INTRODUCTION

Organisms interact with members of other species in myriad ways, including competition for resources, predation, parasitism, herbivory, mutualism and pollination. Phenotypic and genetic variation within and among populations can affect the outcome of these interspecific interactions (Bolnick et al., 2002; Crutsinger et al., 2006; Farkas, Mononen, Comeault, Hanski, & Nosil, 2013; Hendry, 2016; Papkou et al., 2019; Thompson, 2013). For example, a genetic polymorphism for cryptic colour pattern affects the probability that...
Timema cristinae stick insects are predated by birds (Nosil, 2004; Nosil et al., 2018), and allelic variation in Daphnia magna and its bacterial microparasite, Pasteuria ramosa, alters infection rates (Cariou, Little, & Ebert, 2001; Luijckx, Ben-Ami, Mouton, Pasquier, & Ebert, 2011; Luijckx, Fienberg, Duneau, & Ebert, 2013). Until now, many of the most successful studies linking intraspecific variation in genotypes to phenotypes to interactions have focused on relatively simple phenotypes, such as crypsis (e.g., Barrett et al., 2019; Gompert et al., 2014; Nosil et al., 2018). Intraspecific variation can also affect the establishment and evolution (or coevolution) of new interactions, including those that form following species introductions (e.g., Cox, 2004; Lankau, 2012; Mandeville, Parchman, & Thompson, 2017; Strauss, Lau, & Carroll, 2006). We can expect that the establishment of novel interactions will often involve complex phenotypes, including survival and development in particular habitats or on specific resources. Here we utilize recent advances in statistical modelling to investigate such complex phenotypes in the context of a model legume and a specialist herbivore.

Interactions between plants and herbivorous insects have received considerable scientific attention due to their ubiquity (Forister et al., 2015), agricultural relevance (Schoonhoven, Loon, & Dicke, 2010; Via, 1990), and hypothesized contribution to the extreme biodiversity of these taxonomic groups (via coevolutionary diversification; Braga, Guimaraes, Wheat, Soren, & Janz, 2018; Edger et al., 2015; Ehrlich & Raven, 1964; Fordyce, 2010; Mitter, Farrell, & Wiegmann, 1988). These interactions are often affected by genetic variation within species, including variation in plant resistance to insects, and for insect acceptance of, and performance on, potential host plants (e.g., Berenbaum & Zangerl, 1998; Dambroski et al., 2005; Gompert et al., 2015; Mitchell, Brennen, Graham, & Karley, 2016; Nouhaud et al., 2018; Ordas et al., 2009; Rausher & Simms, 1989; Schoonhoven et al., 2010; Spencer, 1988; Stowe, 1998; Via, 1990). Progress in explaining this variation has been made by identifying specific phytochemicals responsible for resistance to insects (e.g., furanocoumarins and glucosinolates), as well as the insect genes and pathways that detoxify these compounds (e.g., cytochrome P450 enzymes, nitrile specifier protein, etc.; Berenbaum & Feeny, 1981; Li, Schuler, & Berenbaum, 2003; Schoonhoven et al., 2010; Wen, Rupasinghe, Niu, Berenbaum, & Schuler, 2006; Wheat et al., 2007). Genomic and metabolomic approaches have begun to provide a more complete view of how within-species variation affects plant-insect interactions (e.g., Harrison et al., 2018; Nallu et al., 2018). As an example, a recent study of intraspecific variation across 770 traits (including 753 chemical features) in alfalfa showed that among-plant variation in insect herbivore communities was best explained by interactions among suites of plant traits with some traits having nonlinear effects on community composition (Harrison et al., 2018). Such findings highlight the need for quantitative, genome-, phenome- and metabolome-scale analyses of the ecological and evolutionary consequences of intraspecific variation in plant-insect systems.

In this study, we combined polygenic genome-wide association (GWA) mapping and genomic prediction with a caterpillar rearing experiment to assess the effect(s) of genome-wide genetic variation in the plant Medicago truncatula on the herbivore Lycaenidae melissa. Medicago truncatula is a model legume with extensive genomic resources, and L. melissa is a legume-specialist butterfly that has been known to utilize (lay eggs and develop on) novel hosts with which it has no evolutionary history (Austin & Leary, 2008). In contrast to standard GWA methods, polygenic GWA methods relate trait variation to hundreds of thousands or millions of genetic variants simultaneously (Zhou, Carbonetto, & Stephens, 2013). Polygenic GWA mapping methods can identify major-effect mutations when they exist, but also generate robust estimates of trait genetic variances (and heritabilities) and genetic covariances regardless of whether few or many genes of large or small effect underlie phenotypic variation (e.g., Moser et al., 2015; Zeng & Zhou, 2017; Zhou et al., 2013). Consequently, the utility of these methods does not depend on a specific, a priori unknown, genetic architecture. These methods can also parse the relative contributions of genetic variants with measurable (modest to large) versus near-infinitesimal effects to trait genetic variances and covariances (e.g., Lucas, Nice, & Gompert, 2018). Such information cannot readily be obtained from standard GWA/QTL methods, or from classic quantitative genetic approaches (e.g., analyses based on the resemblance among relatives or line crosses; Lynch & Walsh, 1998).

The overall aim of this study was to advance our understanding of the genetic architecture of plant-insect interactions. In particular, we wanted to determine how many genetic regions (many or few) in the host plant are relevant to caterpillar development, and whether those regions have simple, pleiotropic or epistatic effects. The complexity/pleiotropy of the genotype-phenotype map determines the speed and trajectory of response to natural selection (Hansen & Houle, 2008; Orr, 2000; Wagner & Zhang, 2011; Walsh & Blows, 2009), which (in the case of plant–insect interactions) is necessary for understanding the evolution of plant defence to herbivores. Herein, we addressed the following specific questions: (a) How much of the variation in L. melissa growth and survival can be explained by genetic variation in M. truncatula?, (b) do genetic loci that affect a set of measured plant traits (some putatively associated with plant vigour or defence) have pleiotropic effects on caterpillar performance?, and (c) how well do the effects of M. truncatula alleles on the measured plant traits explain their effects on caterpillar performance? Thus, we quantify the direct effect of M. truncatula genetic variation on caterpillar performance, and its effect through a set of plant traits. We suggest that this combination of approaches has the potential to (a) provide a more mechanistic understanding of this plant–insect interaction by connecting genetic patterns with plant traits, and (b) discover previously unhypothesized sources of variation in caterpillar performance by identifying alleles associated with caterpillar performance that are not associated with any of the plant traits we measured. Importantly, the methods and approaches we use allow us to generate statistical and functional information about the genetic basis of this interaction even if it is polygenic (as noted above; see Methods and Results for additional details).
TABLE 1 Plant traits along with our predictions about their primary functional roles and relationships with caterpillar performance

| Traits            | Primary putative function | Predicted relationship with caterpillar performance |
|-------------------|---------------------------|-----------------------------------------------------|
| Leaf length       | Growth                    | /                                                   |
| Leaf width        | Growth                    | /                                                   |
| Leaf area         | Growth                    | /                                                   |
| Leaf weight       | Growth                    | /                                                   |
| SLA               | Defence                   | \                                                   |
| Trichome den.     | Defence                   | \                                                   |
| Leaf tough.       | Defence                   | \                                                   |
| Plant height      | Growth                    | / or \                                              |
| IR features       | Growth or defence         | / or \                                              |

Note: We present simplified predictions to guide interpretation, but are aware that the traits potentially have multifaceted relationships to growth and defence. / denotes a positive correlation with caterpillar performance, whereas \ denotes a negative relationship with caterpillar performance. Our classification of SLA is based on its general association with mechanical properties of leaves, including work to shear, tear and penetrate (reviewed in Hanley et al., 2007). All 19 IR chemical features are treated together here, and thus we predict that they include a mixture of features associated with vigour (/) and defence (\). Leaf shape is not included in the table, as its putative function and effects are not known.

2 | MATERIALS AND METHODS

2.1 | Study system

*Lycaeides melissa* butterflies are found throughout western North America where they feed on various legume hosts, particularly from the genera *Astragalus* and *Lupinus* (Scott, 1986). *Medicago sativa* (alfalfa) is a legume native to Eurasia that was introduced to North America ~250 years ago as a forage crop (Michaud, Lehman, & Rumbaugh, 1988). Since then, *L. melissa* has repeatedly colonized *M. sativa*, and numerous *L. melissa* populations now use this plant as their primary host, especially where *M. sativa* has escaped from cultivation along roadsides and trails (Chaturvedi et al., 2018). In the present study, we focus on a close relative of *M. sativa*, the model plant *Medicago truncatula*. *Medicago truncatula* occurs throughout the Mediterranean basin in Europe and is cultivated in Australia (Choi, Kim, et al., 2004a; Choi, Mun, et al., 2004b). Because of its modest genome size (~500 million base pairs), simple diploid genetics, and short generation time (~10 weeks), *M. truncatula* has been developed as the model species for legumes (Young et al., 2011; Young & Udvardi, 2009). Resources for this species include a high-quality reference genome and hundreds of fully sequenced, inbred lines derived from natural accessions (Stanton-Geddes et al., 2013; Young et al., 2011). *Medicago truncatula* is not found in North America and thus is not available as a host for *L. melissa* (i.e., it is not part of *L. melissa’s* realized niche). However, *M. truncatula* is a potential host for other *Lycaeides* species in Eurasia where most of the biodiversity in this butterfly genus is found (North American *Lycaeides* are descended from Eurasian ancestors that came across the Bering land bridge about two million years ago; Gompert, Fordyce, Forister, & Nice, 2008; Vila et al., 2011).

2.2 | Plant propagation and trait measurements

We obtained seeds from 100 *M. truncatula* lines, which are part of the *Medicago* HapMap project (http://www.medicagohapmap.org). Seeds (i.e., germplasm) were obtained from INRA-Montpellier, and from the USDA Agricultural Research Station at Washington State University (Table S1). Each line was derived from a natural accession, but has since been inbred to near complete homozygosity. Whole genome sequences are available for each line (Branca et al., 2011; Stanton-Geddes et al., 2013), and the lines have been used in other GWA mapping studies (GWAS), including GWAS on biomass, drought-related traits, plant defences, flowering time, and nodulation (e.g., Kang, Sakiroglu, & Krom, 2015; Stanton-Geddes et al., 2013).

We planted five replicate pots with seeds from each of the 100 *M. truncatula* lines on 4 and 5 May 2017 (see “Planting and tending *Medicago truncatula*” in the online supplemental material [OSM] for additional details). *Medicago truncatula* plants were grown in a greenhouse under ambient light (~14–15 hr of daylight) at approximately 18–27°C (with variable humidity), and were watered daily or every other day as needed. We thinned the *M. truncatula* seedlings on 26 May (i.e., after germination was complete) to ensure that no pot had more than two plants. This was done to minimize competition among plants, while still providing sufficient plant biomass for the caterpillar rearing experiments. A few plant lines had low germination rates and were dropped from the experiment leaving 94 lines, each with five replicate pots (N = 470).

We measured a series of morphological traits potentially associated with plant vigour or resistance to insects (e.g., putative structural plant defences; Table 1; Hanley, Lamont, Fairbanks, & Rafferty, 2007; Levin, 1973; Malishev & Sanson, 2015). First, 20 days after planting, we measured leaf size (length, width and area), leaf shape (length/width), trichome density, dry leaf weight and specific leaf area (SLA) for each plant line and replicate (pot) (we haphazardly selected one of the two plants in each pot for taking measurements). We chose the second true leaf for these measurements (leaf 1 from branch B0, see figures 1 and 2 from Moreau, 2006). We measured the width (at the widest point) and length (along the midvein) of the middle leaflet with calipers (each leaf comprises three leaflets; measurements were taken to the nearest 1 mm). Next, we calculated leaf area (length × width) and shape (length/width) from these measurements. We then counted the number of trichomes in a 2.5 mm diameter circle directly adjacent to the midvein under a stereoscope (35× magnification). The three leaflets from each plant were then placed in a coin envelope in a bin with desiccant. The dry weight of the middle leaflet from each of these leaves was measured on a Mettler Toledo XPE105 analytical microbalance (Mettler Toledo) to the nearest 0.01 mg. Leaf area and dry weight were used to calculate SLA.
(SLA is the ratio of leaf area to dry mass and is often correlated with leaf mechanical properties, such as work to tear, shear or punch; Hanley et al., 2007).

We measured plant height, from the cotyledons to the tip of the longest branch, 31 days after planting (again, we haphazardly selected one of the two plants in each pot for taking this measurement). Leaf toughness was measured 33 days after planting using a penetrometer. We selected the main leaf from the second primary branch for this assay. The force required to penetrate each of the three leaflets along the midvein was recorded. We took the mean of these three measures as a metric of leaf toughness.

Plant chemistry was quantified with attenuated total reflectance infrared (ATR-IR) spectroscopy. ATR-IR spectroscopy constitutes a quick, cost-effective method to analyze a range of organic chemical compounds in plant and animal tissues. Although the absorbance is directly related to the concentration of specific chemical signatures, there is not a simple one-to-one relationship between IR spectral patterns and specific chemical compounds of interest. Moreover, spectral features are the summation of similar overlapping IR transitions, representative of various compounds within a tissue. Consequently, IR data are often combined with more specific compositional analyses (e.g., HPLC-MS). The combined data can be used to construct a multivariate model linking IR spectral data to chemical compounds (e.g., Costa, Lang, Almeida, Castilho, & Poorter, 2018; Foley et al., 1998; Ramirez et al., 2015). This was not our goal here. We instead used IR spectral features as anonymous chemical markers (akin to anonymous molecular markers for genetic analyses), which could be connected to the presence of specific molecules in future work using compositional methods such as liquid chromatography–mass spectrometry.

Infrared spectra were collected using a Thermo Nicolet 6700 FTIR (a high-resolution instrument with a diamond crystal ATR), which was used to scan 4,000–600 cm⁻¹ of the infrared spectrum. Leaves were placed in direct contact with the diamond crystal, and the average of 32 scans was recorded for each leaf surface with 4 cm⁻¹ resolution. A Norris-Williams second derivative spectrum was calculated for each transmittance measurement using five point smoothing and a gap size of five segments (absorbance is directly proportional to concentration [Beer’s Law], and absorbance = -\log[transmittance]). We focused on the subset of IR features between ~750 and 1,100 cm⁻¹ and with >10% of the phenotypic variation partitioned among plant lines (see Figure S1).

### 2.3 | Caterpillar husbandry and performance assays

We obtained neonate *L. melissa* caterpillars for larval performance assays on the *M. truncatula* accessions. First, 26 female *L. melissa* butterflies were collected on 5 June 2017 from a site along the Bonneville shoreline trail in northern Utah, USA (41.725°N, 111.794°W, 1,513 m elevation). As in previous work (e.g., Forister et al., 2013; Gompert et al., 2015), these butterflies were caged individually in plastic oviposition chambers along with a few sprigs of their host plant (*Medicago sativa*). After 48 hr, *L. melissa* eggs were collected from the host-plant material and placed in unvented Petri dishes in a Percival incubator (model No. 136VL; 27°C; 14 hr light:10 hr dark) until they hatched.

Caterpillars began to emerge on 9 June, and were then placed in individual unvented Petri dishes with a leaf from one of the 94 *M. truncatula* accessions (i.e., on one of the 94 plant lines). Caterpillars were randomly assigned to plant lines so that caterpillar family was not confounded with plant line. We inspected caterpillars daily, adding new leaf material from the same plant line as required (as in Gompert et al., 2015). We rotated the replicate/pot used for each plant line each day. Thus, caterpillars only ate leaves from a single plant line (genotype), but fed on all five replicate pots. Caterpillars were maintained in a Percival incubator at 27°C with 14 hr days (10 hr of dark). We reared 486 caterpillars total (~5 per plant line) from the 19 females that laid viable eggs (mean = 25.6 caterpillars per family, SD = 20.5 caterpillars). We checked all caterpillars daily for survival and recorded survival to pupation and survival to eclosion as adults. As an additional metric of performance, we measured 8- and 16-day caterpillar weight (*L. melissa* caterpillars generally spend 20 to 30 days as larvae) on a Mettler Toledo XPE105 analytical microbalance (Mettler Toledo; weights were recorded to the nearest 0.01 mg). Weight and lifetime fecundity are highly correlated in *L. melissa* (Forister, Nice, Fordyce, & Gompert, 2009).

#### 2.4 | Variance partitioning

Our analyses focussed on the nine plant morphological traits (leaf length, leaf width, leaf area, leaf shape, leaf dry weight, SLA, trichome density, leaf toughness and plant height), 19 IR traits (i.e., anonymous chemical features), and four caterpillar performance traits (weight at 8 days, weight at 16 days, survival to pupation, and survival to eclosion; survival is a binary trait; Table 1). Prior to genetic mapping and genomic prediction, we first quantified the proportion of trait variation found among plant lines (i.e., genotypes) for each of these 32 traits. We worked with replicated, inbred lines, and these were estimates of the broad-sense heritability for each of the traits (with respect to plant not caterpillar genotypes; because caterpillars fed across plants of a genotype, these estimates are upper bounds for the broad-sense heritabilities of the caterpillar performance traits).

We estimated the among-line variance for each trait by fitting linear mixed-effect models via restricted maximum likelihood (REML) (aside from the grand mean, no fixed effects were included in the models). This was done with the `lmer` function in the `lme4` R package (package version 1.1.19, R version 3.4.4; Bates, Mächler, Bolker, & Walker, 2015). We then tested the null hypothesis that the among-line variance was 0 using an exact restricted likelihood ratio test, which was based on 10,000 simulated values to approximate the null distribution (Crainiceanu & Ruppert, 2004; Greven, Crainiceanu, Küchenhoff, & Peters, 2008). This was done with the `exactrlrt` function in the `RLRsim` package in R (version 3.1.3; Scheipl, Greven, & Kuechenhoff, 2008).
We similarly quantified the variance in caterpillar performance traits (weight and survival) among larvae from different families (here a family is defined as the set of larvae from a single female butterfly, regardless of their unknown paternity). We fit and contrasted models with caterpillar family (reduced model) or caterpillar family and plant line (full model) to ask whether plant line explained additional variation in the caterpillar performance traits not explained by caterpillar family. We fit and compared these mixed-effect models using REML (with \( \text{Lmer} \)) and exact restricted likelihood ratio tests (with \( \text{exactREML} \)) as described above.

### 2.5 | Medicago truncatula genomic data

Whole-genome SNP data for the *M. truncatula* accessions were obtained from the *M. truncatula* HapMap project ([http://www.medicagohmap.org/](http://www.medicagohmap.org/); version Mt4.01; Stanton-Geddes et al., 2013). These data comprised 40 million SNPs, which were mapped to the *M. truncatula* reference genome version 4.0 (we used the quality-filtered SNP \( \text{bcf} \) files; Young et al., 2011). We applied additional quality filters to these data with \( \text{vcftools} \) (version 0.1.15; Danecek et al., 2011) such that we only retained biallelic SNPs with minor allele frequencies \( \geq 0.01 \), and with a minimum sequencing depth of 2× per individual, no more than 20% missing data (across the 94 lines analyzed in this study), and a phred-scaled quality score of \( \geq 30 \). We only considered SNPs mapped to the eight *M. truncatula* chromosomes. Approximately 13 million SNPs passed these filters. We then used \( \text{PLINK} \) (version 1.09; Purcell et al., 2007) to remove redundant SNPs, that is SNPs that were in very high linkage disequilibrium (LD) with each other. Specifically, using the \( \text{INDEP‐PAIRWISE} \) command, one of each pair of high-LD SNPs, defined as \( r^2 \geq 0.8 \) in a 10 kilobase (kb) window, was pruned (dropping redundant SNPs, i.e., very high LD SNPs, was necessary for more rapid convergence of the Markov chain Monte Carlo (MCMC) algorithm used in the polygenic GWA analysis described below). After this step, we retained 5,648,722 SNPs for downstream analyses.

The *M. truncatula* HapMap data set included SNP genotype calls and relative genotype likelihoods generated by \( \text{GATK} \) (McKenna et al., 2010). Rather than use the raw genotype calls (which ignore uncertainty in genotypes and information from population allele frequencies), we used an empirical Bayesian approach to obtain estimates of genotypes based on the genotype likelihoods and a prior defined by the allele frequencies at each locus. As in previous studies (e.g., Gompert et al., 2015), we first used an expectation-maximization algorithm to obtain maximum likelihood estimates of the allele frequencies for each SNP. This was done with the computer program \( \text{ESTFEM} \) (in Dryad repository, [https://doi.org/10.5061/dryad.nq67q](https://doi.org/10.5061/dryad.nq67q); Riesch et al., 2017; Soria-Carrasco et al., 2014). This program implements the EM algorithm from Li et al. (2009) and provides allele frequency estimates that account for genotype uncertainty. Prior probabilities for each genotype were then specified based on the allele frequencies, such that \( \Pr(g_j|p) \sim \text{binomial}(p,n=2) \), where \( g_j \) denotes the genotype at locus \( i \) for individual \( j \), and \( p \) denotes the nonreference allele frequency.

Next, we computed the posterior probability of each genotype according to Bayes theorem, and obtained point estimates (posterior means) for genotypes \( g_j = \sum_{k=0,1,2} k L(g_j = k) \Pr(g_j = k|p) \), where \( L(g_j = k) \) is the relative genotype likelihood based on the sequence data and associated quality scores. These genotype estimates take on values between 0 (reference-allele homozygote) and 2 (nonreference-allele homozygote), but are not constrained to be integer values (allowing for noninteger genotypes captures some uncertainty in the genotype estimates, and is a common approach, e.g., when working with imputed genotypes).

### 2.6 | Genome-wide association mapping and genomic prediction

We fit Bayesian sparse linear mixed models (BSLMMs; Zhou et al., 2013) with \( \text{GEMMA} \) (version 0.94.1) to quantify the contribution of *M. truncatula* (i.e., plant) genetic variation to phenotypic variation in the plant traits and *L. melissae* caterpillar performance. Unlike traditional GWA mapping methods, this polygenic GWA method fits a single model with all SNPs simultaneously and thus mostly avoids issues related to testing large numbers of null hypotheses. In particular, trait values are modelled as a function of a polygenic term and a vector of the (possible) measurable effects (associations) of each SNP on the trait \( \phi \) (Zhou et al., 2013). Variable selection is used to estimate the SNP effects; SNPs can be assigned an effect of 0 (not in the model) or a nonzero effect (in the model; Guan & Stephens, 2011). A MCMC algorithm is used to infer the posterior inclusion probability (PIP) for each SNP, that is, the probability that each SNP has a nonzero effect. The polygenic term defines an individual’s expected deviation from the grand phenotypic mean based on all of the SNPs (this assumes all SNPs have near-infinitesimal effects on the trait). It accounts for phenotypic covariances among individuals caused by their relatedness or overall genetic similarity (i.e., observed kinship; Zhou et al., 2013). The kinship matrix also serves to control for population structure and relatedness when estimating the effects of individual SNPs \( \phi \) along with their PIPs. Likewise, SNPs in 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 is captured by the PIPs.

The hierarchical structure of the model provides a way to estimate additional parameters that describe aspects of a trait’s genetic architecture (Guan & Stephens, 2011; Lucas, Nice, & Gompert, 2018; Zhou et al., 2013). These include the proportion of the phenotypic variance explained (PVE) by additive genetic effects (this includes \( \beta \) and the polygenic term, and should approach the narrow-sense heritability), the proportion of the PVE due to SNPs with measurable effects or associations (this is called PGE and is based only on \( \beta \)), and the number of SNPs with measurable associations \( n \sim \gamma \). All of these metrics integrate (via MCMC) over uncertainty in the effects of individual SNPs \( \phi \) along with their PIPs. Likewise, BSLMMs can be used to obtain genomic estimated breeding values (GEBVs), that is, the expected trait value for an individual from the additive effects of their genes as captured by both \( \beta \) and the polygenic term (denoted \( u \)) (Lucas et al., 2018). Most other genomic
prediction methods provide GEBVs based solely on a polygenic term (e.g., Meuwissen, Hayes, & Goddard, 2001; Hayes, Bowman, Chamberlain, & Goddard, 2009; Ober et al., 2012).

We fit BSLMMs for each of 32 traits using *gemma* (version 0.94.1; Zhou et al., 2013) with 15 MCMC chains each with a 500,000 iteration burn-in followed by 2 million sampling iterations with a thinning interval of 20. Survival data were analyzed on the observed (binary) scale rather than on a continuous liability/risk scale with a probit model. We did this following the recommendation of Zhou et al. (2013), which was based on the robustness of the standard BSLMM to model misspecification (i.e., to departures from a Gaussian error distribution) and on the poorer convergence/mixing properties of the probit model. Thus, our estimates of all model parameters, including PVE, were on the observed scale (transformation of survival PVE estimates to the liability following Appendix S3 in Zhou et al. (2013) resulted in higher estimates of PVE; results not shown). GEBVs were obtained using the *-predict 1* option, with predictions averaged over the 15 MCMC chains. GEBVs were used to estimate genetic correlations among traits (i.e., a standardized G-matrix). As a guard against statistical artifacts, we fit BSLMMs to 12 pseudo (randomized)-data sets derived from the caterpillar data (while these methods have been assessed in detail elsewhere, e.g., Zhou et al., 2013; Gompert, Egan, Barrett, Feder, & Nosil, 2017, we were particularly concerned that the low number of survivors and binary data for survival could lead to spurious association; for details, see “BSLMMs fit to randomized data” in the OSM).

We focussed on genetic correlations estimated from the GEBVs described above, as they incorporate SNPs with measurable (captured by β) and near-infinitesimal effects (through the polygenic term, u), and because they account for uncertainty in SNP-trait associations, especially in cases where multiple SNPs might exhibit LD with the same causal variant. However, additional information on the extent to which genetic correlations arise from major effect loci (vs. polygene effects) can be gained by comparing these genetic correlations with other estimates of trait genetic correlations. To this end, we also calculated genetic correlations from (a) GEBVs that only included SNPs with measurable effects/associations (i.e., GEBVs from β alone), and (b) best linear unbiased predictors (BLUPs) of breeding values (we obtained these from the REMI fit for among-line variances using the *lm er* function from *lme4* version 1.1.19; Bates et al., 2015). These three sets of genetic correlations could differ if SNPs with measurable versus near-infinitesimal effects on traits differ in terms of pleiotropy or patterns of LD (a) only includes the former, whereas (b) is based on the latter. We also inferred expected genetic correlations from model-averaged estimates of β (SNP effects/associations weighted by their PIPs) and SNP allele frequencies following equation 3 in Gardner and Latta (2007). These expected genetic correlations ignore LD, and thus will only be nonzero when the same SNPs are associated with multiple traits. Consequently, analysis of expected genetic correlations could help distinguish true pleiotropy from cases where LD among distinct causal mutations that affect different traits as causes of observed genetic correlations (but see the Discussion for caveats). Lastly, we estimated correlations in the model-averaged SNP effects between pairs of traits (this ignores LD and SNP allele frequencies) to simply ask whether SNPs had similar effects on (associations with) multiple traits. We did this at the SNP level, and for windows of 10, 100, 1,000 or 10,000 adjacent SNPs (for these we took the mean model-averaged effect estimates for the set of SNPs in a window).

### 2.7 Connecting plant trait genetics with caterpillar performance

Genetic covariances (correlations) among plant and caterpillar traits (as captured by the G-matrix) can provide evidence of a shared genetic basis for these traits. However, these treat pairs of traits independently and do not formally quantify the total contribution of alleles affecting (or associated with) the measured plant traits to the alleles affecting (or associated with) caterpillar performance. Thus, we next assessed the extent to which we could explain variation in the caterpillar performance GEBVs based on the GEBVs for the plant morphometry and chemistry traits, as well as which plant trait GEBVs were most important for this (here we used GEBVs from *gemma* based on β and the polygenic term). In other words, we wanted to ascertain how well we could explain (or predict) the caterpillar performance GEBVs (that is, the expected performance trait values based on plant genetics) from the subset of genetic variants associated with phenotyped plant traits (as captured by the plant trait GEBVs, and thus weighted by their effects on the plant traits). High explanatory (or predictive) power would imply that most of the *M. truncatula* genetic variants affecting caterpillar performance either had pleiotropic effects on some of the plant traits we measured or were tightly linked to genetic variants that affected these traits. We hoped to identify specific plant traits that share a common genetic basis with (and thus potential causal link to) caterpillar performance. We used two complementary approaches to answer this question: (i) multiple regression with Bayesian model averaging, and (ii) random forest regression. A key distinction between these methods is whether they assume linear (multiple regression) or nonlinear (random forest regression) relationships between predictors and response variables. Note that for each plant and caterpillar trait, there was a single GEBV estimate per plant line, and thus the sample size for these analyses was $N = 94$ plant lines.

We used multiple regression with Bayesian model averaging to identify the subset of predictors (plant GEBVs) that best explained variation in caterpillar performance GEBVs, while accounting for uncertainty in the effects of each covariate including which covariates have nonzero effects. The multiple regression models were fit with the *bms* R package (package version 0.3.4, R version 3.4.2; Zeugner & Feldkircher, 2015). Zellner’s g-prior was used for the regression coefficients with $g = N$, where $N$ is the number of observations ($N = 94$; Zellner, 1986), and a uniform prior was used for the different models (i.e., sets of covariates with nonzero effects; Zeugner & Feldkircher, 2015). Parameter estimates were obtained using MCMC with a 5,000 iteration burnin and 100,000 sampling iterations, and using the birth-death sampler for
exploring model space. We then used 10-fold cross-validation to assess the predictive power of these models (that is, the power of the model to explain observations not used in fitting the model). Predictive power necessarily averages over uncertainty in covariate effects (including which covariates have nonzero effects), and was measured as the Pearson correlation (and squared Pearson correlation) between the observed and predicted caterpillar performance GEBVs. As a simpler metric of explanatory power (not predictive power), we estimated the coefficient of determination ($r^2$) from a standard linear model that included only the subset of predictors (i.e., plant trait GEBVs) with posterior inclusion probabilities (PIPs) greater than 0.5 in the Bayesian model averaging analysis (importantly, here the same data were used to fit the model and assess its explanatory power). This was done with the `lm` function in R.

The random forest regression algorithm was similarly used to determine the influence of the plant trait GEBVs on the caterpillar performance GEBVs, while allowing for interactions among variables and nonlinear relationships between the independent and dependent variables (Breiman, 2001). Random forest creates multiple regression trees and then outputs the importance of each predictor. The number of trees created was left at the default of 500, after determining that changing the number of trees from this number did not significantly reduce error. The number of variables randomly sampled at each split ($m_{try}$) and the number of terminal nodes ($node\_size$) were chosen to minimize the out-of-bag (OOB) error by manually varying these parameters from one to 20 (all possible combinations were considered). To determine variable importance, the predictor of interest was varied and the percent change mean-squared error (%MSE) in predicting the OOB data was determined for each. Those with the greatest effect on %MSE are the most important predictor variables. Random Forest was run using the `randomforest` package (version 4.6–12) in R (Liaw & Wiener, 2002). Random forest regression was run separately with GEBVs for each of the caterpillar performance traits as the response and the GEBVs for plant traits as predictors.

### RESULTS

#### 3.1 Variation in plant traits and caterpillar performance

We documented substantial phenotypic variation for all 32 traits assayed (e.g., Figure 1a,c). Phenotypic correlations among traits were evident, particularly among leaf morphology traits (some of which are functions of each other; Figure 1b) and among some IR chemical traits (Figure S2). Caterpillar survival rates were initially high, with only nine of the 486 caterpillars (1.9%) dying within the first eight days; the mean survival time was 22.3 days (excluding caterpillars that pupated; Figure 1d). But most caterpillars failed to pupate (448, or 92.2%), such that high mortality rates were observed between 20 and 30 days of larval development. Of the 38 caterpillars...
that did pupate, 11 eclosed as adults (29%) (several of the adults were deformed). Mean caterpillar weight at 8 and 16 days was 5.1 mg (SD = 2.5 mg, min. = 0.04 mg, max. = 12.9 mg) and 17.7 mg (SD = 7.7 mg, min. = 3.02 mg, max. = 82.7 mg), respectively.

The 32 traits exhibited significant among-line variation, with the possible exception of survival to eclosion as adults (Table 2). The proportion of variation among lines ranged from 0.15 (SLA) to 0.59 (plant height) for the plant morphology traits, from 0.09 to 0.36 for the plant IR traits, and from 0.05 (survival to eclosion) to 0.41 (16-day weight) for the caterpillar performance traits (Figure S3). With the exception of survival to eclosion (restricted likelihood ratio test [RLRT], \( p = 0.059 \)), the null model could be confidently rejected for all traits (RLRT, all \( p < 0.05 \), most \( < 0.001 \); Table 2).

Caterpillar family explained 11% of the variation in weight at 8 days, but no more than 2% of the variation in the other caterpillar performance traits (Table S2). Again with the exception of survival to eclosion, models that included random effects for both caterpillar family and plant line were preferred over reduced models with only caterpillar family (all \( p < 0.01 \)). And with the exception of 8-day weight, plant line explained ~10 times more of the variation in caterpillar performance than caterpillar family did (Table 3).

### 3.2 Genetic architecture of plant and caterpillar traits

The *Medicago truncatula* SNP data explained a modest to substantial proportion of trait variation (Table S3; Figure 2) (estimates of PVE and REML estimates of broad-sense heritability were similar, as were breeding values inferred from these two methods; Tables 2 and S3; Figure S4). On average, *M. truncatula* genetic variation accounted for a greater proportion of the variation in plant morphology traits (mean PVE = 0.40) than in IR traits (mean PVE = 0.17) or caterpillar performance (mean PVE = 0.24). However, *M. truncatula* genetics explained a particularly large amount of the variation in *Lycaeides melissa* caterpillar 16-day weight (PVE = 0.41, 90% equal-tail probability intervals [ETPIs] = 0.34–0.49; this trait also exhibited high among-line variance, Table 2). Estimates of PVE were generally precise, such that the average width of the 90% ETPIs for these parameters (mean across traits) was 0.13 (range = 0.11–0.15). In contrast, our estimates of the number of genetic loci with measurable effects on each trait (\( n_{\gamma} \)) and of the proportion of the PVE explained by those loci (PGE) were less certain; in particular, the average width of the 90% ETPIs for \( n_{\gamma} \) and PGE (a proportion) were 153.7 loci and 0.82, respectively (Table S3). Thus, uncertainty in these parameter estimates blurs differences in genetic architectures among traits suggested by the differences in parameter point estimates (compare Figure 2 with Table S3). Genetic architecture parameter estimates for permuted (randomized) caterpillar performance data differed markedly from those for the actual data, most notably in terms of PVE. Whereas permutations of the survival to eclosion data did sometimes give modest estimates of

### Table 2 REML estimates for each caterpillar performance trait of the proportion of phenotypic variation found among *Medicago truncatula* lines (“Prop. var.”). Test statistics (LR = likelihood ratios) and \( p \)-values from the null hypothesis test of no line effect are reported

| Traits            | Prop. var. | LR      | \( p \)  |
|--------------------|------------|---------|---------|
| Leaf length        | 0.43       | 115.87  | <0.001  |
| Leaf width         | 0.49       | 144.78  | <0.001  |
| Leaf area          | 0.49       | 147.90  | <0.001  |
| Leaf shape         | 0.21       | 32.23   | <0.001  |
| Leaf weight        | 0.41       | 102.49  | <0.001  |
| SLA                | 0.15       | 17.12   | <0.001  |
| Trichome den.      | 0.49       | 151.82  | <0.001  |
| Leaf tough.        | 0.34       | 69.95   | <0.001  |
| Plant height       | 0.59       | 218.92  | <0.001  |
| IR 1.104.64        | 0.13       | 12.79   | <0.001  |
| IR 1.085.1         | 0.15       | 16.20   | <0.001  |
| IR 1.072.19        | 0.10       | 7.01    | 0.004   |
| IR 1.039.74        | 0.16       | 17.46   | <0.001  |
| IR 1.024.17        | 0.10       | 8.48    | 0.001   |
| IR 1.010.93        | 0.17       | 20.27   | <0.001  |
| IR 998.34          | 0.29       | 54.85   | <0.001  |
| IR 985.1           | 0.12       | 10.78   | <0.001  |
| IR 944.37          | 0.10       | 7.46    | 0.003   |
| IR 937.09          | 0.23       | 35.98   | <0.001  |
| IR 929.14          | 0.15       | 16.76   | <0.001  |
| IR 918.54          | 0.12       | 11.63   | <0.001  |
| IR 892.38          | 0.13       | 11.99   | <0.001  |
| IR 855.96          | 0.13       | 13.19   | <0.001  |
| IR 840.07          | 0.23       | 36.36   | <0.001  |
| IR 830.13          | 0.36       | 80.73   | <0.001  |
| IR 818.21          | 0.24       | 40.84   | <0.001  |
| IR 793.71          | 0.14       | 15.10   | <0.001  |
| IR 757.28          | 0.09       | 6.05    | 0.007   |
| Wgt. 8 days        | 0.07       | 3.99    | 0.019   |
| Wgt. 16 days       | 0.41       | 93.69   | <0.001  |
| Surv. pupation     | 0.24       | 40.33   | <0.001  |
| Surv. eclosion     | 0.05       | 2.39    | 0.059   |

### Table 3 REML estimates for each caterpillar performance trait of the proportion of phenotypic variation found among *Lycaeides melissa* families and *Medicago truncatula* lines (“Prop. var.”). Test statistics (LR = likelihood ratios) and \( p \)-values from the comparison of a reduced model with only *L. melissa* family effects versus the full model with random effects for caterpillar family and plant line are reported

| Traits            | Prop. var. *L. melissa* | Prop. var. *M. truncatula* | LR     | \( p \)  |
|--------------------|--------------------------|-----------------------------|--------|---------|
| Wgt. 8 days        | 0.11                     | 0.09                        | 8.02   | 0.002   |
| Wgt. 16 days       | 0.01                     | 0.42                        | 98.18  | <0.001  |
| Surv. pupation     | 0.03                     | 0.26                        | 48.36  | <0.001  |
| Surv. eclosion     | <0.01                    | 0.05                        | 1.91   | 0.079   |
PVE (the maximum was 0.12, 90% ETPIs = 0.06–0.19), these were still lower than the PVE estimate for the least heritable trait, namely survival to eclosion (PVE = 0.15, 90% ETPIs = 0.09–0.22), and most PVE estimates from permuted data were less than 0.05 (Figure S5).

Consistent with the high (but uncertain) estimates of n-γ for most traits, many SNPs had small but nonzero posterior inclusion probabilities (PIPs) in the BSLMMs (Figure S6). In other words, we were better able to detect than confidently isolate and localize the
effects of individual genetic loci on the traits. There were a few exceptions to this pattern, most notably plant height and survival to eclosion. For plant height, one SNP each on chromosomes 5 and 7 had very high PIPs, ~1.0 (Figure S7). Two nearby SNPs on chromosome 6 were confidently associated with survival to eclosion, but given the unbalanced design (most caterpillars did not survive to eclosion) and the modest difference between PGE (and to a lesser extent PVE) estimates for this trait and permutations of this trait, we do not interpret or discuss these associations with survival further.

We next summarized the genomic distribution of genetic variants affecting each trait by estimating the number of QTL (or QTN) for each trait on each of the eight *M. truncatula* chromosomes (as in Santure et al., 2015; Lucas et al., 2018). This was done by summing the PIPs across all SNPs on each chromosome, and thus is analogous to the parameter \( n_{\gamma} \), except that it refers to specific chromosomes rather than the whole genome (Guan & Stephens, 2011; Lucas et al., 2018; Riesch et al., 2015). As these chromosomes vary little in size (~35 to 55 megabases), the number of QTL per chromosome should be similar across chromosomes if the traits are highly polygenic. Consistent with this prediction, evidence of putative QTL for most traits was not restricted to specific chromosomes but distributed relatively evenly among chromosomes (Figures 3 and S8).

### 3.3 Relationship between plant trait genetics and caterpillar performance

Trait genetic correlations were high for some pairs or sets of traits (high genetic correlations imply pleiotropy or tight linkage of causal variants; Figure 4). For example, genetic correlations among leaf length, width, area and dry weight were all \( r \geq 0.8 \). High, positive genetic correlations were also observed among the caterpillar performance traits, particularly 16-day weight, survival to pupation and survival to eclosion (\( r = 0.47 \) to \( 0.60 \)). Caterpillar performance traits also exhibited nontrivial genetic correlations with several plant traits, most notably with leaf toughness where genetic correlations ranged from ~0.25 for 8-day weight (95% confidence intervals [CIs] = ~0.43 to ~0.05, \( p = 0.016 \)) to ~0.39 for 16-day weight (95% CIs = ~0.55 to ~0.21, \( p < 0.001 \); Figure 4). Weaker, but still consistently negative genetic correlations were observed between caterpillar performance traits and both trichome density and plant height (Figure S9).

More generally, hierarchical clustering revealed sets or modules of traits with high (positive or negative) genetic correlations, particularly for suites of IR spectra traits (Figure S10).

Trait genetic correlations inferred from GEBVs based only on SNPs with measurable effects (i.e., using only \( \beta \) and genetic correlations inferred from BLUP estimates of breeding values (which are derived from among-line phenotypic differences) were similar to each other and to genetic correlations inferred from \( \beta \) and \( u \) described in the previous paragraph (Figures 4 and 5a,b). In contrast, expected genetic correlations based on SNP/QTL effects (but ignoring LD) were much lower (e.g., \(|r| = 0.00\) to 0.87; 2.3% of the correlations exceeded 0.1, whereas this was true for 74.4% of the genetic correlations inferred from the GEBVs). Nonetheless, they were still moderately correlated with G-matrixes inferred from the other methods (Figures 5c and S11). These expected genetic correlations were largely explained by correlations in model-average SNP effects among traits (Figure 5d). Consistent with this, correlations among traits in model-averaged SNP effects were generally low, but increased substantially when effects were averaged over larger windows of contiguous SNPs (e.g., the average Pearson correlation

---

**FIGURE 4** Heat map of (additive) genetic Pearson correlation coefficients for pairs of plant and caterpillar traits. For each trait and plant line, genetic correlations were computed from genomic estimated breeding values (GEBVs) that included loci with measurable effects and the polygenic term (upper triangle), and based only on loci with measurable effects (lower triangle). The two sets of GEBVs were highly correlated with each other (\( r = 0.987, \) 95% CIs = 0.985, 0.990) [Colour figure can be viewed at wileyonlinelibrary.com]
Multiple regression models with Bayesian model averaging had some (albeit modest) predictive power, with correlations between observed and predicted caterpillar performance GEBVs ranging from $r = 0.12$ for survival to pupation (i.e., $r^2 = 1.4\%$ of the variation in observed GEBVs explained by predictions) to $r = 0.42$ for 8-day weight ($r^2 = 17.6\%$ of the variation in the observed GEBVs explained by predictions; Figure 6). The most important predictor for 8-day caterpillar weight was IR 892.38, followed by IR 1,072.19 (IR traits based only on SNPs with measurable effects (panel a and c) and BLUPs (panel b). In addition, we considered the expected genetic correlation based on SNPs with measurable effects but ignoring LD (as in Gardner & Latta, 2007. QTL genetic cors.) (panels c and d), and the correlation in the effect estimates of SNPs with measurable effects (panel d). Coloured points denote whether correlations are between pairs of plant traits (including IR traits), pairs of caterpillar performance traits (insect), or between a plant and insect trait. In each panel, a one-to-one line (solid line) and dashed vertical and horizontal lines delineating positive versus negative correlations are shown for reference [Colour figure can be viewed at wileyonlinelibrary.com].

For 8-day caterpillar weight GEBVs, predictions from random forest regression accounted for 31.9% of the out-of-bag (OOB) variance (OOB variance measures predictive performance) ($\text{MTRY} = 18$, $\text{NODERSIZE} = 2$). The most important predictor variables were IR 892.38, IR 985.1, and plant height (Figure 7a). For 16-day caterpillar weight GEBVs, random forest explained 14.4% of the OOB variance ($\text{MTRY} = 12$, $\text{NODERSIZE} = 9$). The most important predictor variables in this case were leaf toughness, IR 1,104.64, and IR 830.13 (Figure 7b). Only 5.3% of the OOB variance was explained for survival to eclosion, with leaf toughness and IR 830.13 being the most important traits (Figure 7c). Graphical analyses of the random forest regression results suggested nonlinear relationships between GEBVs for many of the top plant and caterpillar traits (Figures 7d–f, S14 and S15). For example, the effects of IR 892.38 and IR 985.1 on 8-day weight exhibited a strong interaction (a similar pattern held for many of the IR chemical features). In contrast, the effect of leaf toughness on 16-day weight was negative and nearly linear (tougher leaves were associated with lower weights), although there was evidence of an asymptote at

![Image](https://wileyonlinelibrary.com/doi/10.1111/mec.14358)
higher values of leaf toughness. We failed to explain a nonzero proportion of the OOB variance in caterpillar survival to pupation with random forest regression, and thus results for this trait are not shown.

4 | DISCUSSION

Because the world is full of newly-formed host-parasite interactions (including plant–insect interactions involving herbivory; Nylin et al., 2018), and because most novel host plants are relatively suboptimal hosts (Yoon & Read, 2016), the results reported here are of interest not only as a step towards understanding the interaction between *Lycaenides melissa* and *Medicago truncatula*, but also as a more general model for the formation of host-parasite interactions. In addition, genetic dissections of plant–insect interactions are important not only for understanding the complexity underlying the formation and persistence of new associations, but also for understanding the evolution of plant defensive traits and phytochemical diversity in terrestrial ecosystems. In our study, genome-wide genetic variation within *M. truncatula* explained a nontrivial proportion of the variation in *L. melissa* caterpillar performance traits, especially 16-day weight (PVE = 0.41) and survival to pupation (PVE = 0.31). Estimates of the variance in plant and caterpillar traits explained (PVE) by plant genetic variation were similar, meaning the two sets of traits were highly correlated

| Trait                  | Coefficient Value | Coefficient Value | Coefficient Value | Coefficient Value |
|------------------------|-------------------|-------------------|-------------------|-------------------|
| Wgt. 8 days, r = 0.42  |                   |                   |                   |                   |
| Wgt. 16 days, r = 0.4  |                   |                   |                   |                   |
| Surv. pupation, r = 0.12 |                   |                   |                   |                   |
| Surv. eclosion, r = 0.25 |                   |                   |                   |                   |

**FIGURE 6** Barplots showing the effect of the genetic component of plant traits on the genetic component of caterpillar performance, specifically (a) weight at 8 days, (b) weight at 16 days, (c) survival to pupation, and (d) survival to eclosion. Bars denote ± one standard deviation of the posterior (analogous to a standard error). Colours distinguish between plant growth and defence traits (green) and IR traits (pink). Pearson correlations between the caterpillar performance GEBVs and estimates of these from 10-fold cross-validation are given in the panel headers (see the main text for corresponding $r^2$ values). See Figure S13 for covariate posterior inclusion probabilities [Colour figure can be viewed at wileyonlinelibrary.com]
of these plant traits are affected by some of the same segregating genetic variants (i.e., pleiotropy), or that modest to high LD exists among genetic variants affecting the plant traits and caterpillar performance. Such high LD would imply tight linkage among many genetic variants, or some alternative process or mechanism for suppressed recombination among genotypes (this could include low rates of gene flow among the natural source populations from which these lines were derived). However, LD is modest and decays within a few kbs to background levels in this mapping population (i.e., mean LD, measured by $r^2$ drops below 0.2 within 20 kbs; Branca et al., 2011) (but see the following section for additional discussion and caveats). Interestingly, the additive effects of alleles on the measured plant traits (as captured by the trait GEBVs) were able to explain or account for the additive effects of $M$. truncatula alleles on caterpillar performance, at least to a modest extent (as expected, explanatory power was lower for cross-validation than in simple linear models). Nonetheless, much of the variation in caterpillar performance GEBVs was not accounted for by the plant trait GEBVs. This implies additional plant traits (and underlying genes) likely contribute to the total variation in caterpillar performance explained by plant genetics. We discuss these results in more detail below.

### 4.1 Polygenic variation with pleiotropic effects: evidence and caveats

Our results suggest that most plant traits and caterpillar performance were affected by multiple causal mutations dispersed across the eight $M$. truncatula chromosomes rather than localized in one or a few regions of the genome. A polygenic basis for caterpillar performance (as a plant trait) was similarly detected in a recent genomic study of Pieris rapae caterpillars reared on Arabidopsis thaliana (Nallu et al., 2018). More generally, polygenic variation for resistance to insects has been documented in numerous other plant species, especially crops (Schoonhoven et al., 2010; Via, 1990), although mostly without genome-scale data and without explicit links to plant traits (reviewed in Schoonhoven et al., 2010).

GWA mapping methods (and to a lesser extent genomic prediction methods) are known to suffer from a failure to detect many small effect variants (Eichler et al., 2010; Yang et al., 2010), and from over-estimating the effects of large effect variants (i.e., the Beavis effect; Beavis, 1998). This certainly applies in our case, as our total sample size ($N = 470$) is modest for GWA mapping. However, major-effect loci are less likely to be missed, and thus our inference of a polygenic basis for most traits is relatively robust. This is true in general as
such loci are easier to detect even with small sample sizes, but especially true here given the high-density genome-wide SNP data set we used (>5 million SNPs, or about one per 100 bps) and thus the high likelihood of LD between at least one of our SNPs and most causal variants. Moreover, two of the plant traits we analyzed, plant height and trichome density, were independently mapped and analyzed in an earlier study of the \textit{M. truncatula} HapMap mapping population (albeit with a different subset of lines; Stanton-Geddes et al., 2013). Results from Stanton-Geddes et al. (2013) and our results were remarkably consistent, with, for example, 58% versus 59% (plant height) and 45% versus 49% (trichome density) of the trait variation partitioned among lines in Stanton-Geddes et al. (2013) versus our study, respectively. This is reassuring, particularly given the variability frequently observed in genetic mapping and quantitative genetic results among mapping populations and environments (e.g., Weiss, 2008; Weinig, Ungerer, & Dorn, 2002; Weinig, Stinchcombe, & Schmitt, 2003).

We inferred similar trait genetic correlations from three different estimates of breeding values: GEBVs based on \( \beta \), GEBVs based on \( \beta \) and \( u \), and BLUPs. As these metrics incorporate genes of major versus minor or even infinitesimal effects to different degrees, this suggests that patterns of pleiotropy (or LD) at loci with very small versus more substantial effects on the traits measured were similar. However, it remains unclear whether the observed genetic correlations (among plant traits and between plant and caterpillar traits) were actually caused by mutations with effects of multiple traits (true pleiotropy) or by LD among physically linked mutations with effects on different traits.

The limited extent of LD in the mapping population is consistent with the pleiotropy hypothesis, but the low estimates of expected genetic correlations (i.e., expectations in the absence of LD), and the similarly low correlations in model-averaged SNP effects support the distinct mutations/LD hypothesis. Even if distinct but tightly linked SNPs are associated with different traits, this might or might not implicate different causal mutations. For example, differences in phenotypic distributions/frequencies among traits could lead to different SNP-trait associations even when the traits are affected by the same causal mutation. This might be especially likely when the effects of multiple SNPs are estimated simultaneously in a single model (as was the case here). Thus, we interpret the observed and expected genetic correlations as upper and lower bounds (respectively) on the true degree of pleiotropy in this system. Moreover, from a functional and analytical perspective, true pleiotropy is perhaps best viewed as one end-point along a LD/linkage continuum that spans no to free recombination among causal mutations. Previous studies that have parsed pleiotropy versus LD from trait mapping results have defined pleiotropy liberally based on overlap in QTL or QTL localizing within large genomic windows (e.g., Albert et al., 2008; Brem, Yvert, Clinton, & Kruglyak, 2002; Gardner & Latta, 2007; Morley et al., 2004). Along these lines, evidence for strict pleiotropy in our system based on correlations in SNP effects/associations increased when large windows of contiguous SNPs were considered.

In the end, genetic manipulations (e.g., CRISPR/Cas9) might be necessary to distinguish true pleiotropy from tight linkage (gene knockouts have been used for this in the past, but still suffer from limited genomic resolution as they affect entire genes not individual nucleotides, e.g., Wang, Liao, & Zhang, 2010).

### 4.2 Possible basis for pleiotropic effects across species, and variance left unexplained

Leaf toughness, and to a lesser extent, trichome density and plant height, exhibited some of the greatest and most consistent negative genetic correlations with \textit{L. melissa} performance. Leaf toughness and trichome density constitute structural (physical) plant defences (Levin, 1973; Schoonhoven et al., 2010), and our results thus support recent calls for greater attention to structural (as opposed to chemical) plant defences (Carmona, Lajeunesse, & Johnson, 2011; Hanley et al., 2007; Malishev & Sanson, 2015). However, some IR chemical features exhibited high genetic correlations with some or many of the caterpillar performance traits. This is consistent with a role for intraspecific variation in phytochemical defences in \textit{M. truncatula} as well, although the IR chemical features could also reflect variation in plant nutritional composition rather than chemical defences per se. Future work should identify the molecules underlying variation at the leading IR chemical features (e.g., IR 892.38 and IR 1,104.64).

Plant trait GEBVs accounted for a moderate amount of the variation in caterpillar weight GEBVs, but relatively little of the variation in caterpillar survival GEBVs. In other words, our results suggest that the alleles affecting the measured plant traits accounted for a greater proportion of the heritable variation in \textit{M. truncatula} for caterpillar weight than caterpillar survival. Nonetheless, in no cases did the variance explained or predictive power of these models approach 100%. In fact, the highest percent variance explained was 40.8%, and predictive power never exceeded 17.6% for the Bayesian multiple regression or 31.9% for the random forest regression. This means that the effects of \textit{M. truncatula} alleles on caterpillar performance are not fully accounted for by the effects of these alleles on the measured plant traits. Additional heritable plant traits not measured in this study must affect \textit{L. melissa} performance, and additional work will be required to identify these. Obvious candidates include defensive phytochemicals or plant nutrients that were not captured by the IR assays. Still, even the modest predictive power of these models allows us to conclude, for example, that the genetic quality of a plant in terms of caterpillar performance can be predicted in part from the additive effects of plant alleles on leaf toughness.

As expected, the plant traits most important in these predictive models tended to be the ones with the largest genetic correlations with caterpillar performance. However, there were a few exceptions that arose because of correlations among the plant trait GEBVs, which rendered a subset of these traits (e.g., trichome density) unimportant in the predictive models. Moreover, the relative ranks of plant traits in terms of their importance (i.e., Bayesian model-averaged effect estimates or percent reduction in MSE) differed between the Bayesian multiple regression models and random
forest regression. We think these differences were most evident in cases where random forest regression identified extreme interactions among plant trait GEBVs or nonlinear relationships between GEBVs for the plant traits and caterpillar performance (e.g., IR 985.1 on 8-day caterpillar weight), as these would not be captured by the Bayesian multiple regression models.

5 | CONCLUSIONS AND FUTURE DIRECTIONS

We have shown that plant genetic variation can have a substantial effect on the outcome of a plant–insect interaction, specifically on whether Lycaenides melissa caterpillars can develop successfully on Medicago truncatula. Genetic variation among M. truncatula plants explained about as much of the variance in caterpillar performance in the current study (9%–41%) as genetic variation among L. melissa caterpillars did in an earlier rearing experiment on Medicago sativa (7%–57%) (Gompert et al., 2015; considerably less variation was detected among caterpillar families in the current study). This suggests that caterpillar and plant genetic variation combined could explain a large proportion (i.e., over half) of the variation in larval performance, which is necessarily a key aspect of the interaction between plants and herbivorous insects. However, M. truncatula and M. sativa are not identical, and it remains to be seen whether similar levels of genetic variation for performance exist in this actual (rather than potential) L. melissa host plant. Moreover, gene by gene epistatic interactions between L. melissa alleles and M. sativa (or M. truncatula) alleles could modulate the total variance in performance explained for the pair of species (in other words, the trait heritabilities with respect to plant and insect genes are not necessarily additive).

Ultimately, we want to accurately predict the mosaic patterns of host use and host adaption in L. melissa across the landscape from a mechanistic understanding of the factors affecting host use (e.g., Chaturvedi et al., 2018). We have reasons to be both optimistic and pessimistic about this aim. Past work on L. melissa has shown that genetic variants associated with performance in the laboratory covary significantly with host use in nature (Chaturvedi et al., 2018; Gompert et al., 2015). Thus, genetic variants affecting performance in the laboratory appear to also be associated with host-plant adaptation in nature. On the other hand, the laboratory environment is necessarily simplified and lacks interactions with predators, competitors and mutualists that could be important determinants of host use in the wild. For example, survival of L. melissa caterpillars on M. sativa in a field experiment depended on the presence of ants that defend the caterpillars from predators (this is a facultative relationship where the ants receive a sugar reward from the caterpillars; Forister, Gompert, Nice, Forsiter, & Fordyce, 2011). Even ignoring such complexities, the relevance of genetic and trait variation in M. truncatula for understanding genetic and trait variation in M. sativa is not certain. Leaf toughness, which was most strongly associated with performance in the current experiment, exhibits a similar range of variation in M. sativa and M. truncatula (albeit with somewhat tougher leaves in M. sativa on average; Harrison et al., 2018). This suggests variation in leaf toughness in M. sativa could have a similar affect on caterpillar performance. In the end, we may fail to generate reliable predictions about host use in nature from simple laboratory experiments, but nonetheless might advance scientific understanding of the importance of intraspecific variation for the evolution and ecology of plant–insect interactions by gaining a better understanding of how and why these predictions fail.

ACKNOWLEDGEMENTS

We thank Shaylen Fidel, Camden Treat, Marissa Kee and Kenneth Tache for their help with the caterpillar rearing experiment, and Marianne Harris for use of greenhouse space and help caring for the plants. Chris Nice, James Fordyce, Alex Buerkle, Amy Springer and two anonymous reviewers provided valuable comments on an earlier draft of this manuscript. Medicago truncatula seeds were provided by INRA-Montpellier and the USDA Agricultural Research Station at Washington State University. This research was funded by the National Science Foundation (DEB-1638768 to ZG and DEB-1638793 to MLF), and by the Utah Agricultural Experiment Station (paper no. 9208). The support and resources from the Center for High Performance Computing at the University of Utah are also gratefully acknowledged.

AUTHOR CONTRIBUTIONS

Z.G., L.K.L. and M.B. designed the study. Z.G., L.K.L., M.B., F.Z., C.P., M.J.T. and M.L.F. conducted the experiment. Z.G., F.C., C.P. and T.S. analyzed the data. Z.G., C.P. and T.S. wrote the manuscript. All authors revised and edited the manuscript.

DATA ACCESSIBILITY

Original data and scripts have been deposited on Dryad, https://doi.org/10.5061/dryad.nj29q8b.

ORCID

Zachariah Gompert https://orcid.org/0000-0003-2248-2488

REFERENCES

Albert, A. Y., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., ... Schluter, D. (2008). The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. Evolution, 62, 76–85.
Austin, G. T., & Leary, P. J. (2008). Larval hostplants of butterflies in Nevada. Holarctic Lepidoptera, 12, 1–134.
Barrett, R. D. H., Laurent, S., Mallarino, R., Pfeifer, S. P., Xu, C. C. Y., Foll, M., ... Hoekstra, H. E. (2019). Linking a mutation to survival in wild mice. Science, 363, 499–504. https://doi.org/10.1126/science.aav3824
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67, 1–48.
inferred from mtDNA. Molecular Phylogenetics and Evolution, 48, 481–490. https://doi.org/10.1016/j.ympev.2008.04.035
Gompert, Z., Jahner, J. F., Scholl, C. F., Wilson, J. S., Lucas, L. K., Soria-Carrasco, V., … Forister, M. L. (2015). The evolution of novel host use is unlikely to be constrained by trade-offs or a lack of genetic variation. Molecular Ecology, 24, 2777–2793. https://doi.org/10.1111/mec.13199
Greven, S., Crainiceanu, C. M., Küchenhoff, H., & Peters, A. (2008). Restricted likelihood ratio testing for zero variance components in linear mixed models. Journal of Computational and Graphical Statistics, 17, 870–891. https://doi.org/10.1198/106186008X36599
Guan, Y., & Stephens, M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. Annals of Applied Statistics, 5, 1780–1815. https://doi.org/10.1214/11-AOAS455
Hanley, M. E., Lamont, B. B., Fairbanks, M. M., & Rafferty, C. M. (2007). Plant structural traits and their role in anti-herbivore defence. Perspectives in Plant Ecology, Evolution and Systematics, 8, 157-178. https://doi.org/10.1016/j.ppees.2007.01.001
Hansen, T.F., & Houle, D.(2008). Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology, 21, 1201-1219. https://doi.org/10.1111/j.1420-9101.2008.01573.x
Harrison, J. G., Philbin, C. S., Gompert, Z., Forister, G. W., Hernandez-Wilson, A., Sullivan, B. W., … Forister, M. L. (2018). Deconstruction of a plant-arthropod community reveals influential plant traits with nonlinear effects on arthropod assemblages. Functional Ecology, 32, 1317–1328. https://doi.org/10.1111/1365-2435.13060
Hayes, B., Bowman, P., Chamberlain, A., & Goddard, M. (2009). Genomic selection in dairy cattle: Progress and challenges. Journal of Dairy Science, 92, 433–443.
Hendry, A. P. (2016). Eco-evolutionary dynamics. Princeton, NJ: Princeton University Press.
Kang, Y., Y. Sakiroglu, M., Krom, N., Stanton-Geddes, J., Wang, M., Lee, Y. C., … Udvardi, M. (2015). Genome-wide association of drought-related and biomass traits with hapmap SNPs in Medicago truncatula. Plant, Cell and Environment, 38, 1997–2011.
Lankau, R. A. (2012). Coevolution between invasive and native plants driven by chemical competition and soil biota. Proceedings of the National Academy of Sciences, 109, 11240–11245. https://doi.org/10.1073/pnas.1201343109
Levin, D. A. (1973). The role of trichomes in plant defense. The Quarterly Review of Biology, 48, 3–15. https://doi.org/10.1086/407484
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics, 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352
Li, W., Schuler, M. A., & Berenbaum, M. R. (2003). Diversification of furanocoumarin-metabolizing cytochrome P450 monoxygenases in two papilionids: Specificity and substrate encounter rate. Proceedings of the National Academy of Sciences, 100, 14593–14598. https://doi.org/10.1073/pnas.031343109
Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2, 18–22.
Lucas, L. K., Nice, C. C., & 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, 892–907.
Luijckx, P., Ben-Ami, F., Mouton, L., Du Pasquier, L., & Ebert, D. (2011). Cloning of the unculturable parasite pasteuria ramosa and its Daphnia host reveals extreme genotype-genotype interactions. Ecology Letters, 14, 125–131. https://doi.org/10.1111/1461-0248.2010.01561.x
Luijckx, P., Fienberg, H., Duneau, D., & Ebert, D. (2013). A matching-allele model explains host resistance to parasites. Current Biology, 23, 1085–1088. https://doi.org/10.1016/j.cub.2013.04.064
Lynch, M., & Walsh, B. (1998). Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer Associates.
Malishev, M., & Sanson, G. D. (2015). Leaf mechanics and herbivory defense: How tough tissue along the leaf body deters growing insect herbivores. Austral Ecology, 40, 300–308. https://doi.org/10.1111/ace.12214
Mandeville, E. G., Parchman, T. L., Thompson, K. G., Compton, R. I., Gelwick, K. R., Song, S. J., & Buerkle, C. A. (2017). Inconsistent reproductive isolation revealed by interactions between catatomus fish species. Evolution Letters, 1, 255–268.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20, 1297–1303. https://doi.org/10.1101/gr.107524.110
Meuwissen, T. H. E., Hayes, B., & Goddard, M. (2001). Prediction of total genetic value using genome wide dense marker maps. Genetics, 157, 1819–1829.
Michaud, R., Lehman, W. F., & Rumbaugh, M. D. (1988). World distribution and historical developments. In A. A. Hanson, D. K. Barnes, & R. R. Hill (Eds.), Alfalfa and alfalfa improvement, vol. 29. Madison: World Distribution and Historical Developments.
Miliborrow, S. (2018). plotmo: Plot a model's residuals, response, and partial dependence plots. R Package Version 3.5.1.
Mitchell, C., Brennan, R. M., Graham, J., & Karley, A. J. (2016). Plant defense against herbivorous pests: Exploiting resistance and tolerance traits for sustainable crop protection. Frontiers in Plant Science, 7, 1132–2000. https://doi.org/10.3389/fpls.2016.01132
Mitter, C., Farrell, B., & Wiegmann, B. (1988). The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification? The American Naturalist, 132, 107–128. https://doi.org/10.1086/284480
Moreau, D. (2006). Morphology, development and plant architecture of M. truncatula. In The Medicago truncatula handbook. Dijon, France: INRA.
Morley, M., Molony, C. M., Weber, T. M., Devlin, J. L., Ewens, K. G., Spielman, R. S., & Cheung, V. G. (2004). Genetic analysis of genome-wide variation in human gene expression. Nature, 430, 743–747. https://doi.org/10.1038/nature02797
Moser, G., Lee, S. H., Hayes, B. J., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2015). Simultaneous discovery, estimation and prediction analysis of complex traits using a Bayesian mixture model. PLoS Genetics, 11, 1–22. https://doi.org/10.1371/journal.pgen.1004969
Nallu, S., Hill, J. A., Don, K., Sahagun, C., Zhang, W., Meslin, C., … Kronforst, M. R. (2018). The molecular genetic basis of herbivory between butterflies and their host plants. Nature Ecology & Evolution, 2, 1418–1427. https://doi.org/10.1038/s41559-018-0629-9
Nosil, P. (2004). Reproductive isolation caused by visual predation on migrants between divergent environments. Proceedings of the Royal Society London Series B-Biological Sciences, 271, 1521–1528.
Nosil, P., Villoutreix, R., de Carvalho, C. F., Farkas, T. E., Soria-Carrasco, V., Feder, J. L., … Gompert, Z. (2018). Natural selection and the predictability of evolution in Timema stick insects. Science, 359, 765–770.
Nouhoud, P., Gautier, M., Gouin, A., Jauquier, J., Pecoud, J., Legeai, F., … Simon, J.-C. (2018). Identifying genomic hotspots of differentiation and candidate genes involved in the adaptive divergence of pea aphid host races. Molecular Ecology, 27, 3287–3300. https://doi.org/10.1111/mec.14799
Nylin, S., Agosta, S., Bensch, S., Boeger, W. A., Braga, M. P., Brooks, D. R., … Janz, N. (2018). Embracing colonization: A new paradigm for species association dynamics. Trends in Ecology & Evolution, 33, 4–14. https://doi.org/10.1016/j.tree.2017.10.005
Ober, U., Ayroles, J. F., Stone, E. A., Richards, S., Zhu, D., Gibbs, R. A., … Simianer, H. (2012). Using whole-genome sequence data to predict quantitative trait phenotypes in Drosophila melanogaster.
SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Gompert Z, Brady M, Chalyavi F, et al. Genomic evidence of genetic variation with pleiotropic effects on caterpillar fitness and plant traits in a model legume. Mol Ecol. 2019;28:2967–2985. https://doi.org/10.1111/mec.15113