivity and consistency of polygenic effects hold generally, as documented in the plant and herbivores studied here, a focus on one species in an interaction might not be misleading, and might inform predictive models, but this hypothesis remains to be tested with other interacting species.

Materials and Methods

Establishing the primary common garden. We planted a common garden comprising 1080 alfalfa plants at the Greenville Experimental Farm near Logan, Utah (41.765° N, 111.814° W) in 2018 (Fig. S1a). Seeds for this garden were collected from 11 naturalized alfalfa sites in the western USA, including five sites where *L. melissa* are found, and six sites lacking *L. melissa* butterflies (Table S1). An average of 4.9 seeds were planted from each of 220 maternal plants (with an average of 97.6 seeds planted from each site, SD = 8.6, range = 77 to 105) (Table S1). See "Establishing the primary common garden" in the SI for additional details.

Caterpillar husbandry and performance assays. We obtained *L. melissa* eggs from gravid females collected from six sites between June 16th and July 4th 2019 (Table S1). As in past work, gravid females were caged with a few sprigs of host plant (*M. sativa*) and allowed to lay eggs (17, 26, 28). Eggs were kept in a Percival incubator (model no. 136VL) at 27°C with 14 hours light:10 hours dark. Upon hatching, caterpillars were assigned randomly to feed on a specific *M. sativa* plant. Each neonate caterpillar was carefully transferred to a Petri dish with a sprig of fresh plant material (a few leaflets) with the stem of the plant tissue wrapped in a damp Kimwipe. We verified each caterpillar was alive and uninjured after transfer. The Petri dish containing the caterpillar was then returned to the incubator. Caterpillars were given fresh leaf tissue *ad libitum* and were checked daily for survival, pupation and eclosion as adults.

As metrics of performance, we measured 8-day and 14-day caterpillar weight, and weight at pupation using a Mettler Toledo XPE105 analytical microbalance (Mettler Toledo). Weights were recorded to the nearest 0.01 mg, and we took the mean of two independent weight measurements. *Lycaeides melissa* caterpillars generally spend 20 to 30 days as larvae (17), and weight and lifetime fecundity are highly correlated in *L. melissa* (26). We then considered the following nine performances metrics: 8-day caterpillar weight (mg), 14-day caterpillar weight (mg), weight at pupation (mg), survival to 8 days (binary), survival to 14 days (binary), survival to pupation (binary), survival to adult (binary), total survival time (integer valued), and truncated survival time (integer valued). For truncated survival time, we truncated survival at the maximum number of days required for any of the caterpillars to reach eclosion; this avoids caterpillars that developed slowly but never pupated or eclosed from having the longer survival times than caterpillar that successfully eclosed as adults.

Generating the genetic data. We extracted DNA from 1236 *M. sativa* plants and 1079 *L. melissa* caterpillars, pupae or adults reared on these plants. We then generated partial genome sequences for each organism using our genotyping-by-sequencing approach (23, 55); this produced ∼2.5 billion reads for *M. sativa* and ∼2.5 billion reads for *L. melissa* (see "DNA extraction and sequencing" in the SI). We then aligned the DNA sequences to the *M. sativa* or *L. melissa* genome and identified SNPs using `samtools` (versions 1.10), `bcftools` (version 1.9) and `GATK` (version 4.1) (56, 57) (see "DNA sequence alignment and variant calling" in the SI). After filtering, we retained 161,008 SNPs for *M. sativa* and 63,194 SNPs for *L. melissa*. We then estimated genotypes using the Bayesian (ad)mixture model implemented in `entropy` (version 2.0) (23, 58) (see "Inference of genotypes and genetic variation" in the SI). Patterns of genetic variation were then summarized with principal component analysis (PCA), and by calculating measures of linkage disequilibrium (LD) and genetic differentiation among samples from different source populations (i.e., F_{ST}) (see "Inference of genotypes and genetic variation" in the SI).

Preparing the caterpillar performance data for genetic mapping. We removed potential confounding variation from the caterpillar performance data prior to analyzing genotype-performance associations. First, we regressed each of the nine caterpillar performance metrics on caterpillar hatch date (to control for temporal effects) and source population (to control for potential non-genetic, e.g., maternal environment, effects). This was done with the `lm` function in `R`. Next, we used distance-based Moran's eigenvector maps to remove possible effects of space (location) within the common garden. This procedure involves creating spatial variables based on a PCA of a truncated (nearest neighbors) Euclidean distance matrix (i.e., a principal coordinates analysis), where distance was defined from the spatial layout of the common garden (59). We then used forward selection of variables following (60) to select spatial variables (eigenvectors) that explained the variation in each trait. Specifically, we first tested for a significant (at $P < 0.05$) fit of a model with all of the spatial variables. If and only if this full model was significant, we began adding spatial variables to a null model one at a time based on the extent to which they increased the total model r^2. This procedure continued until either: (i) the P-value for the most recently added variable was > 0.05, (ii) the total r^2 exceeded the original r^2 from the full model with all variables, (iii) adding the new variable did not increase the model r^2, or (iv) 200 spatial covariates had been added. The final models explained 18 to 51% of the variation in plant traits (mean = 35%) with 22 to 77 covariates retained; however, a model with no spatial covariates was selected for most caterpillar performance traits with 14-day weight being the sole exception (20 covariates explaining 14% of the trait variation). Scaled residuals from the final model for each trait were then used for genetic mapping.

Multilocus genetic mapping of caterpillar performance. We tested for associations between (i) *M. sativa* SNPs (161,008 SNPs), (ii) *L. melissa* SNPs (63,194 SNPs), and (iii) SNPs from both species combined (224,202 SNPs), and each of the nine caterpillar performance metrics (i.e., the residuals from the models described in the previous paragraph). We performed these analyses using Bayesian sparse linear mixed models (BSLMMs), which we fit with `gemma` (version 0.95alpha) (61). A key advantage of this approach for gentoype-phenotype association analyses is that, unlike traditional genome-wide association (GWA) mapping methods that test each genetic marker separately, the BSLMM approach fits all SNPs in a single model and thus mostly avoids issues related to testing large numbers of null hypotheses. The BSLMM method assumes that trait values are determined by a polygenic term and a vector of the (possible) measurable effects of each SNP on the trait ($\boldsymbol{\beta}$) (61). Bayesian Markov chain Monte Carlo (MCMC) with variable selection is used to infer the posterior inclusion probability (PIP) for each SNP, that is, the probability that each SNP has a non-zero effect or association, and the effect size conditional on it being non-zero (62). The polygenic term denotes each individual's expected deviation from the mean phenotype based on all of the SNPs. This term accounts for phenotypic covariances among individuals caused by their relatedness or overall genetic similarity (61). The kinship matrix also serves to control for population structure and relatedness when estimating effects of individual SNPs ($\boldsymbol{\beta}$) along with their PIPs. Similarly, SNPs in linkage disequilibrium (LD) with the same causal variant effectively account for each other, such that only one or the other is needed in the model, and this redundancy is captured by the posterior inclusion probabilities. Moreover, in the context of our study, mapping with plants grown from seed in a common garden and caterpillars reared from eggs in growth chambers substantially reduces some issues related to the confounding effects of population structure, such as genotype-environment correlations, that commonly cause problems in human association-mapping studies (63) and more generally in observational studies of human genetics (64).

The hierarchical structure of the model makes it possible to estimate additional parameters that describe aspects of a trait's genetic architecture (17, 61, 62, 65). These include the percentage of the phenotypic variance explained (PVE) by additive genetic effects (which includes $\boldsymbol{\beta}$ and the polygenic term, and should approach the narrow-sense heritability), the percentage of the PVE due to SNPs with measurable effects or associations (PGE, the percentage of the phenotypic variance explained by genic effects, which is based only on $\boldsymbol{\beta}$), and the number of SNPs with measurable associations (n-γ). All of these metrics use MCMC to integrate over uncertainty in the effects of individual SNPs, including whether these are non-zero. Lastly, using this BSLMM approach, it is also possible to obtain genomic-estimated breeding values (GEBVs) or polygenic scores, that is, the expected trait value for an individual from the additive effects of their genes, as captured by both $\boldsymbol{\beta}$ and the polygenic term (17, 65).

For each of the nine caterpillar performance metrics and three genetic data sets, we conducted 10 MCMC runs with `gemma`, each comprising 1 million iterations and a 200,000 iteration burn-in. Every 10th MCMC sample was retained to form the posterior distribution. Polygenic scores (i.e., genomic-estimated breeding values) were then calculated from the genetic data sets and model-averaged effect estimates for each SNP locus; these incorporate the polygenic term as is standard in genomic prediction methods (e.g., (66, 67)). Genetic covariance matrixes were computed from the estimated polygenic scores.

As noted above, the kinship matrix and multilocus approach of the BSLMM in `gemma` control for confounding effects of population structure and relatedness when testing for individual SNP-phenotype association, but nonetheless, this methods could fail to fully capture complex patterns of structure (see, e.g., (63)). Thus, to verify the robustness of our results, we fit additional models using the BSLMM approach in `gemma` where we included the first 20 genetic PCs as potential covariates to further account for population structure. This was done as described above, except that the analysis was only conducted for *M. sativa* and *L. melissa* SNPs separately. We compared this to our main results both in terms of the percentage of variation in performance explained by genetics (PVE) and specific SNP-performance trait associations.

Direct tests of epistasic genetic effects on caterpillar performance. We tested for epistatic interactions affecting caterpillar performance among (i) the 161,008 *M. sativa* SNPs, (ii) the 63,194 *L. melissa* SNPs, and (iii) the 224,202 SNPs from both species (this includes within and between species epistatic interactions). We conducted these tests with `MAPIT` (https://github.com/lorinanthony/MAPIT) (68). Exhaustive testing of all pairwise SNP-SNP interactions suffers from low statistical power because of the large number of tests involved. The statistical method in `MAPIT` overcomes the problem of low power by instead testing for marginal epistatic effects, that is testing the null hypothesis that a given SNP does not interact with any of the other SNPs (i.e., that the variance component for epistatic effects is 0) (68). This is done without trying to identify the specific SNPs a focal SNP interacts with. We computed P-values for tests of marginal epistasis using the recommended hybrid method that first implements a z-test to compute a P-value and then re-computes the P-value with the Davies method if the initial values is less than 0.05 (as in (18)).

For many of the survival traits, we observed an unexpected excess of very low P-values, especially for *L. melissa* SNPs and for 8 and 14-day survival (Fig. S30). We strongly suspect this is a statistical artefact, especially as these measures constitute residuals from integer or binary traits and the control kinship matrix consists of relatedness based on plant and insect genetics, a combination of complications that could be problematic for this method and inflate type-I errors (note that this differs from the BSLMM in `gemma` where the multilocus approach allows SNPs to serve as controls for each other). Given our concern that these results are not biologically meaningful, we conservatively focus only on the weight measurements when presenting these tests of epistasis, as these do not appear to suffer from the same issue (see Figs. S12 and S13).

Even with the `MAPIT` method, a potential exists for tests of epistasis to be underpowered, especially in terms of achieving strict, genome-wide significance. Thus, we conducted additional analyses using the BSLMM approach from `gemma` to test for associations between *M. sativa* and *L. melissa* genetics and caterpillar performance but where we included pairwise epistatic effects among SNPs with the most evidence of marginal epistasis from the `MAPIT` analyses (similar to (18)). Our goal was to ask whether including these additional epistatic terms improved the explantory power of the model. In these analyses, we considered only the caterpillar weight traits (for the reasons noted above). We included either (i) the top 150 SNPs with the lowest P-values for marginal epistasis within species (for analyses with only *M. sativa* or *L. melissa* SNPs) or (ii) the top 75 SNPs from each species with the lowest P-values for marginal epistasis in the combined species analysis. We then created new genetic covariates for all pairwise interactions between pairs of the 150 SNPs ($\frac{150\times149}{2} = 11,175$ potential epistatic effects). We did this by taking the product of the centered and standardized genotypes for each pair of loci. These were then included in the BSLMM model for `gemma` (though not in the construction of the kinship matrix, which was solely based on additive effects). We fit these models as described above, except we increased the number of MCMC iterations and burnin to 2 million and 400,000, respectively. We then determined the total PVE in weight explained by the models with additive and epistatic effects for *M. sativa* genetics, *L. melissa* genetics, and both *M. sativa* and *L. melissa* genetics combined.

Within-garden cross-validation and genomic prediction. We used 10-fold cross-validation to assess our ability to predict performance traits from *M. sativa* genetic data, *L. melissa* genetic data, and the combined genetic data from *M. sativa* and *L. melissa*. To do this, we first randomly assigned each observation to one of ten test data sets. Then, for each test data set, we estimated genotype-phenotype associations using `gemma` as described above, but based only on the 90% of individuals not in that test data set. For this, we used a single MCMC run comprising 1 million iterations, a 200,000 iteration burn-in, and a thinning interval of 10. We then used `gemma` to predict the phenotypes of the 10% of individuals held back for the test set (these individuals were not used to fit the model); this was done with the predict option in `gemma`. We then quantified predictive performance using the Pearson correlation between the genomic predictions of each performance metric and the observed values.

Gene Miller Life Science Garden set up and genomic prediction. We further tested our ability to predict caterpillar performance trait values from genotypes by generating genomic predictions of performance for caterpillars reared on *M. sativa* from a second, smaller common garden comprising 180 *M. sativa*–The Gene Miller Life Science Garden (see “Establishing the Gene Miller Life Science Garden" in the SI for details). We used leaf tissue from these plants for rearing *L. melissa* caterpillars in the summer of 2019 exactly as described for the main common garden at the Greenville Experimental Farm (see ‘Caterpillar husbandry and performance assays’ above for details). This parallel experiment was conducted at the same time as the main experiment. Plant and caterpillar samples from this parallel experiment were sequenced along with the samples from the Greenville Experimental Farm experiment. We successfully obtained genetic data from 172 *M. sativa* and 156 caterpillars of the 180 involved in this experiment. These genetic data were processed along with those from the main garden (see ‘DNA sequence alignment and variant calling’ above for details).

We then used the estimated, model-averaged effects from the BSLMM fits in `gemma` from the main garden to predict performance traits based on plant, caterpillar, or plant and caterpillar genotypes for these individuals. We compared these genomic predictions (i.e., polygenic scores computed from the main-garden models) to the observed performance trait values for these caterpillars. This was done using residuals after removing effects of hatch date and block (i.e., plot) within the USU garden. As with the within-garden cross-validation analyses described in the previous section, predictive power was measured by the Pearson correlation between the predicted and observed performance trait values.

Plant trait measurements and phytochemical analysis. We measured a series of morphological traits potentially associated with plant vigor or resistance to insects (e.g., putative structural plant defenses)(17, 69, 70) for each of the 1080 *M. sativa* plants in the Greenville Experimental Farm common garden: plant height, leaf length, leaf width, leaf area, leaf shape, leaf weight, specific leaf area, leaf toughness, trichome density, levels of herbivory on the plants in the field, and 1750 plant chemistry metabolites, which were quantified and characterized using liquid-chromatography combined with mass spectrometry (LC-MS). See “Plant trait measurements" and “Sample extraction and phytochemical analysis" in the SI for details. We further annotated the 20 phytochemicals that were most stronlgy associated with caterpillar performance (see “Structural annotations of phytochemicals" in the SI).

Multilocus genetic mapping of plant traits. We tested for associations between the *M. sativa* SNPs (161,008 SNPs) and 1760 plant traits: leaf length, leaf width, leaf area, leaf shape, leaf weight, SLA, trichome density, leaf toughness, plant height, field herbivory and 1750 metabolomic chemical features (see the previous two sections for details). This was done using the 1080 *M. sativa* plants from the main common garden at the Greenvile Experimental Farm in Logan, UT. We first removed possible effects of spatial location within the garden as captured by distance-based Moran’s eigenvector maps using forward selection of variables (60), exactly as described for the caterpillar performance traits above (see 'Preparing the caterpillar performance data for genetic mapping'). As with the caterpillar performance traits, genotype-plant trait associations were estimated by fitting BSLMMs with `gemma` (version 0.95alpha) (61). For each of the 1760 plant traits, we conducted 10 MCMC runs with `gemma`, each comprising 1 million iterations and a 200,000 iteration burn-in. Every 10th MCMC sample was retained to form the posterior distribution. Polygenic scores were then calculated from the genetic data sets and model-averaged effect estimates for each SNP locus. Genetic covariance matrixes were computed from the estimated polygenic scores. The model-fitting procedure was repeated with 1760 randomized plant trait data sets (i.e., values of each of the original traits were permuted among plants) to verify that the distribution genotype-phenotype associations from the real data set differed from null expectations.

LASSO regression models. We used least absolute shrinkage and selection operator (LASSO) regression to (i) identify the subset of plant traits with polygenic scores that best predicted caterpillar-performance polygenic scores and (ii) estimate the direction and magnitude of these associations (see "LASSO regression models" in the SI). We fit additional LASSO models (i) using PCs of the 1760 plant trait polygenic scores as covariates and (ii) to determine whether plant-trait polygenic scores could explain and predict caterpillar performance at the phenotypic level. Lastly, we fit an additional models to evaluate the extent to which plant-genetic effects interacted with caterpillar genetics to affect performance (see "LASSO regression models" in the SI).

Complementary USU greenhouse and Nevada common garden rearing experiments. An additional rearing experiment was conducted using *M. sativa* grown in a USU greenhouse to (i) replicate the general effect of *M. sativa* genotype on caterpillar performance and (ii) determine whether different plant genotypes had consistent effects of caterpillar performance across different butterfly populations and species (i.e., *Colias eurytheme* and *Vanessa cardui*). We performed yet another rearing experiment with the same three species of caterpillars using an experimental *M. sativa* garden experimental garden at the University of Nevada, Reno (Fig. S1). Together, these experiments provide additional tests of additivity versus epistasis with respect to genetic differences among butterfly populations and among deeply divergent species. See "Complementary USU greenhouse experiment" and "Complementary Nevada common garden rearing experiment" in the SI for details.

ACKNOWLEDGMENTS. Thanks to Caitlyn Allgrunn, Jennifer Bryan, Hannah Curtis, Savannah Daines, Skylar Deschene, Shaylen Fidel, Daniel Johnson, Adair Schruhl, Amy Springer, Abbie Weight, and Roecale Yazzie for their help with planting and maintaining the common gardens and conducting the caterpillar rearing experiments. This manuscript was improved by comments from two anonymous reviewers. This work was funded by NSF awards to Z. Gompert (DEB 1638768), M. Forister (DEB 1638793), C. A. Buerkle (DEB 1638602), J. A. Fordyce and S. L. Lebeis (1638922) and C. C. Nice (1638773) and by a generous donation from Dr. Gene Miller that funded the Gene Miller Science Garden at USU.

1. Manceau M, Domingues VS, Mallarino R, Hoekstra HE (2011) The developmental role of Agouti in color pattern evolution. *Science* 331(6020):1062–1065.
2. Van't Hof AE, et al. (2016) The industrial melanism mutation in british peppered moths is a transposable element. *Nature* 534(7605):102–105.
3. Zhang L, Mazo-Vargas A, Reed RD (2017) Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *Proceedings of the National Academy of Sciences* 114(40):10707–10712.
4. Villoutreix R, et al. (2020) Large-scale mutation in the evolution of a gene complex for cryptic coloration. *Science* 369(6502):460–466.
5. Nylin S, et al. (2018) Embracing colonizations: a new paradigm for species association dynamics. *Trends in ecology & evolution* 33(1):4–14.
6. Randolph HE, et al. (2021) Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. *Science* 374(6571):1127–1133.
7. Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. *Evolution* pp. 586–608.
8. Wheat CW, et al. (2007) The genetic basis of a plant–insect coevolutionary key innovation. *Proceedings of the National Academy of Sciences* 104(51):20427–20431.
9. Futuyma DJ, Agrawal AA (2009) Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences* 106(43):18054–18061.
10. Fordyce JA (2010) Host shifts and evolutionary radiations of butterflies. *Proceedings of the Royal Society B: Biological Sciences* 277(1701):3735–3743.
11. Edger PP, et al. (2015) The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences* 112(27):8362–8366.
12. Spencer KC (1988) *Chemical Mediation of Coevolution.* (Elsevier).
13. Holzinger F, Frick C, Wink M (1992) Molecular basis for the insensitivity of the Monarch (*Danaus plexippus*) to cardiac glycosides. *FEBS letters* 314(3):477–480.
14. Berenbaum MR, Zangerl AR (1998) Chemical phenotype matching between a plant and its insect herbivore. *Proceedings of the National Academy of Sciences* 95(23):13743–13748.
15. Hawthorne DJ, Via S (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412(6850):904–907.
16. Zhang Z, Ober JA, Kliebenstein DJ (2006) The gene controlling the quantitative trait locus EPITHIOSPECIFIER MODIFIER1 alters glucosinolate hydrolysis and insect resistance in *Arabidopsis*. *The Plant Cell* 18(6):1524–1536.
17. Gompert Z, et al. (2019) Genomic evidence of genetic variation with pleiotropic effects on caterpillar fitness and plant traits in a model legume. *Molecular Ecology* 28(12):2967–2985.
18. Messina FJ, Lish AM, Gompert Z (2021) Disparate genetic variants associated with distinct components of cowpea resistance to the seed beetle *Callosobruchus maculatus*. *Theoretical and Applied Genetics* 134(9):2749–2766.
19. Bai Y, et al. (2022) Natural history–guided omics reveals plant defensive chemistry against leafhopper pests. *Science* 375(6580):eabm2948.
20. Nallu S, et al. (2018) The molecular genetic basis of herbivory between butterflies and their host plants. *Nature Ecology & Evolution* 2(9):1418–1427.
21. Dyer LA, et al. (2018) Modern approaches to study plant–insect interactions in chemical ecology. *Nature Reviews Chemistry* 2(6):50–64.
22. Thompson JN, Burdon JJ (1992) Gene-for-gene coevolution between plants and parasites. *Nature* 360(6400):121–125.
23. Gompert Z, et al. (2014) Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology* 23(18):4555–4573.
24. Chaturvedi S, et al. (2018) The predictability of genomic changes underlying a recent host shift in Melissa blue butterflies. *Molecular Ecology* 27(12):2651–2666.
25. Forister M, et al. (2020) Predicting patch occupancy reveals the complexity of host range expansion. *Science Advances* 6(48):eabc6852.
26. Forister ML, Nice CC, Fordyce JA, Gompert Z (2009) Host range evolution is not driven by the optimization of larval performance: the case of *Lycaeides melissa* (Lepidoptera: Lycaenidae) and the colonization of alfalfa. *Oecologia* 160(3):551–561.
27. Forister ML, Scholl CF (2012) Use of an exotic host plant affects mate choice in an insect herbivore. *The American Naturalist* 179(6):805–810.
28. Gompert Z, et al. (2015) The evolution of novel host use is unlikely to be constrained by trade-offs or a lack of genetic variation. *Molecular Ecology* 24(11):2777–2793.
29. Yoon S, Read Q (2016) Consequences of exotic host use: impacts on Lepidoptera and a test of the ecological trap hypothesis. *Oecologia* 181(4):985–996.
30. Harrison JG, et al. (2018) Deconstruction of a plant-arthropod community reveals influential plant traits with nonlinear effects on arthropod assemblages. *Functional Ecology* 32(5):1317–1328.
31. Harrison JG, et al. (2016) The many dimensions of diet breadth: phytochemical, genetic, behavioral, and physiological perspectives on the interaction between a native herbivore and an exotic host. *PloS ONE* 11(2):e0147971.
32. Forister ML, et al. (2020) Caterpillars on a phytochemical landscape: The case of alfalfa and the melissa blue butterfly. *Ecology and Evolution* 10(10):4362–4374.
33. Zhang Q, Ragnauth C, Greener MJ, Shanahan CM, Roberts RG (2002) The nesprins are giant actin-binding proteins, orthologous to *Drosophila melanogaster* muscle protein MSP-300. *Genomics* 80(5):473–481.
34. Xu L, et al. (2017) Silencing of a lipase maturation factor 2-like gene by wheat-mediated RNAi reduces the survivability and reproductive capacity of the grain aphid, *Sitobion avenae*. *Archives of Insect Biochemistry and Physiology* 95(3):e21392.
35. Zhang SZ, et al. (2020) A novel digestive proteinase lipase member HA in *Bombyx mori* contributes to digestive juice antiviral activity against *B. mori* nucleopolyhedrovirus. *Insects* 11(3):154.
36. Williams CM (1961) The juvenile hormone. II. its role in the endocrine control of molting, pupation, and adult development in the *Cecropia* silkworm. *The Biological Bulletin* 121(3):572–585.
37. Nijhout HF (1998) *Insect Hormones.* (Princeton University Press).
38. Wang H, et al. (2015) Mapping quantitative trait loci associated with soybean resistance to common cutworm and soybean compensatory growth after defoliation using SNP marker-based genome-wide association analysis. *Molecular Breeding* 35(8):1–15.
39. Zhang JM, Huang GQ, Li Y, Zheng Y, Li XB (2014) Cotton photosynthesis-related psak1 protein is involved in plant response to aphid attack. *Molecular Biology Reports* 41(5):3191–3200.
40. Espeland M, et al. (2018) A comprehensive and dated phylogenomic analysis of butterflies. *Current Biology* 28(5):770–778.
41. Thompson JN (1997) Conserving interaction biodiversity in *The Ecological Basis of Conservation.* (Springer), pp. 285–293.
42. Hardy NB, Kaczvinsky C, Bird G, Normark BB (2020) What we don't know about diet-breadth evolution in herbivorous insects. *Annual Review of Ecology, Evolution, and Systematics* 51:103–122.
43. Agrawal AA, Zhang X (2021) The evolution of coevolution in the study of species interactions. *Evolution* 75(7):1594–1606.
44. Walsh B, Lynch M (2018) *Evolution and selection of quantitative traits.* (Oxford University Press).
45. Nozzolillo C, Arnason JT, Campos F, Donskov N, Jurzysta M (1997) Alfalfa leaf saponins and insect resistance. *Journal of Chemical Ecology* 23(4):995–1002.
46. Cai F, et al. (2017) *Medicago truncatula* oleanolic-derived saponins are correlated with caterpillar deterrence. *Journal of Chemical Ecology* 43(7):712–724.
47. Hussain M, et al. (2019) Role of saponins in plant defense against specialist herbivores.

Molecules 24(11):2067.
48. Stuart J (2015) Insect effectors and gene-for-gene interactions with host plants. *Current Opinion in Insect Science* 9:56–61.
49. Strauss SY, Sahli H, Conner JK (2005) Toward a more trait-centered approach to diffuse (co) evolution. *New Phytologist* 165(1):81–90.
50. Hahn PG, Maron JL (2016) A framework for predicting intraspecific variation in plant defense. *Trends in Ecology & Evolution* 31(8):646–656.
51. Pilson D (1992) Aphid distribution and the evolution of goldenrod resistance. *Evolution* 46(5):1358–1372.
52. Forister ML, Gompert Z, Nice CC, Forister GW, Fordyce JA (2011) Ant association facilitates the evolution of diet breadth in a Lycaenid butterfly. *Proceedings of the Royal Society B: Biological Sciences* 278(1711):1539–1547.
53. Carius HJ, Little TJ, Ebert D (2001) Genetic variation in a host-parasite association: potential for coevolution and frequency-dependent selection. *Evolution* 55(6):1136–1145.
54. Råberg L, Sim D, Read AF (2007) Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science* 318(5851):812–814.
55. Gompert Z, et al. (2012) Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* 66(7):2167–2181.
56. Li H, et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
57. McKenna A, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20(9):1297–1303.
58. Shastry V, et al. (2021) Model-based genotype and ancestry estimation for potential hybrids with mixed-ploidy. *Molecular Ecology Resources* 21(5):1434–1451.
59. Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* 196(3-4):483–493.
60. Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. *Ecology* 89(9):2623–2632.
61. Zhou X, Carbonetto P, Stephens M (2013) Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genetics* 9(2):e1003264.
62. Guan Y, Stephens M (2011) Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *The Annals of Applied Statistics* 5(3):1780–1815.
63. Peterson RE, et al. (2019) Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell* 179(3):589–603.
64. Lewontin RC, Rose S, Kamin LJ (1984) *Not in our genes: Biology, ideology, and human nature.* (Pantheon Books).
65. Lucas LK, Nice CC, Gompert Z (2018) Genetic constraints on wing pattern variation in lycaeides butterflies: A case study on mapping complex, multifaceted traits in structured populations. *Molecular ecology resources* 18(4):892–907.
66. Meuwissen TH, Hayes BJ, Goddard M (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4):1819–1829.
67. Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics* 10(6):381–391.
68. Crawford L, Zeng P, Mukherjee S, Zhou X (2017) Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *PLoS Genetics* 13(7):e1006869.
69. Levin DA (1973) The role of trichomes in plant defense. *The Quarterly Review of Biology* 48(1, Part 1):3–15.
70. Hanley ME, Lamont BB, Fairbanks MM, Rafferty CM (2007) Plant structural traits and their role in anti-herbivore defence. *Perspectives in Plant Ecology, Evolution and Systematics* 8(4):157–178.