the geneticssociety

www.nature.com/hdy

ARTICLE

Check for updates

Differences in gene expression and genetic variation underlying preference-performance mismatches: insights from a specialized native herbivore on an invasive toxic plant

Nitin Ravikanthachari[1,2,4]✉ and Carol L. Boggs[1,2,3]

Specialist phytophagous insects have a narrow hostplant range for optimal development and survival. Mismatches between female oviposition preference and larval performance can lead to high fitness costs. Understanding the mechanistic basis of this decoupling can help us understand evolutionary constraints and aid in predicting outcomes of error-prone oviposition. We investigated the causes for preference-performance mismatches in a specialist native herbivore laying eggs on an invasive toxic plant. Transcriptomic analyses revealed host-plant-specific gene expression signatures in larvae feeding on different plants, while there was no differential gene expression in gustatory/olfactory organs of adult females with different oviposition preferences. However, genomic analysis revealed significant genetic differentiation in several genes underlying signal transduction in adult females with different oviposition preferences. The larvae feeding on toxic plants showed lower expression of specialized detoxification enzymes and higher expression of general digestive enzymes, indicating the inability of larvae to detoxify toxic compounds present in the toxic plants. We additionally found that genes related to successful detoxification and adaptive feeding were enriched in larvae feeding on native plants, while genes related to toxic responses, apoptosis, and accelerated development were enriched in larvae feeding on toxic plants. Our findings dissect the underlying mechanisms behind a preference-performance mismatch, quantifying the impact of error-prone oviposition on larval performance in a specialized species interaction.

Heredity; https://doi.org/10.1038/s41437-025-00777-7

INTRODUCTION

Herbivorous insects and the plants they feed on are a classic example of co-evolution (Ehrlich and Raven 1964; Thompson 1999a, 1999b). Herbivorous insects have evolved to feed on a narrow range of host plants by adapting to circumvent specialized plant defenses (Jaenike 1990; Joshi and Thompson 1995; Hardy and Otto 2014). In holometabolous insects such as Lepidoptera, where the larval stages are relatively immobile, the success of the herbivore depends on the female choosing a suitable host that maximizes the survival of the resulting offspring (Berdegué et al. 1998; Bossart and Scriber 1999; Ladner and Altizer 2005). Females primarily rely on cues from their gustatory senses, consisting of gustatory receptors, gustatory neurons, and their underlying genes tuned to identify suitable host plants for oviposition. Variation in hostplant preference can arise due to underlying variation in gustatory genes tuned to different compounds (Ozaki et al. 2011), or differences in gene expression in the gustatory genes in response to hostplant stimuli (Orsucci et al. 2018), or due to differences in neural sensitivity for specific hostplant compounds (Sollai et al. 2018). As a result, oviposition cues in specialized herbivores are highly co-evolved with their host plants, leading to strong selection for female oviposition preference for suitable host plants. This idea, known as the preference-performance hypothesis, postulates that the preference for a host plant by the female is positively correlated with the performance of the larvae on that host plant (Jaenike 1978; Valladares and Lawton 1991). Female host choice is dependent on a multitude of factors, including the ability to distinguish between suitable host plants and toxic non-host plants, hostplant specialization, hostplant density, and presence of natural enemies (Larsson and Ekbom 1995; García-Robledo and Horvitz 2012; Carrasco et al. 2015).

Support for the preference-performance hypothesis ranges from none to strong correlation in various insect groups (Rausher 1979; Valladares and Lawton 1991; Menacer et al. 2021). A meta-analysis showed that oligophagous insects (insects feeding on several plant genera within a plant family) had a stronger preference-performance correlation compared to polyphagous insects (insects feeding on several plant families) (Gripenberg et al. 2010). However, for specialized insects, the evidence is mixed. One underlying hypothesis for this mixed evidence is that cue similarity between hosts and non-hosts often leads to error-prone oviposition in females (Trowbridge and Todd 2001; Steward and Boggs 2020).

Error-prone oviposition can result in several scenarios (Wiklund 1975). First, erroneous oviposition can be adaptive if the larvae can survive on the novel host plant, thus leading to an increase in the diet breadth of the herbivore and adaptation to new host plants

[1]Department of Biological Sciences, University of South Carolina, Columbia, SC, USA. [2]Rocky Mountain Biological Laboratory, Crested Butte, CO, USA. [3]School of Earth, Ocean & Environment, University of South Carolina, Columbia, SC, USA. [4]Present address: Division of Biological Sciences, University of Montana, Missoula, MT, USA. Associate editor: Sara Goodacre. ✉email: ravikanth.nitin@gmail.com

Received: 26 January 2025 Revised: 10 June 2025 Accepted: 12 June 2025
Published online: 03 July 2025

(Janz et al. 1994; Stefanescu et al. 2012). In fact, error-prone oviposition occurs in many species (Nylin and Janz, 1993, 2009; Janz et al. 1994; Nylin et al. 2000) and has often been cited as the fuel driving hostplant range in Lepidoptera (Larsson and Ekbom 1995; Stefanescu et al. 2012). The success of the larvae on the novel host depends on the degree of chemical similarity of the novel host to the hosts that the larvae have co-evolved with and the insect's underlying variation in detoxification enzymes, as determined by the evolutionary history of host use (Futuyma and Agrawal 2009; Erbilgin et al. 2014; Celorio-Mancera et al. 2016; Calla et al. 2017). The secondary compounds present in the plant are detoxified upon ingestion by the larvae by various enzymes that act on specific substrates involving complex redox reactions and conjugation of chemical substrates (Heidel-Fischer and Vogel 2015). The detoxification machinery of insects involves three important phases (Heckel 2014): phase 1, which includes the cytochrome p450 genes and carboxylesterases are involved in redox reactions of plant secondary metabolites (Heidel-Fischer and Vogel 2015; Dermauw et al. 2020). Phase 2 enzymes include glutathione-S-transferase, sulfo-transferases, and glucuronosyl-transferase (UDPGT). Enzymes in phase 2 act on the byproducts of phase 1 enzymes. In specialist herbivores, the plant compounds are directly acted upon by phase 2 enzymes by circumventing the conjugation step on phase 1 enzymes (Groen and Whiteman 2022). Phase 3 enzymes, which include ATP-binding cassettes, act on the byproducts of phase 2 enzymes and excrete the toxic compounds out of the cell. In addition to these enzymes, trypsin and insect cuticle proteins are involved in improving digestion efficiency (Dermauw and Van Leeuwen 2014).

Alternatively, error prone oviposition can lead to high fitness costs if the larvae do not possess behavioral or physiological mechanisms to utilize the novel host plant, thus leading to higher mortality and maladaptation (Casagrande and Dacey 2007; Keeler and Chew 2008; Yoon and Read 2016; Steward et al. 2019; Steward and Boggs 2020). Repeated error-prone oviposition has two possible outcomes. The resulting selection pressure on larval performance can lead to the larvae evolving the ability to feed on the plant, provided that new mutations or underlying standing genetic variation provide the grist for the evolution of key innovations in detoxification machinery (Nylin and Janz 1993; Wheat et al. 2007).

However, in cases where suitable genetic variation in larval performance is lacking and repeated erroneous oviposition occurs, evolutionary traps arise. In particular, abrupt or rapid environmental changes can lead to instances where evolved reliable cues fail, resulting in repeated oviposition errors and larval mortality (Keeler and Chew 2008; Yoon and Read 2016; Singer and Parmesan 2019; Steward and Boggs 2020; Singer 2021). This could occur either due to females' inability to distinguish between hosts and non-hosts to avoid costly mistakes or the failure of larvae to incorporate the novel host due to behavioral or physiological constraints (Chew 1977a; Huang et al. 1994; Casagrande and Dacey 2007; Keeler and Chew 2008; Singer and Parmesan 2019; Steward et al. 2019). Identifying the underlying mechanisms that lead to decoupling cue-response systems, as well as preference-performance, can shed light on the constraints of evolution in species interactions.

Although evolutionary traps have been documented in many systems, the mechanistic processes underlying persistent evolutionary traps are not well understood (Robertson and Hutto 2006; Robertson et al. 2013; Hale et al. 2016; Robertson and Chalfoun 2016). Here we use a system of a specialized native herbivore that oviposits on a lethal, invasive plant that causes high mortality in the larvae to dissect the underlying genetic mechanisms affecting adult female oviposition and larval feeding on the plant. While preference-performance mismatches can arise due to various factors, we specifically tested whether preference-performance resulted from (a) differences in gene expression (in adult female oviposition preference and larval performance) and/or (b) genetic differences (in female oviposition preference).

STUDY ORGANISMS AND SPECIFIC HYPOTHESES

Pieris macdunnoughii Remington 1954 (Pieridae; formerly *P. napi macdunnoughii*) (Chew and Watt 2006) is a montane butterfly distributed in the Southern Rocky Mountains in North America and is a specialist herbivore: the females lay eggs on and the larvae feed on native Brassicaceae (Chew 1977a, 1977b, 1980; Nakajima and Boggs 2015). *Pieris macdunnoughii*, like other species in the *Pieris* species complex, has evolved detoxification enzymes to overcome the toxicity of glucosinolates (secondary metabolites) in Brassicaceae (Wheat et al. 2007; Edger et al. 2015). However, larvae experience high mortality when they encounter novel mustards whose glucosinolates/secondary metabolites differ from those with which they have locally co-evolved (Chew 1977b; Haribal and Renwick 1998; Haribal et al. 2001; Keeler and Chew 2008; Steward et al. 2019).

Thlaspi arvense (L.) (Brassicaceae) is a plant native to temperate Eurasia that was introduced to North America in the 1800s (Chew 1977a). It was introduced to Gunnison County between 1850s to 1870s, with the earliest herbarium records dating back to 1929 when the Rocky Mountain Laboratory's herbarium was established. Thus, *T. arvense* has been present in the habitat for at least 94 years, and possibly up to 170 years (Chew 1977a). *Thlaspi arvense* is an early successional plant that colonizes disturbed soil and is known to occur up to 2900 m (Best and McIntyre 1975; Warwick et al. 2002).

The glucosinolate profile of *T. arvense* is comprised mainly of the aliphatic glucosinolate sinigrin, in contrast to native mustards that contain both aliphatic and aromatic glucosinolates (Rodman and Chew 1980). *Pieris macdunnoughii* females recognize *T. arvense* as a potential host plant in areas where both co-occur due to cue similarity with the native plants (Chew 1977a; Nakajima and Boggs 2015). Sinigrin acts both as an oviposition attractant as well as a larval feeding stimulant in the genus *Pieris* (David and Gardiner 1966; van Loon et al. 1992; Städler et al. 1995; Müller et al. 2015). However, higher concentrations of sinigrin in the host plants act as larval feeding deterrents (Steward et al. 2019).

Although larvae may slowly try to eat the plant, no larvae survive to the adult stage, so fitness resulting from any eggs laid on *T. arvense* is 0 (Nakajima et al. 2013). Further, the magnitude of fitness loss to females ovipositing on *T. arvense* is directly proportional to the relative abundance of *T. arvense* in the habitat. Previous work involving hostplant surveys and egg distribution of *P. macdunnoughii* in 2000 in Gothic, Colorado, quantified a relative abundance of 2.7% of *T. arvense* in the habitat, which resulted in 3% fitness loss to females for ovipositing on *T. arvense* (Nakajima et al. 2013; Nakajima and Boggs 2015). Therefore, there is strong selection on females of *P. macdunnoughii* to avoid laying eggs on the plant and for escape from the evolutionary trap.

Recently, female oviposition choice experiments highlighted sex-linked heritable additive genetic variation related to oviposition preference within a population when tested on whole plants. However, additional population genomic analyses of *P. macdunnoughii* from habitats with and without *T. arvense* identified signatures of selection only on genes underlying larval development in habitats where the butterfly and the toxic plant co-occurred. (Ravikanthachari et al. 2024). These experiments indicate that both adult oviposition preference and/or larval performance on the exotic plant are under selection.

To identify whether both life stages or either of the life stages are under selection, we first employed a transcriptomic approach to quantify differential gene expression to address the hypotheses that a) larvae that fed on the toxic plant exhibit transcriptomic signatures of impaired feeding and toxicity (upregulation of stress responses) compared to those that fed on the native host plant and b) females

that preferred the toxic lethal plant differed in their expression repertoire of sensory/gustatory genes compared to those preferring the native plant. Second, we quantified allele frequency differences in (a) sensory/gustatory transcripts of females with differing oviposition preference and (b) in whole genomes of females from habitats with and without previous exposure to the toxic plant to test if fixed allele frequency differences were driving oviposition differences between females. Our study adds to the extensive literature on preference-performance studies by identifying the molecular differences driving the decoupling of mismatches in female preference and larval performance, which leads to ecological and evolutionary traps.

METHODS

Oviposition choice experiment

Gravid females were collected in the East River Valley, Gunnison County, Colorado (38.9664° N, 106.9896° W, 2896 m a.s.l) in 2019 using an aerial net. In the lab, the females were fed twice a day with a 25% honey-water solution. Females were initially housed in the dark at 27c for 24 h to reset learned hostplant preferences. Previous research in several insects has suggested that learned preferences for host plants can be erased when the stimulus is removed for extended periods, due to dynamic changes in mushroom body volumes in insect brains (Barth and Heisenberg 1997; Zars 2000; Snell-Rood 2007; Balkenius and Hansson 2012; van Dijk et al. 2017). After a period of 24 h in the dark, females were housed in plastic cages in an environmental chamber at 27 °C during the day and at 18 °C at night on an 18:6 L:D cycle. The females were each provided with one whole plant of *T. arvense* and *Cardamine cordifolia* A. Gray (Brassicaceae), a primary native host plant, placed in a floral water pick. The larval host plants were visually matched by approximate leaf area and by plant phenology (pre-flowering stage). The females were allowed to lay eggs on the host plants, and the eggs on each plant were counted every morning. The host plants provided to the females were rotated every day to change the orientation inside the cages. Additionally, females were provided with new host plants every day when previous day's plants were removed for counting the number of eggs laid. Preference for a hostplant was quantified as follows: preference for *T. arvense* if >80% of eggs laid on *T. arvense*; preference for *C. cordifolia* if >80% of eggs laid on *C. cordifolia*, and equal preference if no plant received >80% eggs laid. The females were removed while actively ovipositing once they had laid a total of fifty eggs, and their forelegs, antennae, and head were stored in RNAlater at −20 °C.

Larval feeding assays

The larvae from eggs laid by females in the oviposition choice trails were reared in an environmental chamber under the same conditions used during the female oviposition experiment. The larvae were reared on young leaves of *Raphanus sativus* L. (Brassicaceae) until they reached the third instar to avoid bias in our gene expression data and to ensure that larvae were naïve to compounds present on the native and the exotic plant. Five third instar larvae from each female were provided either with a potted whole plant of *C. cordifolia* or *T. arvense* after 12 h of starvation. Visual examination indicated that larvae had no plant material in their guts on placement on the plants. The larvae were allowed to feed for 24 h. After 24 h, the larvae were visually examined to confirm that their midguts were at least partially filled with plant material. The larvae were then dissected in PBS solution, degutted, and plant material removed, and their mouth parts and guts were stored in RNAlater.

RNA extraction and sequencing

Total RNA was extracted from 10 females (head, antennae, and forelegs) each preferring *T. arvense* and *C. cordifolia*, and from 10 third instar larvae (gut and mouthparts) each fed on *C. cordifolia* and *T. arvense* using a QIAGEN RNeasy kit following manufacturer's protocol. Total RNA from all samples was sent to MedGenome for cDNA library preparation and RNA sequencing. cDNA library preparation was carried out using Illumina TruSeq stranded mRNA kit followed by sequencing on a Novaseq S4 platform.

Transcriptome read mapping

Demultiplexed raw Illumina reads were checked for quality using FastQC (Andrews 2010). Reads were trimmed using Trimmomatic v. 0.39 (Bolger et al. 2014) with the following options: Sliding window: 4:20, minimum length: 25, and adapter sequences were cleaned using the option TruSeq3. Cleaned transcripts were mapped to the reference genome (Steward et al. 2021) using genome-guided STAR assembly (Dobin et al. 2013; Dobin and Gingeras 2015). We first generated genome indices and the genome directory, followed by mapping reads to the genome with default options. 94% of the reads were correctly mapped to the genome. The resulting BAM files were sorted by genome coordinates using SAMtools (Li et al. 2009) sort and indexed using SAMtools index. Mapped reads were then quantified using featureCounts (useMetaFeatures geneid) (Liao et al. 2014) and run through Rsubread (Liao et al. 2019) in the R statistical environment v. 4.0 (R Development Core Team 2008). We ran featureCounts separately on adult female (mean read counts across libraries: 1293.2, range of read counts across libraries: 0 - 1171617) and larval (mean read counts across libraries: 1535.5, range of read counts across libraries: 0 - 2017845) mapped reads to obtain their respective expression profiles.

Differential gene expression analysis

Differential gene expression (DGE) analyses for adult females and larvae were carried out separately in edgeR (Robinson et al. 2010). We used the CPM-TMM log2 transformation for further filtering our dataset. The raw counts were filtered based on abundance and then normalized using count-per-million (CPM) to account for differences in library sizes among samples in edgeR. We additionally used the trimmed mean of M values (TMM) for cross-sample normalization using the option "calcNormFactors" in edgeR. Genes that were lowly expressed were filtered using the HTSFilter (which uses a pairwise Jaccard similarity index between pairs of replicates in each experimental condition to identify genes with constant, low-level expressions; Fig. S1), followed by differential gene expression analysis using the Fisher's exact test approach in edgeR. The differentially expressed genes were further filtered using a false discovery rate of 1% using the BH correction method. We set a cut-off of a minimal fold change (FC) of 1.5 between treatments and an FDR of p-value < 1e−3 for assessing significant differential expression. The cut-off values for FC and FDR were set based on prior studies to increase the signal-to-noise ratio and to reduce non-hostplant-specific patterns (Breeschoten et al. 2019, 2022). We used the hclust function in the R package stats for hierarchical clustering of significantly differentially expressed transcripts with a k value of 3 based on results from the sum of square means (SI Fig. 2). We additionally visualized differentially expressed genes using a heatmap using the pheatmap package in R (Kolde and Kolde 2015).

Gene expression in plant compound detoxification genes

We obtained the identity of the genes that were differentially expressed in the larval treatments by linking the gene names to their respective protein names in the protein family (Pfam) database (Finn et al. 2014). We extracted genes that corresponded to the six gene families involved in detoxification of plant compounds (Phase 1: Cytochrome p450s and Carboxylesterase; Phase 2: Glutathione-S-Transferase and Glucuronosyltransferase; Phase 3: ATP Binding Cassettes and General digestion efficiency enzymes: Trypsin, and insect cuticle protein) and grouped them to their respective clusters using all CPM normalized counts (dataset prior to TMM normalization; see above for normalization process). We then ran a mixed linear model using the *lme4* package (Bates et al. 2015) with normalized counts as dependent variable and with host plant identity and gene family identity and their interactions as independent variables and included a nested random intercept for gene family identity within individual identity to account for repeated measures of gene families across individuals.

Weighted gene co-expression analysis

We utilized the WGCNA package in R (Langfelder and Horvath 2008) to identify plant-specific co-expressed genes in the larval and adult treatments separately using TMM normalized gene counts (see above for description of normalization). For larval and adult weighted gene correlation network construction, we selected a soft clustering power of 6 and 9, respectively, based on a scale-free topology model fit (SI Fig. 3a and b; r2 > 0.9). The networks were built using signed networks with Pearson correlation, requiring a minimum of 40 genes per module with a deep split value of 3 and a merge cut height of 0.3. Subsequently, we tested the network for module-trait correlations with plant treatment as a binary variable (*T. arvense* = 1, *C. cordifolia* = 0). We ran a linear fit model using the *lmfit* function of the *limma* package (Ritchie et al. 2015) to identify modules of interest (module-trait correlations that showed the highest difference between plant treatments).

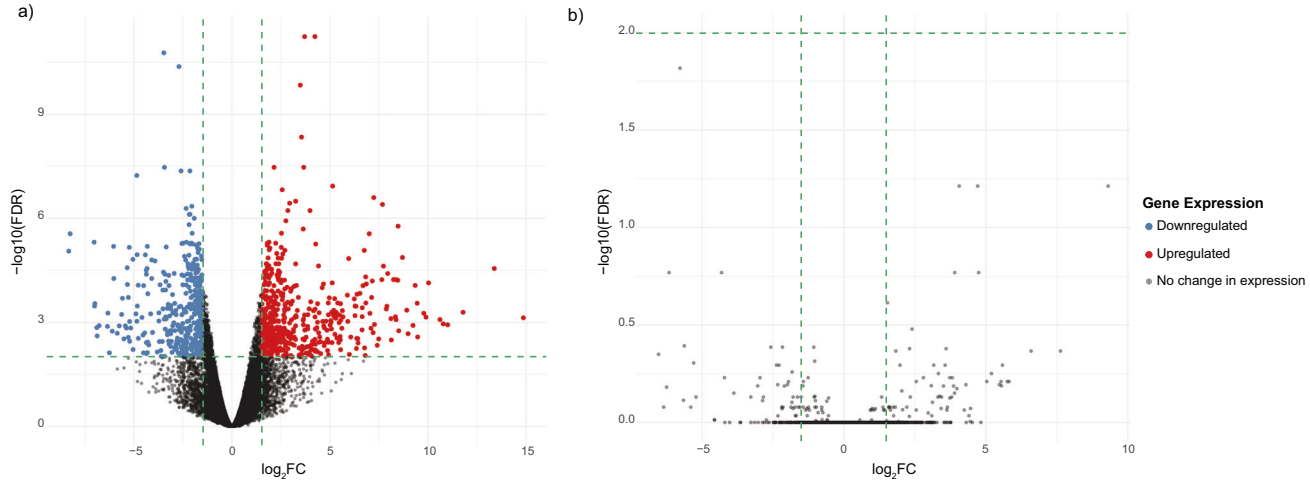

Fig. 1 Differential gene expression in *Pieris macdunnoughii*. Gene expression differences in **a** larvae feeding on and **b** adults laying eggs on *Thlaspi arvense* vs *Cardamine cordifolia*. Red indicates transcripts upregulated in larvae feeding on *T. arvense*, and blue indicates transcripts downregulated in larvae feeding on *T. arvense*. The dashed lines indicate cut-offs for identifying differentially expressed transcripts (FC: > 1.5 and FDR < 0.01).

We used two matrices to further identify specific genes associated with oviposition preference and larval performance. Since none of the modules identified for adult oviposition preference were associated with *T. arvense*, we focused only on the modules that were identified for larval performance associated with *T. arvense*. First, we obtained gene significance and module membership for each gene in our dataset from the WGCNA package. We calculated gene significance for a gene as the correlation between the normalized gene expression of the gene and the treatment variable. We calculated module membership for each gene as the correlation between the normalized gene expression of the gene and the module eigengene values. Additionally, we calculated gene connectivity (a measure of node density for each gene) using the *signedKME* function in the WGCNA package. We considered genes as associated with feeding on *T. arvense* if they showed >0.6 correlation between gene significance vs module membership and scored >0.8 for gene connectivity, based on author recommendations. Second, we used the *chooseTopHubsInEachModule* function in the WGCNA package to identify the top genes in the modules of interest. Hubs are central genes in each module that are highly connected.

Gene set enrichment

We generated gene ontology (GO) annotation of the *P. macdunnoughii* genome using eggNOG-Mapper (Huerta-Cepas et al. 2017) by linking the annotated genes to their respective protein names using the protein family (Pfam) database to obtain GO terms. The GO annotation consisted of 15,477 genes that were used as input for topGO (Alexa and Rahnenführer 2009) to quantify gene set enrichment in the differentially expressed and co-expressed genes associated with larval feeding. We used the parent-child algorithm with a two-tailed Fisher's test to identify enriched GO terms based on biological processes (BP). We calculated gene set enrichment separately for (a) each of the 3 distinct hierarchical gene clusters in the differentially expressed gene dataset and (b) for filtered genes in each module of interest. We considered GO terms as significant for those that overlapped between the DGE dataset and the WGCNA dataset. The final list of enriched GO terms was then run through REVIGO (Supek et al. 2011) to cluster the terms and to reduce redundancy by identifying similarity between the GO terms.

SNP (single-nucleotide polymorphism) calling

To test whether differences in female oviposition preferences were driven by SNP differences rather than gene expression differences, we called SNPs for females differing in oviposition preference using the female transcriptome data. The transcriptome coordinate-sorted BAM files from the RNA read mapping were marked using PICARD tools ('Picard toolkit' 2019). We used SplitNCigarReads in GATK (McKenna et al. 2010) to split reads containing Ns (reads spanning splice events in RNAseq data) to remove mismatching overhangs and to reassess mapping qualities. SNPs were called across all samples using the GATK HaplotypeCaller to generate individual intermediate gVCF files that were then imported using GATK GenomicsDBImport and were finally genotyped using GATK GenotypeGVCFs. We then used GATK SelectVariants to include only those variants that met the criteria: QD > 2, base quality > 30, SOR < 3.0, FS < 60, MQ > 40, MQ Rank Sum > 12.5, Read Pos Rank Sum > 8.0. We used VCFtools (Danecek et al. 2011) to filter the SNPs with the following parameters: minor allele count = 4, max missing = 0.20, min $q = 30$. Our final sample size consisted of 19 individuals with 27,806,300 sites (both invariant and variant sites), with 572,246 SNPs.

Nucleotide summary statistics

We used Pixy (Korunes and Samuk 2021) to calculate a) nucleotide diversity (π) separately for females preferring *T. arvense* and those preferring *C. cordifolia* for all sites (invariant and variant sites combined as suggested by Pixy) using a 5000 bp sliding window approach; and b) genome-wide and 5000 bp sliding window F_{ST} between females with differing preferences using only the variant sites. We ran 100 simulations of 5000 bp sliding window F_{ST} estimates by randomly permuting preference identity of the females to calculate a null distribution to identify the 99th percentile cut-off for delineating significant outlier SNPs. We then used SnpEff (Cingolani et al. 2012) to annotate the outlier SNPs. We used the annotated variants as input in UniProt (The UniProt Consortium 2025) to obtain protein names and used OrthoDB (Kriventseva et al. 2019) to obtain the functional category of these proteins. We finally ran an enrichment analysis using Fisher's exact test with Benjamini–Hochberg correction (to account for false discovery rates) with the functional categories represented in the outlier genes as the foreground set and the functional categories represented for the entire *P. macdunnoughii* genome to identify the functions that were overrepresented between females with differing host plant preferences.

RESULTS

Hostplant-specific differential gene expression changes with larval performance but not with adult oviposition preference

We found 1489 (Fig. 1a; FDR < 0.01) overall differences in gene expression profiles, of which 910 transcripts (Fig. 1a; dashed lines: $\log_2$FC > 1.5 and FDR < 0.01) showed significantly greater differences between larvae that fed on *T. arvense* versus *C. cordifolia*. Hierarchical clustering of individuals showed that replicates of larvae of each treatment were broadly similar to each other, compared to those from the other treatments (Fig. 2a), suggesting host plant-specific transcriptome patterns. However, two larvae that fed on *T. arvense* clustered with larval profiles from *C. cordifolia*. These transcripts split into three distinct clusters (Fig. 2a). Genes in clusters 1 ($n = 404$ genes; Fig. 2b; SI Table 1; paired *t*-test, $p < 0.001$) were upregulated in larvae

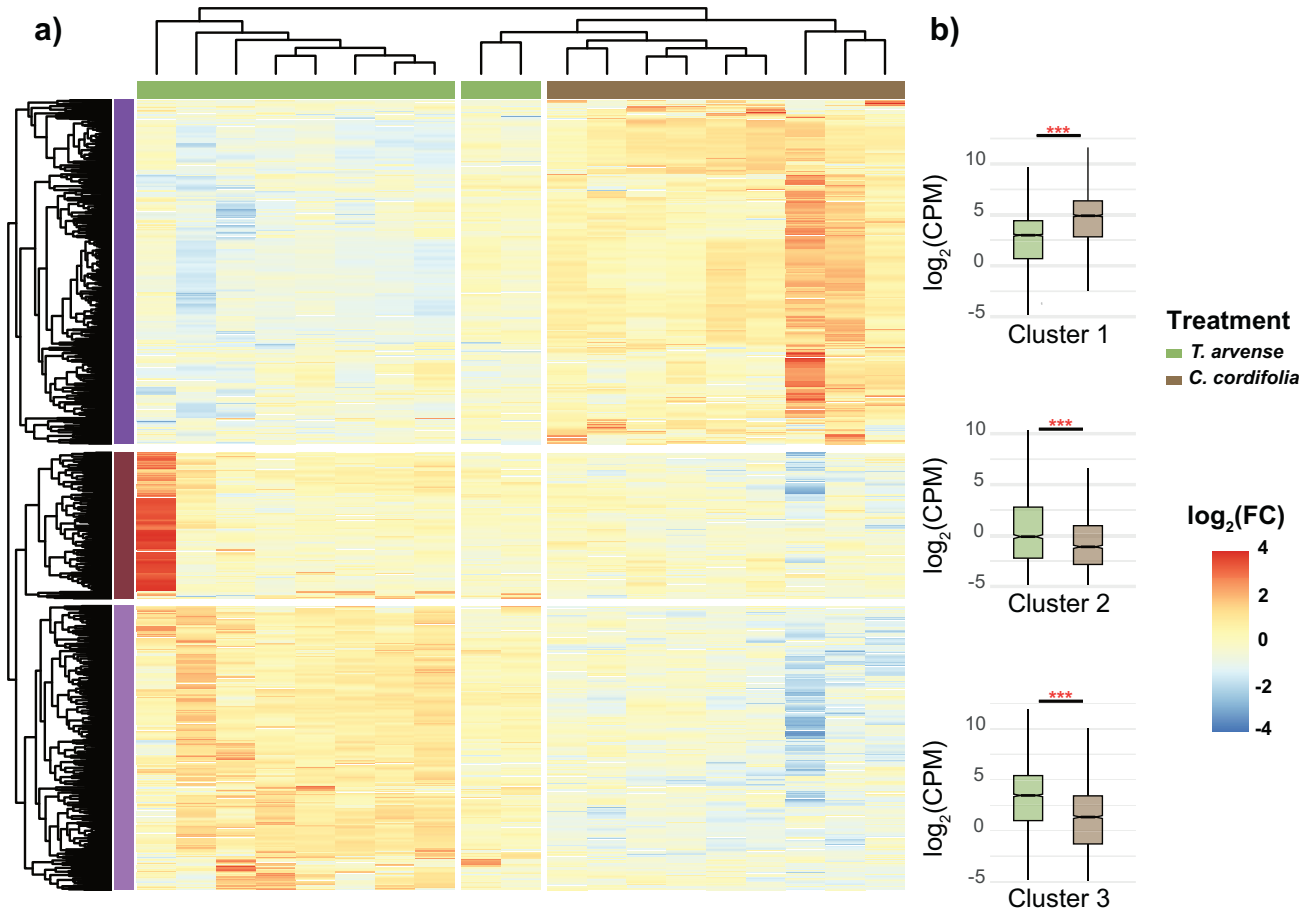

Fig. 2 Transcriptional differences in larval performance on different host plants. a Heatmap showing log (FC) in *Pieris macdunnoughii* larvae in response to feeding on *C. cordifolia* (brown) and *T. arvense* (green). Individuals (columns) and transcripts (rows) are arranged by hierarchical clustering of expression profiles. **b** Individual gene expression profile corresponding to each cluster of gene expression represented in **a**.

that fed on *C. cordifolia*. Genes in cluster 2 ($n = 172$ genes; Fig. 2b; SI Table 1; paired *t*-test, $p < 0.001$) and 3 ($n = 334$ genes; Fig. 2b; SI Table 1; paired t-test, $p < 0.001$) were upregulated in larvae that fed on *T. arvense*.

In contrast, although we observed variation in female oviposition preference (14 females preferred *Cardamine cordifolia*, and 10 females preferred *Thlaspi arvense*. Nine females had equal preference for both plants, and 25 females did not meet the threshold of 50 eggs to quantify preference), the gene expression profiles of female *P. macdunnoughii* preferring *T. arvense* and *C. cordifolia* were similar (Fig. 1b).

Signatures of elevated genetic differentiation across several functions involved in hostplant sensation in females with contrasting oviposition preferences

While our genome-wide F_{ST} and π estimates revealed no genetic differentiation (data not shown), we identified several regions across the genome with significantly elevated genetic differentiation (Fig. 3a–d) in females with contrasting oviposition preferences. We found 898 SNPs spanning 103 windows across the genome with significant genetic differentiation (Fig. 3b). Our SnpEff analysis revealed that these 898 SNPs were distributed across 165 genes. Our OrthoDB analysis classified these 165 genes into 18 functional COG (cluster of orthologs) categories (SI Table 2). We additionally found that all genes were involved in more than one function, with the majority of them significantly enriched for signal transduction (104 genes), followed by posttranslational modification (48 genes) and amino acid transport/metabolism (59 genes) (Fig. 4).

Genes underlying the detoxification of plant compounds differentially expressed between larval diet treatments

While gene expressions across all detoxification gene families between larvae feeding on *C. cordifolia* vs *T. arvense* were similar (SI Table 3; linear mixed model, $p = 0.1$), we found differences in specific gene families between the two treatments. Specifically, the larvae feeding on *C. cordifolia* had significantly higher expression of phase 2 enzymes: Glutathione-S-transferase (Fig. 5; SI Table 3; linear mixed model, $p < 0.001$) and Glucuronosyltransferase (Fig. 5; SI Table 3; linear mixed model, $p = 0.009$) compared to those feeding on *T. arvense*. Gene expression of enzymes in phases 1 and 3 was similar between the larval treatments.

Co-expression of genes indicates host plant-specific functional gene modules in larval performance but not in female oviposition preference

Our network analysis identified thirty-three modules (SI Fig. 4a) for larval performance and thirty-nine modules (SI Fig. 4b) for adult oviposition preference. Five of the thirty-three modules identified for larval performance showed significant correlation with hostplant use (SI Fig. 5a), while none of the thirty-nine modules identified with female oviposition preference were correlated with either of the host plants (SI Fig. 5b).

Three of the five modules in larval performance (module blue, module black, and module green) were positively correlated with *T. arvense*, and two modules (module magenta and module turquoise) were negatively correlated with *T. arvense* (i.e., upregulated in *C. cordifolia*). We identified 2353 genes as significantly associated with *T. arvense* (gene significance vs

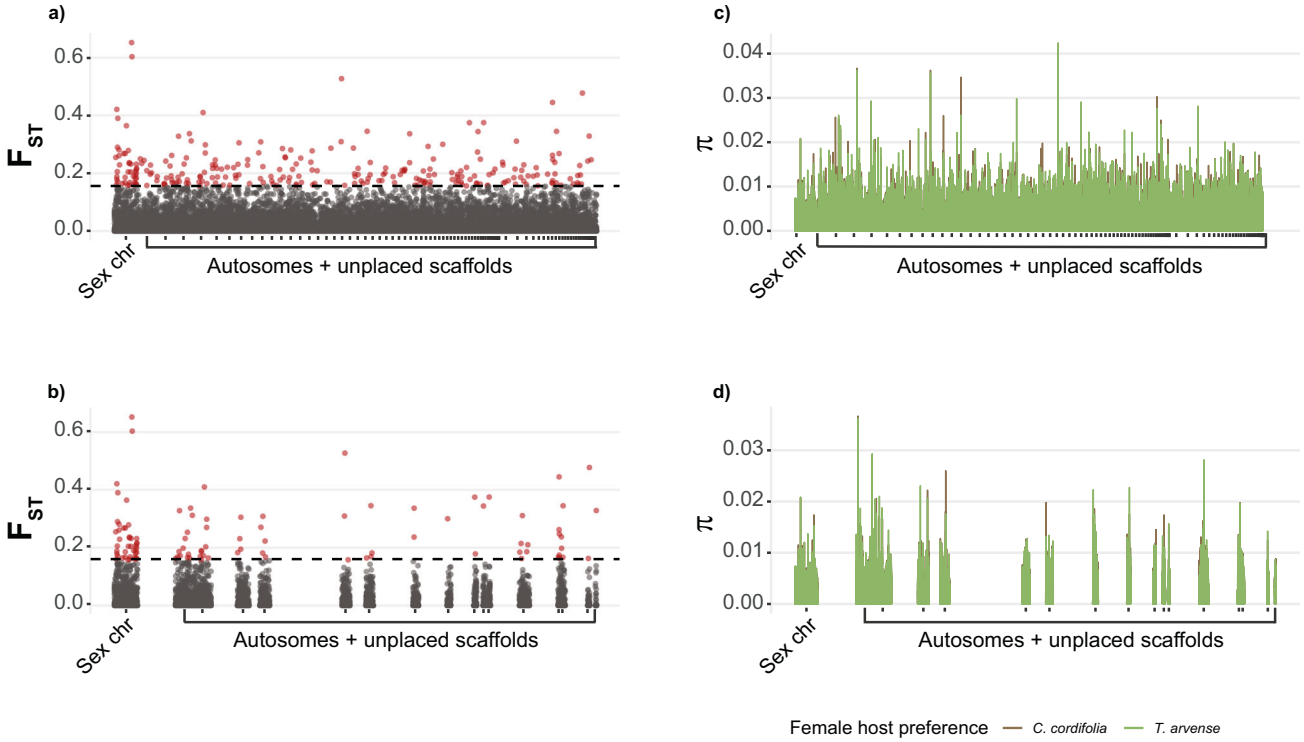

Fig. 3 Genetic differences underlying female oviposition on *T. arvense*. **a** Genome-wide estimates of genetic differentiation (F_{ST}) and **c** genetic diversity (π) between *Pieris macdunnoughii* females of contrasting oviposition preference, calculated using a 5000 bp sliding window approach. **b** Estimates of genetic differentiation (F_{ST}) in scaffolds containing 99th percentile outlier loci, identified by permutation analysis, and **d** genetic diversity estimates (π) underlying outlier loci identified in (**a** and **b**). Outlier loci are highlighted in red in **a** and **b**.

module membership $r^2 > 0.6$ and gene connectivity $r^2 > 0.8$) across the five modules. 1795 genes showed a positive association with *T. arvense* (genes in modules blue, black, and green), and 1921 genes showed a negative association with *T. arvense* (genes in modules magenta and turquoise) (Fig. SI Fig. 6).

Our analysis identified several hub genes of interest (implicated in metabolism, development, and/or immune process) in each treatment (SI Table 4). Broadly, in modules that were positively correlated with feeding on *T. arvense*, genes related to lipid metabolism (Low density lipoprotein receptor-related protein 4 and C2 domain-containing protein), general metabolism (ABC transporter domain-containing protein, dipeptidase, ubiquitin hydrolase 1 and E3 ubiquitin-protein ligase), response to stress (Nicotinic acetylcholine receptor alpha 3 subunit, receptor protein serine/threonine kinase), and immune response (CUB domain-containing protein and EGF-like domain-containing protein) were identified as the top hub genes. In larvae feeding on *C. cordifolia*, genes related to DNA/RNA synthesis (WKF domain-containing protein, Exoribonuclease phosphorolytic domain-containing protein, Ribosomal protein L1, Menin, and Tudor domain-containing protein) and carbohydrate metabolism (GH18 domain-containing protein) were the top hub genes.

Gene set enrichment analysis reveals host-plant specific effects on critical larval traits

Overlap between co-expression modules and differential gene expression (clusters 2 and 3) in larvae feeding on *T. arvense* (SI Table 5) showed that gene ontology terms involved in glucose metabolism, lipid metabolism, amino acid/nitrogen metabolism (protein metabolism, nucleic acid metabolism, response to stress, cell death and apoptosis, immune response and development were enriched.

Overlap between co-expression modules and differential gene expression (cluster 1) in larvae feeding on *C. cordifolia* (SI Table 6) showed that gene ontology terms related to polysaccharide catabolism, regulation of methylation, lipid metabolism, transport of molecules, development and/or cell/organ growth, reproduction, feeding behavior, response to endogenous stimuli, locomotion, signaling and regulation of transcription/translation were enriched.

DISCUSSION

We compared gene expression patterns of third instar larvae of *Pieris macdunnoughii* when feeding on *Thlaspi arvense* (a toxic and invasive plant) with those of larvae feeding on *Cardamine cordifolia* (a native host plant). We also compared gene expression patterns and quantified single-nucleotide polymorphism (SNP) differences between adult females when actively laying eggs on the above two plants. Our analysis revealed 910 host-plant-specific transcripts that were differentially expressed in the larvae between the two treatments. Analysis of plant compound detoxification genes in the larvae indicated that larvae that fed on *C. cordifolia* had higher expression of genes involved in detoxification of specialized plant secondary compounds compared to larvae feeding on *T. arvense*. We additionally found five host-plant-specific functional gene modules: two in larvae feeding on *C. cordifolia* and three in larvae feeding on *T. arvense*. We found that gene expression in the larvae feeding on *C. cordifolia* was enriched for genes related to feeding behavior, polysaccharide breakdown, development, and upregulation of nucleic acid synthesis, suggesting favorable hostplant feeding. Conversely, the gene expression in larvae feeding on *T. arvense* was enriched for response to stress, macromolecule metabolism, apoptosis, and immune response indicating signs of impaired feeding. Finally, even though we saw variation in oviposition choice by the females in our behavioral assays, we did not find any differential gene expression between the females preferring different plant species.

However, we found that several genes underlying signal transduction, post-translational modification, and amino acid transport were significantly differentiated between females with differing hostplant choice. Our results identify the molecular differences driving the decoupling of preference-performance mismatches, where genetic differences and not differential gene regulation drive female oviposition preference, while host plant performance is driven by genetic differences as well as differential gene regulation in the presence of a novel resource. Overall, our work provides new insights into the consequences of novel resources on native herbivores and highlights ontogenetically specific stages where adaptation is likely to occur.

Larval responses to *Thlaspi arvense*

Our analysis revealed contrasting transcriptomic patterns in larvae feeding on *C. cordifolia* compared to *T. arvense*. We found that larvae feeding on *C. cordifolia* exhibited signs of adaptive feeding as expected, while larvae feeding on *T. arvense* showed signs of toxicity and inefficient metabolism of compounds present in *T. arvense*.

Our analysis of genes involved in detoxification highlighted these patterns. Larvae that fed on *C. cordifolia* had higher expression of a phase 2 enzymes, GST and UDPGT, compared to those feeding on *T. arvense*. This suggests that *P. macdunnoughii* larvae have evolved efficient mechanisms to detoxify the compounds present in *C. cordifolia*, but they are unable to do the same with *T. arvense*. In addition, larvae are less likely to feed on *T. arvense*, and when they do consume the plant, they eventually feed slowly and consume less biomass compared to the native plants, indicating a combination of feeding deterrents and toxins in the plant (Steward et al. 2019). Our results provide additional support to this finding, where we see several hallmarks of higher gene expression related to response to stress, response to starvation, response to inflammation, and cell apoptosis. Although previous research has suggested that the deterrent in *T. arvense* is a result of higher concentration of sinigrin in the plant (Steward et al. 2019), the gene expression profiles of larvae feeding on *T. arvense* leads us to hypothesize that the failure of larvae to develop on the plant is a result of complex interactions among various defenses of the plant, which also involve sinigrin. Future work involving a combination of transcriptome responses of *T. arvense* in response to *P. macdunnoughii* herbivory, and *P. macdunnoughii* response to various concentrations of sinigrin and/or methanol extracts of *T. arvense* plant compounds will help in identifying the nature of deterrent and the toxins in the plant that is responsible for poor larval performance on the plant.

Our analysis of enriched genes during larval feeding provided additional insights into the specific processes underlying differences in performance on both host plants. The larvae that fed on *C. cordifolia* had genes enriched that suggested successful detoxification of plant compounds and adaptive feeding. For example, genes related to feeding behavior and polysaccharide catabolism were enriched, suggesting successful breakdown of complex carbohydrates present in the tissues of *C. cordifolia* (Watanabe and Tokuda 2010). In addition, several genes related to macromolecule metabolism (including lipid, linolic acid, 5-methylcytosine, and protein depolymerization) were enriched, suggesting that larvae were successfully able to assimilate nutrients present in the native host plant. Subsequently, genes related to nucleic acid metabolism were enriched, suggesting that cellular processes were not disrupted when feeding on *C. cordifolia*. Lastly, genes related to humoral immune response were also enriched. Immune cell proliferation is also known to be upregulated in other phytophagous insects when feeding on suitable host plants (Klemola et al. 2007; Diamond and Kingsolver 2010).

We identified several densely connected key genes that were associated with larvae feeding on *C. cordifolia*. Our analysis revealed that N-acetyltransferase domain-containing protein (NAT) and uS12 prolyl 3-hydroxylase (PHD) were the top hub genes in the modules associated with *C. cordifolia* feeding larvae. NAT is involved in the transacetylation of acetyl-CoA to arylalkylaminase. NAT knockouts in several insects have indicated that it is involved in melatonin synthesis, aromatic neurotransmitter inactivation, and cuticle sclerotization. Mutant insects lacking NAT protein have aberrant pigmentation and improper cuticle formation (Noh et al. 2016). PHD proteins possess post-translational activity and are involved in many functions, including the regulation of oxygen and regulation of collagen stability. Normal expression of PHD proteins is critical for preventing cellular hypoxia and is a critical protein in development (Fong and Takeda 2008). Other genes that were identified as top hub genes were involved in normal cellular processes, including cell cycle regulation, DNA/RNA synthesis, and chromosome structure/function.

In contrast, larvae that fed on *T. arvense* had genes enriched that indicated poor performance and hostplant-induced toxicity in the larvae. For example, genes related to general stress response, oxidative stress response, response to reactive oxygen species, and acute inflammatory responses were enriched when larvae fed on *T. arvense*. Insect herbivores experience oxidative stress from the secondary compounds present in their host plants (Aucoin et al. 1995; Bi and Felton 1995). However, insects adapted to their host plants counteract the oxidative stress through various mechanisms, including efficient detoxification through phase 2 enzymes, employing superoxide dismutase and increased levels of reduced glutathione (Velki et al. 2011). Larvae that are not adapted to novel host plants are unable to efficiently employ the above-mentioned strategies and hence experience high levels of oxidative stress and accumulation of reactive oxygen species (ROS), which can lead to reduced performance and higher mortality. The larvae feeding on *T. arvense* also had genes related to apoptotic processes enriched, further supporting the toxic effects of ROS induced by *T. arvense* on cell death. Apart from these, genes related to immune cell differentiation were enriched. Natural killer cell differentiation and interleukin-12 production were particularly enriched in larvae when feeding on *T. arvense*. These responses indicate immune responses to fungal and/or bacterial pathogens in the midgut (Shikano 2017). Plant defenses are known to interact with gut microbiome and can alter the composition from mutualistic bacteria to pathogenic bacteria, especially when larvae feed on low-quality or non-host species, by initiating leaky gut syndrome (Hammer and Bowers 2015; Mason et al. 2019). Finally, genes related to metamorphosis, including ecdysteroid production, response to ecdysone, and morphogenesis, were upregulated. We hypothesize that larvae feeding on *T. arvense* respond to the toxicity of the host plant by shortening their development time to escape suboptimal host plant quality. Previous research in many lepidopteran insects indicated that larvae respond to unfavorable conditions by shortening their developmental period and eclosing with smaller adult body size (Pöykkö and Hyvärinen 2012; Salgado and Saastamoinen 2019; Seehausen et al. 2024).

In larvae that fed on *T. arvense*, densely connected genes included low-density lipoprotein receptor-related protein 4 (LDL 4), J domain-containing protein, and C2H2-type domain protein in the functional gene modules associated with *T. arvense*. LDL 4 is involved in the transport of cellular cholesterol and is a crucial component in the production of ecdysone, suggesting that larvae feeding on *T. arvense* showed accelerated development (Zhao et al. 2024). J domain-containing protein is part of the hsp70 chaperone machinery involved in several cellular functions, including protein folding, transport of polypeptides across organelles, degrading proteins, and coordinating responses to stress (Zhang et al. 2014). C2H2-type domain protein belongs to the zinc finger domain commonly found in transcription factors and is involved in regulating gene expression. These proteins play

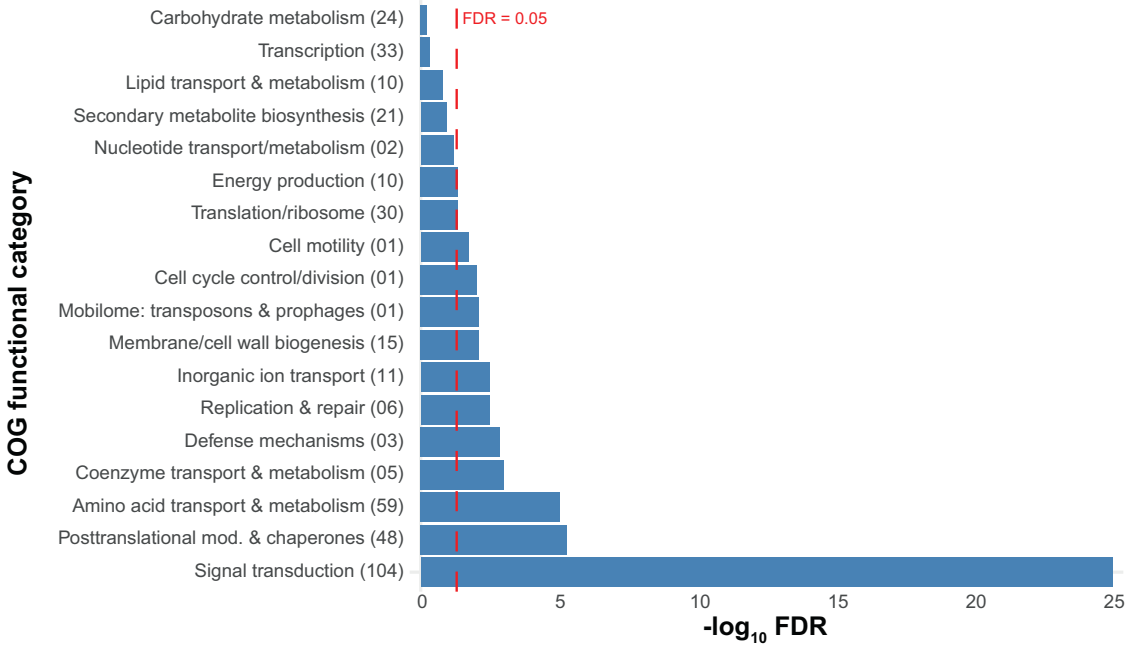

Fig. 4 Enrichment analysis of COG (Clusters of Orthologs) functional categories identified by OrthoDB for genes underlying differences in *Pieris macdunnoughii* female oviposition preference on *C. cordifolia* and *T. arvense*. Numbers indicated in parentheses reflect the number of genes identified for the respective categories. The red dashed line indicates significance at FDR of 0.05, based on Fisher's exact test between categories represented in the oviposition preference dataset and the entire genome.

an important role in regulating gene expression in response to environmental stress (Guo et al. 2021). Together, these hub genes indicate that larvae feeding on *T. arvense* are unable to process the compounds present in the plant and are upregulating genes involved in stress response and inflammation. For instance, *Pieris oleracea*, a congener of *P. macdunnoughii*, experiences a similar evolutionary trap while interacting with *Alliaria petiolata* (another invasive mustard), where larvae are unable to feed on the plant, specifically during the plant's rosette stage (Rodman and Chew 1980; Huang et al. 1994; Keeler and Chew 2008). Later experiments revealed that two glycosides, isovitexin 6″-O-glucoside and allirinoside, which acted as feeding deterrents, were responsible for the poor larval performance on the plant (Haribal and Renwick 1998; Haribal et al. 2001). These findings, in combination with current and previous experiments on *P. macdunnoughii larvae*, indicate that feeding deterrents coupled with higher concentrations of glucosinolates can lead to higher expression of genes related to stress response, starvation, inflammation, and cell apoptosis.

Female responses to *Thlaspi arvense*

Pieris macdunnoughii has been experiencing *T. arvense* for at least 95 generations (earliest herbarium record of *T. arvense* in Gunnison Valley) and possibly for 150+ generations (since *T. arvense* was first introduced to the Gunnison Valley) (Best and McIntyre 1975; Chew 1977a). Given that the fitness costs of laying eggs on *T. arvense* are high (Nakajima et al. 2013), the ability of females to differentiate between the native plants and *T. arvense* should be under strong selection. Previous research has identified sex-linked heritable additive genetic variation in oviposition preference of *P. macdunnoughii* females on *T. arvense* (Steward et al. 2022). However, the sex-linked heritability was only significant when tested on whole plants or cut stems, but not on leaf extracts of *T. arvense*, suggesting that other factors in addition to differences in glucosinolates between the host plants were driving oviposition preference. Our current analysis of female transcriptome and genomic responses to oviposition on native plants and *T. arvense* revealed that differences in oviposition preference were driven by differences in genes primarily underlying signal transduction and post-translational modification, and not due to differences in gene expression between females with contrasting oviposition preference or due to allelic/gene expression differences in NSP/MA alleles (Fig. 4a–d). Additionally, our analysis revealed that more than a third of the SNP differences were located on the sex chromosome (315/898 SNPs) (Fig. 4b). These results in combination with our current findings highlight that both sex-linked as well as autosomal allelic differences, independent of differences in NSP/ MA proteins (which respond to glucosinolates in the host plants (Wheat et al. 2007; Edger et al. 2015)), influence female oviposition preference within the East River valley population.

Although lepidopteran females use a combination of olfactory, visual, and gustatory cues to lay eggs on suitable host plants, the ultimate step involves gustatory recognition (Pivnick et al. 1994; Steward and Boggs 2020). Gustatory receptors are localized in the forelegs of the females, and olfactory and visual receptors are located on the antennae and eyes, respectively (McIndoo 1929; Xu 2020). Thus, differences in oviposition preference should translate to heritable changes underlying genes of the sensory system. Additionally, differences in oviposition preference can be driven by differences in sensitivity to the oviposition stimulants present in the host plants. Electrophysiological readings from antennae and fore tarsi in various herbivorous insects during oviposition have demonstrated differences in neural excitation activity when females experience various oviposition stimulants and can differentiate between relative concentrations of the stimulants (Den Otter et al. 1978; Rojas 1999; Beck et al. 2014).

Since a majority of the allelic differences in our analysis underlie signal transduction, post-translational modifications, and/or amino acid transport (Fig. 5), females may experience either altered or no stimulation if they land on a host plant for which they do not possess the required sensory alleles. Future experiments, including those that measure the neural excitation of gustatory and antennal receptors in females preferring different host plants, will help identify the role of these allelic differences in shaping female sensory pathways in response to different host plants.

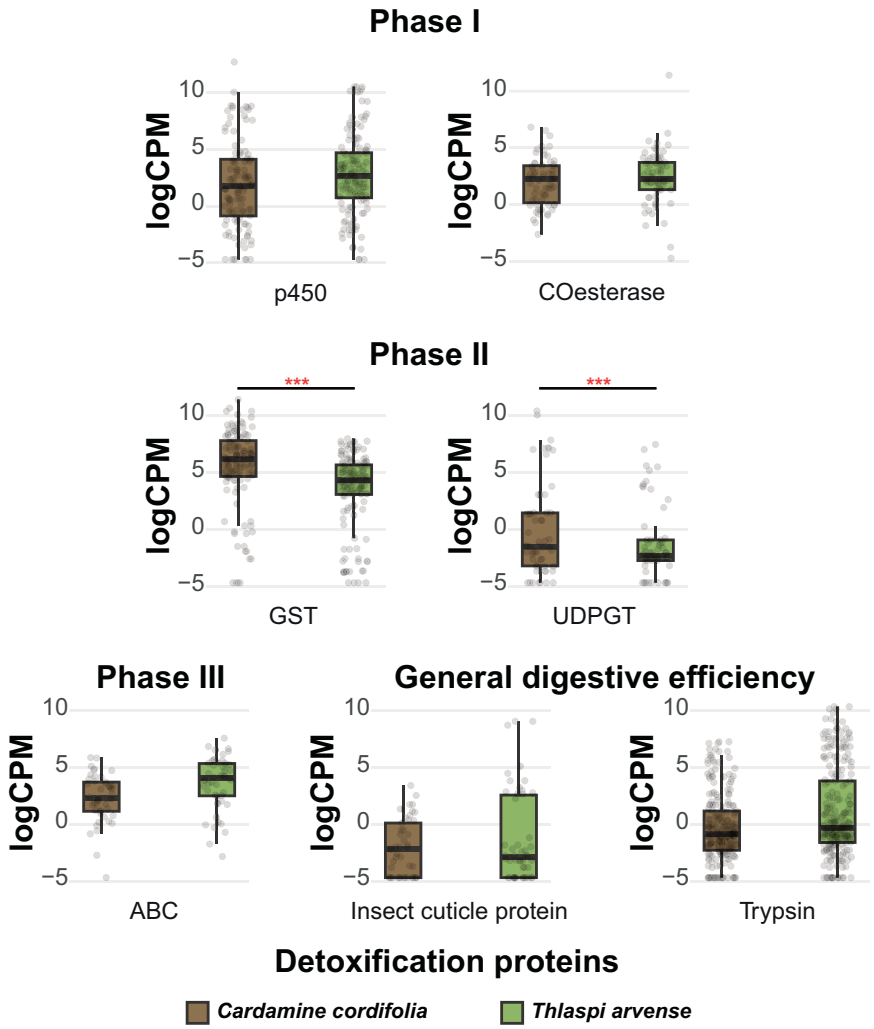

Fig. 5 Gene expression of various plant compound detoxification genes grouped based on Pfam database annotation between *Pieris macdunnoughii* larvae feeding on *T. arvense* and *C. cordifolia*.

Implications for preference-performance hypothesis and the outcome of the evolutionary trap

Our results demonstrate the decoupling of female oviposition preference and larval performance in the presence of *T. arvense*. The fitness costs related to laying eggs on *T. arvense* is dependent on the fine-grained spatial structure of *T. arvense* in the habitat and its proximity to native host plants (Nakajima et al. 2013; Nakajima and Boggs 2015). *Thlaspi arvense* prefers dry habitats and occurs in disturbed environments (Warwick et al. 2002). The abundance and range of *T. arvense* could increase due to continued anthropogenic disturbance in native habitats, thus increasing the frequency at which *P. macdunnoughii* females encounter the invasive plant. Our current and published results in this system show that genetic differences alone drive females' preference for *T. arvense* over *C. cordifolia* for oviposition, while larvae exhibit both host-plant-specific transcriptomic and genetic differences. These results suggest that selection can act on multiple components of preference-performance in this system. For example, selection can act to reduce the frequency of females with alleles associated with preference for *T. arvense* over time. Additionally, selection can simultaneously or independently act on larval ability to feed on *T. arvense* either through modulating gene expression of critical detoxification enzymes on *T. arvense* or through increasing the frequency of alleles that enable larvae to develop successfully on *T. arvense*. Support for female oviposition preference leading to adaptation to novel resources has been documented in several systems, including the formation of host races in the dipteran fly *Rhagoletis pomonella* (Linn et al. 2003; Hood et al. 2020), in several lepidopteran species (Jahner et al. 2011), including *Polygonia spp*. (Nylin and Janz 1993), *Euphydryas editha* (Singer et al. 1988; Singer and Parmesan 2019), *Spodoptera frugiperda* (Groot et al. 2010), and distinct host races in pea aphids (Ferrari et al. 2006; Peccoud et al. 2009), which in extreme cases can lead to speciation. Alternatively, several cases where repeated error-prone oviposition by females has resulted in larvae incorporating the novel host plant have been documented, including in the related species *Pieris oleracea* and its interaction with the invasive mustard *Alliaria petiolata*. After decades of maladaptation on the plant, *P. oleracea* larvae are now able to develop on *A. petiolata* during its bolting stage but not the rosette stage (Huang et al. 1994; Haribal et al. 2001). Indeed, recent population genomic analysis of *P. macdunnoughii* from habitats with and without *T. arvense* has shown signatures of selection in *P. macdunnoughii* in response to *T. arvense* in areas where they co-occur. The loci that were under selection were identified to be those related to metabolism and larval development, providing support for selection acting on larval survival (Ravikanthachari et al. 2024). One potential route through which larvae can escape the evolutionary trap in our case, would be for the larvae to survive longer on *T. arvense*. Several

experiments in the lab have demonstrated variation in time to death in the larvae while feeding on *T. arvense* (Boggs C. L., unpublished) and that larvae can survive longer on *T. arvense* plants that experience heat and drought stress (Ravikanthachari N., unpublished). Additionally, two larvae feeding on *T. arvense* in our experiment exhibited variation in their gene expression profiles, where they clustered with those larvae feeding on *C. cordifolia*, suggesting variation in tolerance to *T. arvense* among larvae feeding on the invasive plant. These instances, coupled with proximity to a native plant (Nakajima et al. 2013; Nakajima and Boggs 2015), might result in some larvae moving to suitable native host plants in the later instars, resulting in larval rescue from the evolutionary trap.

DATA AVAILABILITY

The raw reads used in the project are deposited at NCBI under the Bioproject accession id PRJNA1076332. The source code and associated data files are available on GitHub https://github.com/nitinra/rna_seq_pieris_preference_performance#.

REFERENCES

Alexa A, Rahnenführer J (2009) Gene set enrichment analysis with topGO. Biocond Improv 27:1–26.

Andrews S (2010) A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Aucoin R, Guillet G, Murray C, Philogène BJR, Arnason JT (1995) How do insect herbivores cope with the extreme oxidative stress of phototoxic host plants?. Arch Insect Biochem Physiol 29:211–226.

Balkenius A, Hansson B (2012) Discrimination training with multimodal stimuli changes activity in the mushroom body of the Hawkmoth Manduca sexta. PLOS One 7:e32133.

Barth M, Heisenberg M (1997) Vision affects mushroom bodies and central complex in Drosophila melanogaster. Learn Mem 4:219–229.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H et al. (2015) Package 'lme4'. convergence 12:2.

Beck JJ, Light DM, Gee WS (2014) Electrophysiological responses of male and female Amyelois transitella antennae to pistachio and almond host plant volatiles. Entomol Exp Appl 153:217–230.

Berdegué M, Reitz SR, Trumble JT (1998) Host plant selection and development in *Spodoptera exigua*: do mother and offspring know best?. Entomol Exp Appl 89:57–64.

Best KF, McIntyre GI (1975) The biology of Canadian weeds: 9. *Thlaspi arvense* L. Can J Plant Sci 55:279–292.

Bi JL, Felton GW (1995) Foliar oxidative stress and insect herbivory: Primary compounds, secondary metabolites, and reactive oxygen species as components of induced resistance. J Chem Ecol 21:1511–1530.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Bossart JL, Scriber JM (1999) Preference variation in the polyphagous Tiger Swallowtail Butterfly (Lepidoptera: Papilionidae). Environ Entomol 28:628–637.

Breeschoten T, Ros VID, Schranz ME, Simon S (2019) An influential meal: host plant dependent transcriptional variation in the beet armyworm, *Spodoptera exigua* (Lepidoptera: Noctuidae). BMC Genomics 20: 845.

Breeschoten T, Schranz ME, Poelman EH, Simon S (2022) Family dinner: Transcriptional plasticity of five Noctuidae (Lepidoptera) feeding on three host plant species. Ecol Evol 12:e9258.

Calla B, Noble K, Johnson RM, Walden KKO, Schuler MA, Robertson HM et al. (2017) Cytochrome P450 diversification and hostplant utilization patterns in specialist and generalist moths: Birth, death and adaptation. Mol Ecol 26:6021–6035.

Carrasco D, Larsson MC, Anderson P (2015) Insect host plant selection in complex environments. Curr Opin Insect Sci 8:1–7.

Casagrande RA, Dacey JE (2007) Monarch Butterfly Oviposition on Swallow-Worts (*Vincetoxicum spp*.). Environ Entomol 36:631–636.

Celorio-Mancera M, de la P, Wheat CW, Huss M, Vezzi F, Neethiraj R, Reimegård J et al. (2016) Evolutionary history of host use, rather than plant phylogeny, determines gene expression in a generalist butterfly. BMC Evol Biol 16: 59.

Chew FS (1977a) The effects of introduced mustards (Cruciferae) on Some Native North American Cabbage Butterflies (Lepidoptera: Pieridae). Atala 5:13–19.

Chew FS (1977b) Coevolution of Pierid Butterflies and Their Cruciferous Foodplants. II. The Distribution of Eggs on Potential Foodplants. Evolution 31:568–579.

Chew FS (1980) Foodplant preferences of Pieris caterpillars (Lepidoptera). Oecologia 46:347–353.

Chew FS, Watt WB (2006) The green-veined white (*Pieris napi* L.), its Pierine relatives, and the systematics dilemmas of divergent character sets (Lepidoptera, Pieridae). Biol J Linn Soc 88:413–435.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms. Fly 6:80–92.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158.

David WAL, Gardiner BOC (1966) The effect of Sinigrin on the feeding of *Pieris brassicae* L. larvae transferred from various diets. Entomol Exp Appl 9:95–98.

Den Otter CJ, Schuil HA, Oosten AS-V (1978) Reception of host-plant odours and female sex pheromone in Adoxophyes Orana (lepidoptera: Tortricidae): Electrophysiology and morphology. Entomol Exp Appl 24:570–578.

Dermauw W, Van Leeuwen T (2014) The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. Insect Biochem Mol Biol 45:89–110.

Dermauw W, Van Leeuwen T, Feyereisen R (2020) Diversity and evolution of the P450 family in arthropods. Insect Biochem Mol Biol 127:103490.

Diamond SE, Kingsolver JG (2010) Host plant quality, selection history and trade-offs shape the immune responses of *Manduca sexta*. Proc R Soc B: Biol Sci 278:289–297.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S et al. (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

Dobin A, Gingeras TR (2015) Mapping RNA-seq Reads with STAR. Curr Protoc Bioinforma 51:11.14.1–11.14.19.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE et al. (2015) The butterfly plant arms-race escalated by gene and genome duplications. Proc Natl Acad Sci 112:8362–8366.

Ehrlich PR, Raven PH (1964) Butterflies and plants: a study in coevolution. Evolution 18:586–608.

Erbilgin N, Ma C, Whitehouse C, Shan B, Najar A, Evenden M (2014) Chemical similarity between historical and novel host plants promotes range and host expansion of the mountain pine beetle in a naïve host ecosystem. N Phytol 201:940–950.

Ferrari J, Godfray HCJ, Faulconbridge AS, Prior K, Via S (2006) Population differentiation and genetic variation in host choice among Pea Aphids from eight host plant genera. Evolution 60:1574–1584.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR et al. (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230.

Fong G-H, Takeda K (2008) Role and regulation of prolyl hydroxylase domain proteins. Cell Death Differ 15:635–641.

Futuyma DJ, Agrawal AA (2009) Macroevolution and the biological diversity of plants and herbivores. Proc Natl Acad Sci 106:18054–18061.

García-Robledo C, Horvitz CC (2012) Parent–offspring conflicts, "optimal bad motherhood" and the "mother knows best" principles in insect herbivores colonizing novel host plants. Ecol Evol 2:1446–1457.

Gripenberg S, Mayhew PJ, Parnell M, Roslin T (2010) A meta-analysis of preference-performance relationships in phytophagous insects. Ecol Lett 13:383–393.

Groen SC, Whiteman NK (2022). Ecology and Evolution of Secondary Compound Detoxification Systems in Caterpillars. In: Marquis RJ, Koptur S (eds) *Caterpillars in the Middle: Tritrophic Interactions in a Changing World*, Springer International Publishing: Cham, pp 115–163

Groot AT, Marr M, Heckel DG, Schöfl G (2010) The roles and interactions of reproductive isolation mechanisms in fall armyworm (Lepidoptera: Noctuidae) host strains. Ecol Entomol 35:105–118.

Guo H, Wang L, Wang C, Guo D, Xu B, Guo X et al. (2021) Identification of an Apis cerana zinc finger protein 41 gene and its involvement in the oxidative stress response. Arch Insect Biochem Physiol 108: e21830.

Hale R, Morrongiello JR, Swearer SE (2016) Evolutionary traps and range shifts in a rapidly changing world. Biol Lett 12:2016–2019.

Hammer TJ, Bowers MD (2015) Gut microbes may facilitate insect herbivory of chemically defended plants. Oecologia 179:1–14.

Hardy NB, Otto SP (2014) Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. Proc R Soc B: Biol Sci 281:20132960.

Haribal M, Renwick JAA (1998) Isovitexin 6″-O-β-d-glucopyranoside: A feeding deterrent to *Pieris napi oleracea* from *Alliaria petiolata*. Phytochemistry 47:1237–1240.

Haribal M, Yang Z, Attygalle AB, Renwick JAA, Meinwald J (2001) A Cyanoallyl Glucoside from *Alliaria petiolata*, as a feeding deterrent for larvae of *Pieris napi oleracea*. J Nat Prod 64:440–443.

Heckel DG (2014) Insect Detoxification and Sequestration Strategies. Annu Plant Rev 47:77–114.

Heidel-Fischer HM, Vogel H (2015) Molecular mechanisms of insect adaptation to plant secondary compounds. Curr Opin Insect Sci 8:8–14.

Hood GR, Powell THQ, Doellman MM, Sim SB, Glover M, Yee WL et al. (2020) Rapid and repeatable host plant shifts drive reproductive isolation following a recent human-mediated introduction of the apple maggot fly, *Rhagoletis pomonella*. Evolution 74:156–168.

Huang XP, Renwick JAA, Chew FS (1994) Oviposition stimulants and deterrents control acceptance of *Alliaria petiolata* by *Pieris rapae* and *P. napi oleracea*. Chemoecology 5:79–87.

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von, Mering C et al. (2017) Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. Mol Biol Evol 34:2115–2122.

Jaenike J (1978) On optimal oviposition behavior in phytophagous insects. Theor Popul Biol 14:350–356.

Jaenike J (1990) Host specialization in phytophagous insects. Annu Rev Ecol Syst 21:243–273.

Jahner JP, Bonilla MM, Badik KJ, Shapiro AM, Forister ML (2011) Use of exotic hosts by Lepidoptera: Widespread species colonize more novel hosts. Evolution 65:2719–2724.

Janz N, Nylin S, Wedell N (1994) Host plant utilization in the comma butterfly: sources of variation and evolutionary implications. Oecologia 99:132–140.

Joshi A, Thompson JN (1995) Trade-offs and the evolution of host specialization. Evol Ecol 9:82–92.

Keeler MS, Chew FS (2008) Escaping an evolutionary trap: preference and performance of a native insect on an exotic invasive host. Oecologia 156:559–568.

Klemola N, Klemola T, Rantala MJ, Ruuhola T (2007) Natural host-plant quality affects immune defence of an insect herbivore. Entomol Exp Appl 123:167–176.

Kolde R, Kolde MR (2015) Package 'pheatmap'. R package 1:790.

Korunes KL, Samuk K (2021) Pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol Ecol Resour 21:1359–1368.

Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA et al. (2019) OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res 47:D807–D811.

Ladner DT, Altizer S (2005) Oviposition preference and larval performance of North American monarch butterflies on four *Asclepias* species. Entomol Exp Appl 116:9–20.

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinforma 9: 559.

Larsson S, Ekbom B (1995) Oviposition mistakes in herbivorous insects: confusion or a step towards a new host plant?. Oikos 72:155–160.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930.

Liao Y, Smyth GK, Shi W (2019) The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res 47:e47.

Linn C, Feder JL, Nojima S, Dambroski HR, Berlocher SH, Roelofs W (2003) Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. Proc Natl Acad Sci 100:11490–11493.

Mason CJ, Ray S, Shikano I, Peiffer M, Jones AG, Luthe DS et al. (2019) Plant defenses interact with insect enteric bacteria by initiating a leaky gut syndrome. Proc Natl Acad Sci 116:15991–15996.

McIndoo NE (1929). Tropisms and sense organs of Lepidoptera. *Smithsonian Miscellaneous Collections*.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303.

Menacer K, Cortesero AM, Hervé MR (2021) Challenging the Preference–Performance Hypothesis in an above-belowground insect. Oecologia 197:179–187.

Müller C, van Loon J, Ruschioni S, De Nicola GR, Olsen CE, Iori R et al. (2015) Taste detection of the non-volatile isothiocyanate moringin results in deterrence to glucosinolate-adapted insect larvae. Phytochemistry 118:139–148.

Nakajima M, Boggs CL (2015) Fine-grained distribution of a non-native resource can alter the population dynamics of a native consumer. PLoS ONE 10:1–17.

Nakajima M, Boggs CL, Bailey S, Reithel J, Paape T (2013) Fitness costs of butterfly oviposition on a lethal non-native plant in a mixed native and non-native plant community. Oecologia 172:823–832.

Noh MY, Koo B, Kramer KJ, Muthukrishnan S, Arakane Y (2016) Arylalkylamine N-acetyltransferase 1 gene (TcAANAT1) is required for cuticle morphology and pigmentation of the adult red flour beetle, *Tribolium castaneum*. Insect Biochem Mol Biol 79:119–129.

Nylin S, Bergström A, Janz N (2000) Butterfly host plant choice in the face of possible confusion. J Insect Behav 13:469–482.

Nylin S, Janz N (1993) Oviposition preference and larval performance in *Polygonia c-album* (Lepidoptera: Nymphalidae): the choice between bad and worse. Ecol Entomol 18:394–398.

Nylin S, Janz N (2009) Butterfly host plant range: an example of plasticity as a promoter of speciation?. Evol Ecol 23:137–146.

Orsucci M, Audiot P, Nidelet S, Dorkeld F, Pommier A, Vabre M et al. (2018) Transcriptomic response of female adult moths to host and non-host plants in two closely related species. BMC Evol Biol 18: 145.

Ozaki K, Ryuda M, Yamada A, Utoguchi A, Ishimoto H, Calas D et al. (2011) A gustatory receptor involved in host plant recognition for oviposition of a swallowtail butterfly. Nat Commun 2: 542.

Peccoud J, Ollivier A, Plantegenest M, Simon J-C (2009) A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. Proc Natl Acad Sci, USA 106:7495–7500.

Picard toolkit (2019). *Broad Institute, GitHub repository*.

Pivnick KA, Jarvis BJ, Slater GP (1994) Identification of olfactory cues used in host-plant finding by diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). J Chem Ecol 20:1407–1427.

Pöykkö H, Hyvärinen M (2012) To grow fast or to grow big? Time-limited larvae of *Eilema depressum* speed up their growth and reduce number of instars. Entomol Exp Appl 142:145–152.

R Development Core Team (2008). R: A language and environment for statistical computing.

Rausher MD (1979) Larval habitat suitability and oviposition preference in three related butterflies. Ecology 60:503–511.

Ravikanthachari N, Steward RA, Boggs CL (2024) Patterns of genetic variation and local adaptation of a native herbivore to a lethal invasive plant. Mol Ecol 33:e17326.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47.

Robertson BA, Chalfoun AD (2016) Evolutionary traps as keys to understanding behavioral maladapation. Curr Opin Behav Sci 12:12–17.

Robertson BA, Hutto RL (2006) A framework for understanding ecological traps and an evaluation of existing evidence. Ecology 87:1075–1085.

Robertson BA, Rehage JS, Sih A (2013) Ecological novelty and the emergence of evolutionary traps. Trends Ecol Evol 28:552–560.

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

Rodman JE, Chew FS (1980) Phytochemical correlates of herbivory in a community of native and naturalized Cruciferae. Biochem Syst Ecol 8:43–50.

Rojas JC (1999) Electrophysiological and behavioral responses of the cabbage moth to plant volatiles. J Chem Ecol 25:1867–1883.

Salgado AL, Saastamoinen M (2019) Developmental stage-dependent response and preference for host plant quality in an insect herbivore. Anim Behav 150:27–38.

Seehausen ML, Jones IM, Smith SM, Bourchier RS (2024) Foliage age affects pupal weight and development time, but not diapause induction in the biological control agent *Hypena opulenta* (Lepidoptera: Erebidae). Environ Entomol 54:39–45.

Shikano I (2017) Evolutionary ecology of multitrophic interactions between plants, insect herbivores and entomopathogens. J Chem Ecol 43:586–598.

Singer MC (2021) Preference Provides a Plethora of Problems (Don't Panic). Annu Rev Entomol 66:1–22.

Singer MC, Ng D, Thomas CD (1988) Heritability of oviposition preference and its relationship to offspring performance within a single insect population. Evolution 42:977–985.

Singer MC, Parmesan C (2019) Butterflies embrace maladaptation and raise fitness in colonizing novel host. Evol Appl 12:1417–1433.

Snell-Rood EC (2007). Costs of plasticity in host use in butterflies. Ph.D. Thesis, The University of Arizona.

Sollai G, Biolchini M, Crnjar R (2018) Taste sensitivity and divergence in host plant acceptance between adult females and larvae of *Papilio hospiton*. Insect Sci 25:809–822.

Städler E, Renwick JaA, Radke CD, Sachdev-Gupta K (1995) Tarsal contact chemoreceptor response to glucosinolates and cardenolides mediating oviposition in *Pieris rape*. Physiol Entomol 20:175–187.

Stefanescu C, Jubany J, Dantart J (2012) Egg–laying by the butterfly *Iphiclides podalirius* (Lepidoptera, Papilionidae) on alien plants: a broadening of host range or oviposition mistakes?. Anim Biodivers Conserv 29:83–90.

Steward RA, Boggs CL (2020) Experience may outweigh cue similarity in maintaining a persistent host-plant-based evolutionary trap. Ecol Monogr 90:e01412.

Steward RA, Epanchin-Niell RS, Boggs CL (2022) Novel host unmasks heritable variation in plant preference within an insect population. Evolution 76:2634–2648.

Steward RA, Fisher LM, Boggs CL (2019) Pre- and post-ingestive defenses affect larval feeding on a lethal invasive host plant. Entomol Exp Appl 167:292–305.

Steward RA, Okamura Y, Boggs CL, Vogel H, Wheat CW (2021) The genome of the margined White Butterfly (*Pieris macdunnoughii*): Sex chromosome insights and the power of polishing with PoolSeq Data. Genome Biol Evol 13:evab053.

Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. PLOS ONE 6:e21800.

The UniProt Consortium (2025) UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res 53:D609–D617.

Thompson JN (1999a) Specific hypotheses on the geographic mosaic of coevolution. Am Nat 153:S1–S14.

Thompson JN (1999b) The evolution of species interactions. Science 284:2116–2118.

Trowbridge CD, Todd CD (2001) Host-plant change in marine specialist herbivores: Ascoglossan sea slugs on introduced macroalgae. Ecol Monogr 71:219–243.

Van Dijk LJA, Janz N, Schäpers A, Gamberale-Stille G, Carlsson MA (2017) Experience-dependent mushroom body plasticity in butterflies: consequences of search complexity and host range. Proc R Soc B: Biol Sci 284:20171594.

Van Loon JJA, Blaakmeer A, Griepink FC, van Beek TA, Schoonhoven LM, de Groot A (1992) Leaf surface compound from Brassica oleracea (Cruciferae) induces oviposition by *Pieris brassicae* (Lepidoptera: Pieridae). Chemoecology 3:39–44.

Valladares G, Lawton JH (1991) Host-plant selection in the Holly Leaf-Miner: Does mother know best?. J Anim Ecol 60:227–240.

Velki M, Kodrík D, Večeřa J, Hackenberger BK, Socha R (2011) Oxidative stress elicited by insecticides: A role for the adipokinetic hormone. Gen Comp Endocrinol 172:77–84.

Warwick SI, Francis A, Susko DJ (2002) The biology of Canadian weeds. 9. *Thlaspi arvense* L. (updated). Can J Plant Sci 82:803–823.

Watanabe H, Tokuda G (2010) Cellulolytic systems in insects. Annu Rev Entomol 55:609–632.

Wheat CW, Vogel H, Wittstock U, Braby MF, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant–insect coevolutionary key innovation. Proc Natl Acad Sci 104:20427–20431.

Wiklund C (1975) The evolutionary relationship between adult oviposition preferences and larval host plant range in *Papilio machaon* L. Oecologia 18:185–197.

Xu W (2020) How do moth and butterfly taste?—Molecular basis of gustatory receptors in Lepidoptera. Insect Sci 27:1148–1157.

Yoon S, Read Q (2016) Consequences of exotic host use: impacts on Lepidoptera and a test of the ecological trap hypothesis. Oecologia 181:985–996.

Zars T (2000) Behavioral functions of the insect mushroom bodies. Curr Opin Neurobiol 10:790–795.

Zhang Y, Liu Y, Guo X, Li Y, Gao H, Guo X et al. (2014) sHsp22.6, an intronless small heat shock protein gene, is involved in stress defence and development in Apis cerana cerana. Insect Biochem Mol Biol 53:1–12.

Zhao Y, Liu W, Zhao X, Yu Z, Guo H, Yang Y et al. (2024) Low-density lipoprotein receptor-related protein 2 (LRP2) is required for lipid export in the midgut of the migratory locust, *Locusta migratoria*. J Integr Agric 23:1618–1633.

ACKNOWLEDGEMENTS

We thank M. Olson, H. Walton, E. Wagner, O. Shrestha, and S. McDaniel for assisting with animal rearing. We thank R. Steward for providing advice on differential gene expression analysis. We thank R. Steward, W. Watt, J. Quattro, B. Hollis, P. Andolfatto, and M. Asher for providing comments on earlier versions of the manuscript, and several anonymous reviewers for productive suggestions or comments on the manuscript. This work was supported by the University of South Carolina to CLB and the Rocky Mountain Biological Laboratory (RMBL) graduate student fellowship to NR. Field research permits were arranged by RMBL.

AUTHOR CONTRIBUTIONS

NR and CLB designed the research. NR conducted oviposition choice experiments in 2019 and 2020. NR and CLB conducted larval preference assays in 2019, and NR conducted the assays in 2020. NR did the molecular work and analyzed the data. NR wrote the draft manuscript. NR and CLB contributed to manuscript revisions. CLB provided oversight for all stages of the work.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS

Research ethics committee's approval was not required for this study, as the experimental organisms used were unregulated invertebrate species.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41437-025-00777-7.

Correspondence and requests for materials should be addressed to Nitin Ravikanthachari.