Title

Transfer to different host grapevines causes rapid and widespread transcriptional plasticity in the specialist insect *Daktulosphaira vitifoliae*

Author information

W. Nate Collison [1,2], Syed Adeel Zafar [1], Simon C. Groen [1,2], and Paul D. Nabity [1,3]

[1] Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA, USA

[2] Department of Nematology, University of California, Riverside, Riverside, CA, USA

[3] School of BioSciences, University of Melbourne, Melbourne, Australia

Corresponding Author: Paul D. Nabity, School of BioSciences, University of Melbourne, Melbourne, Australia, Paul.Nabity@unimelb.edu.au

Abstract

Insect herbivores employ many adaptive mechanisms to successfully colonize host plants. Transcriptional plasticity is one important mechanism that enables host use in polyphagous insects that maintain a broad host range. However, whether specialist herbivores also rely on transcriptional plasticity to feed successfully on different hosts remains largely unexplored. Here, we measure genome-wide transcriptional plasticity of grape phylloxera, *Daktulosphaira vitifoliae*, a grapevine (*Vitis* spp.) specialist, after transfer onto different host grapevines for a single generation using RNA-seq. We show that the transcriptome of *D. vitifoliae* is highly plastic and responsive to host genotype – 4,820 genes (37% of the 12,876 expressed genes) are significantly differentially expressed in at least one contrast comparing host genotypes. Genes with functions related to host manipulation (predicted secretory effectors), cuticle and chitin, detoxification of xenobiotics, and development were enriched among host-responsive genes, with predicted secretory effectors being most strongly enriched (Fisher's exact test odds ratio = 8.3, BH-adjusted p-value = 0). We further linked gene co-expression modules to *D. vitifoliae* fitness using weighted gene co-expression network analysis. The co-expression module most strongly correlated with fitness was enriched for predicted secretory effectors, including a novel group of proline-rich genes that exist as a tandem gene cluster. The only module negatively correlated with fitness was enriched for genes encoding glycoside hydrolase enzymes that function in carbohydrate metabolism. Overall, our results show that the phylloxerid transcriptome is highly plastic

and rapidly reprogrammed when transferred onto distinct hosts, which may be an adaptive mechanism underlying host use in this clonally reproducing herbivore.

Key Words

grape phylloxera (*Daktulosphaira vitifoliae*), grapevines (*Vitis* spp.), ecological specialization, host use, transcriptional plasticity, effectors

Introduction

As sessile organisms, plants face significant pressure to evolve effective defenses against a breadth of herbivores. Plant defenses may occur at the morphological level, such as trichomes and thorns, the chemical level, e.g. biosynthesis of toxic defensive metabolites, or an indirect level, e.g. attraction of natural enemies that target herbivores. The evolution of plant defense mechanisms imparts selective pressure on herbivores to reciprocally evolve mechanisms to modulate or evade plant defense. Two general strategies represent key adaptations to dampen plant defense – 1) tolerance or avoidance of defense chemicals via detoxification or sequestration, and (2) direct manipulation of host defenses via delivery of effector molecules into or onto host cells or tissues. In insect herbivores, plant defensive metabolites are often detoxified in two phases [1]. In Phase I, plant-derived toxins are modified into less reactive forms via enzymes such as cytochrome P450 monooxygenases. In Phase II, the toxin derivatives are converted into hydrophilic forms to facilitate sequestration or excretion via UDP-glycosyltransferases or other conjugating enzymes. The use of

66 secretory effector proteins that function as host-manipulating 'molecular weapons' has
67 been well-studied for microbial plant pathogens [2]. More recently, research on molecular
68 plant-insect interactions has revealed that insects also use secretory proteins as host-
69 modifying weapons. This is particularly true for herbivores that maintain intimate
70 associations with host plants, such as sap-feeding and gall-forming insects [3,4]. Although
71 experiments that functionally validate insect effector proteins are limited, these proteins
72 appear to play broad roles in modulating host biological processes including E3-
73 complex-/ubiquitin-mediated protein turnover, calcium signaling, and phytohormone-
74 mediated signaling [3–5].

75 Insect herbivores are among the most species-rich lineages of metazoans, and
76 this diversity is likely driven by insect adaptations to specific host plants [6]. Host-specific
77 adaptations of insects are a mode of ecological specialization that may inflict trade-offs
78 in fitness when an herbivorous insect feeds on ancestral or alternative hosts [7]. This
79 hypothesis is supported by the fact that most insect herbivores are specialists, feeding
80 on one or very few species of closely related hosts. How generalist, polyphagous
81 herbivores maintain compatibility with a breadth of divergent hosts remains largely
82 unknown. One hypothesis is that generalists maintain a high degree of genomic
83 flexibility, which enables polyphagy [8]. Indeed, the genomes of generalist insect
84 herbivores contain extraordinary expansions of genes that function in detoxification of
85 plant defensive metabolites [9,10,]. Additionally, transcriptional plasticity of expanded gene
86 families, including ones encoding detoxification proteins and secretory effectors, has
87 been linked to fitness in generalist insect herbivores during interaction with divergent

host plants [11,12]. Following this same logic, specialist herbivores should possess genomes that are specialized for specific hosts and thus under less pressure to expand and acquire new functions relative to generalists. Transcriptional plasticity of genes that function in host plant use may also be limited or constrained in specialist insect herbivores, although few studies have tested this hypothesis. Here, we investigate the transcriptional basis of host use in grape phylloxera *Daktulosphaira vitifoliae* as it feeds on different host grapevines for a single generation. We generated genome-wide gene expression and fitness data with the aim of (1) understanding how the oligophagous specialist *D. vitifoliae* successfully colonizes the non-native host *Vitis vinifera* (European wine grape), (2) determining which genes show transcriptional plasticity upon transfer to different host grapevines, and (3) correlating transcriptional plasticity with *D. vitifoliae* fitness to determine whether plasticity in expression of different genes is adaptive, neutral, or maladaptive.

Results

D. vitifoliae fitness differs significantly across host grapevine genotypes

To determine the fitness of *D. vitifoliae* on different grapevine hosts, we experimentally introduced a clonal line of *D. vitifoliae* to a panel of grapevine genotypes, which included the non-native host *V. vinifera* (cv Cab 33) and interspecific hybrids of native North American grapevine species (Figure 1A). We collected fitness data repeatedly over an 8-day period starting 15 days after *D. vitifoliae* introduction, when 1st nymphal instars had matured to adulthood and begun reproducing. We found that *D.*

110 *vitifoliae* fitness varied significantly across host grapevine genotypes. Genotypes SW
111 (*V. riparia* X *V. rupestris*), 5BB (*V. riparia* X *V. berlandieri*), and 3309 (*V. riparia* X *V.*
112 *rupestris*) were incompatible with the *D. vitifoliae* genotype used in our study; *D.*
113 *vitifoliae* failed to induce galls and subsequently died. Attempted feeding by *D. vitifoliae*
114 induced small necrotic lesions on leaves of genotypes SW and 5BB within 6 days after
115 experimental infestation (Figure 1D and 1E). Interestingly, genotype 3309 showed
116 symptomless resistance after 6 days (Figure 1F), but eventually showed necrotic
117 lesions in leaves where *D. vitifoliae* had attempted to feed within 3 weeks after
118 infestation (Supplemental Figure 1). *D. vitifoliae* was compatible with the remaining five
119 genotypes including the non-native host *V. vinifera* cv Cab 33. To determine whether *D.*
120 *vitifoliae* fitness was significantly different between any of the five susceptible host
121 genotypes, we performed pair-wise Wilcoxon tests. *D. vitifoliae* fitness was significantly
122 lower on both *V. vinifera* Cab33 and SO4 (*V. berlandieri x V. riparia*) compared to the
123 other three hosts (Figure 1G). All other pairwise comparisons between host genotypes
124 revealed no significant differences in *D. vitifoliae* fitness.
125
126 <u>*D. vitifoliae* gene expression is highly dynamic and responsive to host grapevine</u>
127 <u>genotype</u>
128 To understand the genome-wide variation in *D. vitifoliae* gene expression
129 induced upon transfer onto different host plants, we isolated RNA from *D. vitifoliae*
130 adults after feeding on new hosts for a single generation and performed short-read
131 RNA-seq. Four replicates for each *D. vitifoliae* treatment group were sequenced,

132 totaling 24 samples. These included eight replicates for insects on host *V. vinifera* cv
133 Cab33 across two different *D. vitifoliae* generations. We targeted 30 million 150-bp
134 reads per sample (29 +/- 3 million) and 85% of reads uniquely mapped to gene models
135 in the reference genome after basic quality and adapter trimming (Supplemental Table
136 1). Gene count matrices were transformed and normalized using a standard pipeline in
137 edgeR [13], then unsupervised clustering of samples was plotted using multidimensional
138 scaling (MDS) (Figure 1H). Sample clustering indicated host environment induces
139 distinct changes in *D. vitifoliae* gene expression; however, the SO4 genotype group
140 showed high variability relative to the other 4 treatment groups.

141 We then performed differential gene expression analysis to identify genes that
142 vary in expression across different host genotypes. We performed pairwise contrasts for
143 each genotype-by-genotype comparison and concatenated the results to compile a list
144 of genes differentially expressed in one or more contrasts, i.e. host-responsive genes.
145 Genes were considered differentially expressed based on a 1.5-fold-change difference
146 in expression and adjusted p-value cutoff of 0.05. This resulted in 4,820 host-responsive
147 genes (Figure 2A), which is 37% of the 12,876 expressed genes in our dataset. We
148 then performed a Fisher's exact test to identify gene functional groups that are enriched
149 among the host-responsive genes relative to the genome-wide background, which
150 included all 25,808 *D. vitifoliae* gene models. Genes encoding predicted secretory
151 effectors (adj. p-value = 0, odds ratio = 3.8), cuticle- and chitin-associated proteins (adj.
152 p-value = 2.2e-47, odds ratio = 17.9), developmental regulators (adj. p-value = 2.2e-12,
153 odds ratio = 3.8), and genes with predicted functions in detoxification of plant

154 metabolites and other environmental molecules (adj. p-value = 2.1e-10, odds ratio =
155 2.7) were enriched among host-responsive genes (Figure 2A).
156 Pair-wise analysis of gene differential expression (DE) revealed that host *V.*
157 *vinifera* cv Cab33 induced the strongest change in gene expression compared to other
158 hosts; 4,817 out of the total 4,820 host-responsive genes show DE in pairwise contrasts
159 between *V. vinifera* cv Cab33 and the other four host genotypes (Figure 2B).
160 Furthermore, 227 genes show DE in all contrasts with *V. vinifera* cv Cab33 (Figure 2B).
161 Genes encoding cuticle and chitin associated proteins are significantly enriched among
162 the 227 overlapping DE genes (adj. p-value = 4.1e-38, odds ratio = 50.1) (Supplemental
163 Figure 2A). Enrichment analysis for Kyoto Encyclopedia of Genes and Genomes
164 (KEGG) terms showed that genes encoding oxidoreductase enzymes are enriched
165 among the 227 overlapping DE genes (Supplemental Figure 2B). We also investigated
166 the types of genes that are up- and down-regulated when *D. vitifoliae* feeds on *V.*
167 *vinifera* cv Cab33. Interestingly 3,933 (81.6 %) of the host-responsive genes are
168 upregulated during interaction with *V. vinifera* cv Cab33. Genes encoding predicted
169 secretory effectors, cuticle and chitin associated proteins, and detoxication enzymes are
170 enriched among genes upregulated on *V. vinifera* cv Cab33 (Supplemental Figure 3).
171 Enriched KEGG terms for upregulated genes include peroxidase and oxidoreductase
172 enzymes, sugar and organic acid transporters, integral membrane ion channels, and
173 fatty acid elongation enzymes (Supplemental Figure 4). Genes downregulated during
174 interaction with *V. vinifera* cv Cab33 are enriched for functions related to development
175 (Supplemental Figure 3) and carbohydrate metabolism (Supplemental Figure 4).

176 Overall, these results suggest that *D. vitifoliae* transcript expression is highly plastic,
177 especially at genes with predicted functions in host manipulation (secretory effectors)
178 and interaction with the environment.
179
180 Four gene co-expression modules co-vary with *D. vitifoliae* fitness
181 To identify groups of genes that are strongly correlated in expression pattern and
182 likely function together in shared biological processes, we performed weighted gene co-
183 expression network analysis using WGCNA [14]. We included 13,917 genes for network
184 construction after removing 'zero genes' that showed negligible expression (only genes
185 with ≥0.5 counts per million gene counts (CPM) in at least 4 samples were kept). We
186 constructed a signed co-expression network with 88 modules based on hierarchical
187 clustering of gene connectivity dissimilarity (Figure 3A-B). To identify groups of co-
188 regulated genes (modules) that might underlie differences in *D. vitifoliae* fitness across
189 host genotypes, we tested for correlation between module eigengenes (first principal
190 component of the module and the most representative gene) and fitness. Because our
191 *D. vitifoliae* RNA-seq samples were collected from a single host genotype we ran linear
192 regression tests for averaged insect fitness and averaged eigengenes for each host
193 genotype and co-expression module. Five eigengenes significantly correlated with *D.*
194 *vitifoliae* fitness (Figure 3C-F); since the 'grey' module is a bin for genes that cannot be
195 assigned to a co-expression module and are unlikely to be co-regulated, we included
196 only the 'thistle', 'darkolivegreen4', 'darkorange2', and 'antiquewhite4' modules for
197 further analysis. The darkolivegreen4 module was positively correlated with *D. vitifoliae*

198 fitness (R^2 = 0.91, p-value = 0.012) (Figure 3D). Darkolivegreen4 is moderately
199 overrepresented for genes related to detoxification of plant metabolites and other
200 environmental molecules (adjusted p-value = 0.084) (Supplemental Table 3). The
201 antiquewhite4 module was also positively correlated with *D. vitifoliae* fitness (R^2 = 0.79,
202 p-value = 0.044) (Figure 3F). A Fisher's exact test for gene annotation group enrichment
203 and KEGG term enrichment analysis showed no enrichment for the antiquewhite4
204 module. However, antiquewhite4 showed an interesting pattern of gene expression
205 across treatments (Supplemental Figure 5). Genes within the module clustered into two
206 groups with distinct expression patterns based on hierarchical clustering of vst-
207 normalized expression. A small cluster of nine genes increased in expression during
208 interaction with *V. vinifera* cv Cab33 relative to native hosts 101-14, 1103, and 99R,
209 whereas a larger cluster of 57 genes decreased in expression during interaction with *V.*
210 *vinifera* cv Cab33 relative to native hosts. Thus, the antiquewhite4 module appears to
211 be downregulated in *D. vitifoliae* during interaction with *V. vinifera* relative to native
212 hosts, which may contribute to the successful colonization of this non-native host.
213
214 <u>The darkorgange2 module is enriched for genes that are host-responsive and encode</u>
215 <u>glycoside hydrolase family 1 proteins</u>
216 We performed Fisher's exact tests to identify gene co-expression modules that
217 were enriched for host-responsive genes. 21 out of the 88 modules were enriched for
218 host-responsive genes (odd ratio >2, BH-adjusted p-value <0.05) (Supplemental Table
219 1). The darkorange2 module was the only module that was both enriched for host-

220 responsive genes and significantly co-varying with *D. vitifoliae* fitness. Darkorange2 is
221 also the only module that was negatively correlated with *D. vitifoliae* fitness (Figure 3C-
222 F). Based on KEGG term enrichment analysis, darkorange2 is over-represented for
223 genes that encode glycoside hydrolase family 1 enzymes (PFAM GH1), which mediate
224 carbohydrate metabolism (Supplemental Figure 6). Similar to the antiquewhite4 module,
225 genes within the darkorange2 module show distinct patterns of expression across host
226 genotypes (Figure 6). A small cluster of 24 genes is highly expressed during interaction
227 with native hosts 101-14 and 1103 and decreases in expression during interaction with
228 *V. vinifera* cv Cab33 and SO4, which are hosts that maintain relatively low *D. vitifoliae*
229 fitness. 4 of the 5 genes with glycoside hydrolase KEGG annotation are within this 24-
230 gene cluster and show similar patterns of expression (Figure 6). The remaining 66
231 genes in the module are grouped into a larger cluster that shows the exact opposite
232 trend in expression, largely increasing in expression during interaction with *V. vinifera* cv
233 Cab33 and SO4 relative to 101-14 and 1103. These results show the darkorange2
234 module is highly plastic and variation in module expression may explain differences in
235 *D. vitifoliae* fitness across distinct host genotypes.
236
237 <u>The thistle module is enriched with predicted effectors including a novel group of</u>
238 <u>proline-rich secretory proteins</u>
239 The thistle module was most strongly correlated with *D. vitifoliae* fitness (R^2 =
240 0.94, p-value = 0.005) and) and was positively correlated with fitness (Figure 3C). The
241 thistle module is also enriched with genes encoding predicted secretory effectors based

242 on Fisher's exact test (Supplemental Table 3), including a group of four effectors with
243 homology to the *Acyrthosiphon pisum* lysine-rich arabinogalactan protein 19-like protein
244 based on BLAST searches
245 (https://bipaa.genouest.org/sp/daktulosphaira_vitifoliae/download/annotation/v3.2/func_
246 annot/)[15]; we termed these genes 'ARGL' effectors, for **AR**abino**G**alactan-**L**ike. We also
247 noticed that the ARGL genes are in a tandem cluster in the genome on scaffold 40
248 (Supplemental Figure 7). Within this same ~50-kb spatial cluster are two more predicted
249 effectors and three 'null' genes; two of the three null genes are also assigned to the
250 thistle module, suggesting they are co-regulated with the ARGL genes. We performed
251 new BLASTp searches for ARGL effectors and found that they show homology to one
252 another and to other proteins from *D. vitifoliae*, *Adelges cooleyi*, and multiple aphid
253 species (Supplemental Table 4). The ARGL effectors and all proteins that are significant
254 BLAST hits contain no predicted functional domains based on InterProScan searches.
255 To determine if the ARGL effectors have highly similar sequences and shared
256 protein motifs, we performed Multiple Sequence Alignment (MSA) of the seven ARGL
257 proteins using clustalOmega[16]. The sequence logo plot for the ARGL protein MSA
258 revealed these proteins have a high proportion of proline residues and are highly
259 conserved, especially at proline residues (Figure 5A). To understand the evolutionary
260 relationship among ARGL effectors, we performed MSA for all 2,743 predicted effector
261 proteins and constructed a phylogenetic tree. Despite being highly conserved at the
262 amino acid level, AGRL effectors did not cluster together in a monophyletic group
263 (Supplemental Figure 8), suggesting ARGL proteins have some structural divergence.

264 The ARGL effectors show similar patterns of transcript expression across treatments,
265 generally increasing in expression during interaction with high-fitness native hosts 101-
266 14, 1103, and 99R relative to *V. vinifera* cv Cab33. Overall, our analysis of the thistle
267 co-expression module suggests ARGL genes may have shared function and play a role
268 in *D. vitifoliae* host use, albeit through unknown mechanisms.
269

270 **Discussion**

271 We have presented analyses of a large RNA-seq data set collected from *D.*
272 *vitifoliae* adults after feeding on different host grapevines (*Vitis* spp.) for a single
273 generation, with the aim of understanding the transcriptional basis of host use in this
274 specialist herbivore. We found that 37% of all expressed genes are host-responsive,
275 with most genes being upregulated during interaction with the non-native host *V.*
276 *vinifera* cv Cab33 relative to native hosts. Genes encoding predicted secretory effectors,
277 cuticle and chitin associated proteins, and detoxification-related proteins were enriched
278 among host-responsive genes, suggesting that genes that function in host-manipulation,
279 detoxification of xenobiotics, and protection against the environment are dynamically
280 regulated for successful colonization of the non-native host *V. vinifera*. Weighted gene
281 co-expression network analysis revealed modules co-regulated genes that covary with
282 *D. vitifoliae* fitness. Most notably, the co-expression module 'thistle' was strongly
283 positively correlated with *D. vitifoliae* fitness and enriched with genes encoding
284 predicted secretory effectors. Five of these genes were co-located in a ~25-kb spatial
285 cluster in the genome and encode highly conserved and proline-rich secretory proteins,

286 termed ARGL proteins. These proteins have no known functional domain, thus ARGL
287 gene function remains unclear. Our findings show that the *D. vitifoliae* transcriptome is
288 highly plastic and is reprogrammed within a single generation on different host
289 grapevines. Our analysis of genome-wide gene expression and fitness for *D. vitifoliae*
290 on different hosts also indicates transcriptional plasticity for a subset of genes may be
291 adaptive and enable successful colonization of the non-native wine grape *V. vinifera*.
292

293 **Materials and Methods**

294 Propagation of host grapevines

295 Nine different grapevine genotypes were used as experimental hosts for *D.*
296 *vitifoliae* in our experiments, all of which were clonally propagated from cuttings
297 purchased as rooted, overwintered 1-year-old plants from Inland Desert Nursery
298 (Benton City, WA). Genotypes included self-rooted rootstock hybrids of *V. berlandieri* x
299 *V. rupestris*), (99R, 1103P), *V. berlandieri* x *V. riparia* (SO4, 5BB), and *V. riparia* x *V.*
300 *rupestris* (101-14, SW, 3309). All plants were potted in 3-gallon round pots with 'UC Mix
301 3' soil (https://agops.ucr.edu/soil-mixing#uc-soil-mix-iii) and grown in a greenhouse
302 (24:27C +/- 3C for day:night with 14h daylight). After leaf bud swelling (8 days after
303 potting), half of each genotype were infested with the remaining plants potted and
304 infested similarly 6 days later. Plantings were replicated through time to optimize
305 sample collection effort given the time required to sort and count insects.
306

307 Phylloxera rearing and experimental infestation of host grapevines

308 A clonal line of leaf-galling *D. vitifoliae* was established from a single female
309 collected on *V. rupestris* (cv St. George) near Davis, CA, in 2021. The clonal population
310 was maintained for successive generations on both grafted (*Vitis* genotype 1616 and
311 3309 rootstocks) and self-rooted *V. vinifera* cv Cabernet Sauvignon 33 (Cab33) in UCR
312 greenhouses in Riverside, CA (27+/-3C) until experimentation over 2022-2023. For
313 infestation, galled leaves from the starting *D. vitifoliae* colony (hereafter generation 0, or
314 'G0') were excised and placed on the branch nodes of new host plants. Infestation was
315 monitored every 12 hours to note hatch and dispersal of 1st nymphal instars (hereafter
316 generation 1, 'G1') onto budding leaves. Galled leaves used for infestation were
317 removed from branch nodes after two days to prevent further infestation.
318
319 <u>Insect performance assay and statistical analysis</u>
320 To assess plant vigor and insect performance, 'G1' galls were collected 15 days
321 after infestation (DAI). Galls and tissue from ungalled leaves of similar age and node on
322 un-infested plants were collected using a 4mm diameter hole punch, placed in tubes on
323 ice for transport back to the lab. Galls were dissected under a dissecting microscope to
324 count eggs per gall. Dissected gall tissues and 10 leaf discs per plant were also
325 weighed for fresh weight, dried at 50C until dry (1-2 days), and then assessed for dry
326 weight. Each plant was sampled five different times, at 15, 17, 19, 21, and 23 DAI. Each
327 time the five largest galls from each plant were sampled, selecting galls from leaves that
328 had not been previously sampled unless the largest remaining galls were on a
329 previously sampled leaf. Egg production was assessed using a mixed effects linear

330 model with genotype and day as predictors and plant id as a random effect. Post hoc
331 comparisons among genotypes were performed using pairwise Wilcoxon tests (R v.
332 4.3.1).
333
334 RNA isolation and RNA sequencing
335 To quantify *D. vitifoliae* transcriptome plasticity during host switching events,
336 twenty adults were collected from galls from unique plants and pooled per each sample.
337 Insects were placed in tubes on ice during collection then immediately stored at -80 °C
338 until RNA isolation. RNA was isolated from adults using the RNeasy Mini Kit (Qiagen)
339 following manufacturer's instructions. We ran gel electrophoresis to confirm RNA
340 intactness and only kept samples that had intact 28S and 18S ribosomal RNA bands.
341 We used a Nanodrop spectrophotometer to measure RNA purity and only used samples
342 with 260 nm / 280 nm absorbance ratios above 2 and 260 nm / 230 nm absorbance
343 ratios above 1.7 for library preparation. Libraries were prepared using a
344 NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England
345 Biolabs). Paired-end RNA libraries were sequenced using an Illumina NovaSeq X
346 system at the DNA Technologies and Expression Analysis Core of the UC Davis
347 Genome Center, Davis, CA.
348
349 RNA-seq data preprocessing
350 Fastqc v. 0.11.9 (https://github.com/s-andrews/FastQC) was used to generate
351 read and base quality statistics for raw and processed reads. Raw RNA-seq reads were

352 quality trimmed and filtered using BBTools bbduk (v.38.95; https://jgi.doe.gov/data-and-
353 tools/software-tools/bbtools/) with default settings plus options for kmer length (k=23)
354 and hamming distance (hist=1). Trimmed and filtered reads were aligned to the
355 reference genome [15] using STAR v. 2.7.10b [17] with default settings plus options for
356 maximum intron length (--alignIntronMax 15000) and removing alignments that
357 contained non-canonical junctions (--outFilterIntronMotifs RemoveNoncanonical). Only
358 uniquely mapped reads were used for downstream analysis.
359
360 Differential gene expression analysis
361 Gene count matrices were generated using the --quantMode GeneCounts option
362 in STAR, which contained gene counts that coincide with counts produced by htseq-
363 count with default parameters. Gene count matrices were processed and analyzed
364 using edgeR (v.4.0.16) [13]. Genes with low counts were removed from downstream
365 analysis using the filterByExpr command with default parameters. Linear model fitting of
366 count data was performed using limma's voom and lmFit functions [18]. Genes were
367 considered differentially expressed at adjusted p-value < 0.05 and log fold-change $\geq$
368 1.5. Pie charts were made using the R graphics package (v.4.3.1) to visualize functional
369 annotations of genes that were differentially expressed. Fisher's exact testing for gene
370 functional group enrichment was performed with the R stats package (v.4.3.1).
371 Functional groups were considered significantly enriched based on a BH-adjusted p-
372 value cutoff of 0.05 and an odds ratio ≥ 2.
373

374 Gene co-expression network analysis

375 Genes with 0.5 CPM (about 10 raw gene counts) in at least four samples were
376 included for network construction using WGCNA [14]. Gene count matrices were
377 normalized using DESeq2 vst (variance stabilizing transformation) function (v. 1.42.1). A
378 signed gene co-expression network was constructed based on hierarchical clustering of
379 dissimilarity in gene connectivity (1-TOM) with the soft power parameter (β) set to 16 to
380 ensure the network had approximate scale-free topology. Genes were assigned to
381 modules using cutreeDynamic function with minimum module size = 20 and module
382 detection sensitivity (deepSplit) = 2. Heat maps of module eigengenes and normalized
383 gene expression (z-scores of vst-transformed count data) were made using pheatmap
384 (v. 1.0.12) (https://github.com/raivokolde/pheatmap); rows were clustered based on
385 Euclidean distance. Correlation between *D. vitifoliae* fitness and gene co-expression
386 modules were tested by running a linear regression test for average fitness on each
387 host genotype and average module eigengene expression using the lm function in the R
388 stats package v. 4.3.1: fitness = $\beta 0 + \beta 1 \times$ eigengene + ϵ. Scatter plots of eigengene-
389 fitness correlations were printed with results from the linear regression test using
390 ggplot2 v, 3.5.0 (https://ggplot2.tidyverse.org/). Gene co-expression modules were
391 tested for overrepresentation/enrichment of gene functional groups based on Fisher's
392 exact tests as described above.

393

394 KEGG term enrichment analysis

395 eggNOG-mapper v. 2 [20] was used to assign KEGG terms to the predicted
396 proteome of *D. vitifoliae* [15] based on precomputed orthology assignments. eggNOG-
397 mapper assigned KEGG terms to 7,532 proteins out of 15,076 queried. To test for
398 KEGG term enrichment for subsets of genes, e.g. differentially expressed genes vs all
399 KEGG-annotated genes, we used the enrich function of the R package clusterProfiler v.
400 4.10.1 [21]. KEGG terms were considered significantly enriched based a BH-adjusted p-
401 value cutoff <0.05 and a gene count ≥4.
402
403 Multiple Sequence Alignment and phylogenetic analysis
404 To determine sequence similarity and phylogenetic relationships of ARGL
405 proteins, we performed Multiple Sequence Alignment (MSA) using the clustalOmega
406 web application (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) [16]. Alignment results
407 were visualized in a sequence logo plot using the WebLogo web application
408 (https://weblogo.threeplusone.com) [22]. Phylogenetic trees were constructed from the
409 clustalOmega MSA output using the Interactive Tree of Life (iTOL) v. 6 web application
410 (https://itol.embl.de) [23].
411
412 **Acknowledgments**
413 We thank Conner Lay for help with collection of phenotypic data, Hung D. Lai for
414 greenhouse maintenance, and Holly Clark, Clay Clark, and Wei Zhang at the UC
415 Riverside genomics core for RNA library preparation. USDA-NIFA Hatch grant CA-R-

416 BPS-5199-H and CDFA grant CDFA-IAB #21-0426-000-SA were awarded to PDN to
417 support this project.
418

419 **References**

420 1. Heidel-Fischer, H. M. & Vogel, H. Molecular mechanisms of insect adaptation to plant
421 secondary compounds. *Current Opinion in Insect Science* **8**, 8–14 (2015).
422 2. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329
423 (2006).
424 3. Stuart, J. Insect effectors and gene-for-gene interactions with host plants. *Current*
425 *Opinion in Insect Science* **9**, 56–61 (2015).
426 4. Kaloshian, I. & Walling, L. L. Hemipteran and dipteran pests: Effectors and plant host
427 immune regulators. *J. Integr. Plant Biol.* **58**, 350–361 (2016).
428 5. Wang, H., Shi, S. & Hua, W. Advances of herbivore-secreted elicitors and effectors in
429 plant-insect interactions. *Front. Plant Sci.* **14**, 1176048 (2023).
430 6. Simon, J.-C. *et al.* Genomics of adaptation to host-plants in herbivorous insects.
431 *Briefings in Functional Genomics* **14**, 413–423 (2015).
432 7. García-Robledo, C. & Horvitz, C. C. Jack of all trades masters novel host plants:
433 positive genetic correlations in specialist and generalist insect herbivores expanding
434 their diets to novel hosts: Quantitative genetics of diet expansion. *Journal of*
435 *Evolutionary Biology* **25**, 38–53 (2012).
436 8. Birnbaum, S. S. L. & Abbot, P. Gene Expression and Diet Breadth in Plant-Feeding
437 Insects: Summarizing Trends. *Trends in Ecology & Evolution* **35**, 259–277 (2020).

438 9. Grbić, M. *et al.* The genome of Tetranychus urticae reveals herbivorous pest
439 adaptations. *Nature* **479**, 487–492 (2011).
440 10. Pearce, S. L. *et al.* Genomic innovations, transcriptional plasticity and gene loss
441 underlying the evolution and divergence of two highly polyphagous and invasive
442 Helicoverpa pest species. *BMC Biol* **15**, 63 (2017).
443 11. Snoeck, S., Wybouw, N., Van Leeuwen, T. & Dermauw, W. Transcriptomic
444 Plasticity in the Arthropod Generalist *Tetranychus urticae* Upon Long-Term
445 Acclimation to Different Host Plants. *G3 Genes|Genomes|Genetics* **8**, 3865–3879
446 (2018).
447 12. Mathers, T. C. *et al.* Rapid transcriptional plasticity of duplicated gene clusters
448 enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol*
449 **18**, 27 (2017).
450 13. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor
451 package for differential expression analysis of digital gene expression data.
452 *Bioinformatics* **26**, 139–140 (2010).
453 14. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation
454 network analysis. *BMC Bioinformatics* **9**, 559 (2008).
455 15. Rispe, C. *et al.* The genome sequence of the grape phylloxera provides insights
456 into the evolution, adaptation, and invasion routes of an iconic pest. *BMC Biol* **18**, 90
457 (2020).
458 16. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of
459 many protein sequences. *Protein Science* **27**, 135–145 (2018).

460 17. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–
461 21 (2013).
462 18. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock
463 linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
464 19. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
465 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
466 20. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas,
467 J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain
468 Prediction at the Metagenomic Scale. *Molecular Biology and Evolution* **38**, 5825–
469 5829 (2021).
470 21. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for
471 Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of*
472 *Integrative Biology* **16**, 284–287 (2012).
473 22. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: A
474 Sequence Logo Generator: Figure 1. *Genome Res.* **14**, 1188–1190 (2004).
475 23. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the
476 phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**, W78–W82
477 (2024).
478

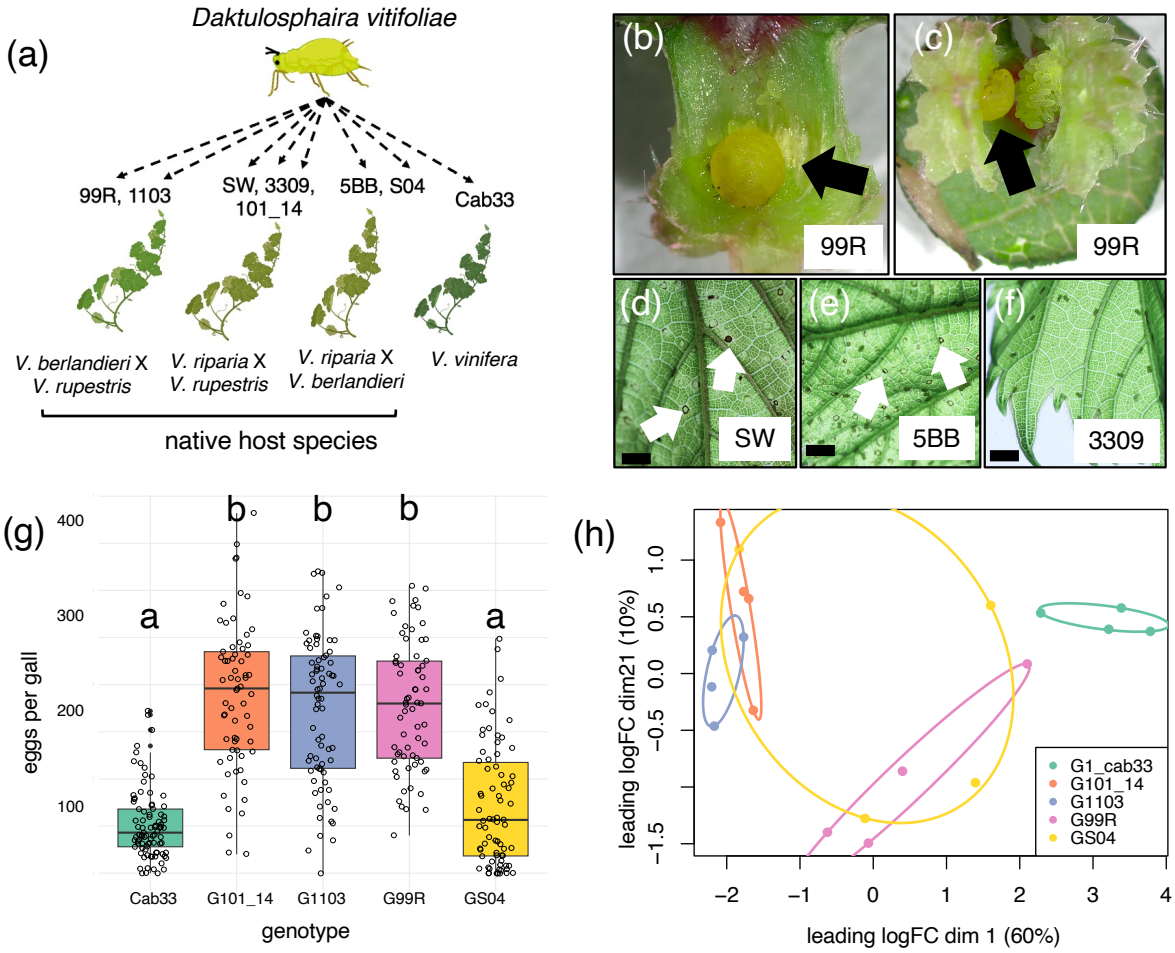


Figure 1: (a) Overview of the initial experimental setup. (b-c) A phylloxera (*D. vitifoliae*) adult (black arrows) within its gall at 3 weeks after infestation of host genotype 99R. (d-f) Young leaves of resistant host plants at 6 days after infestation; white arrows indicate sites of leaf necrosis induced by *D. vitifoliae* feeding/probing; horizontal black scale bars are 1 mm in length. (g) Box plot of *D. vitifoliae* fitness (eggs per gall) across susceptible host genotypes; genotypes with letters ‘a’ are significantly different from those with letter ‘b’ based on pair-wise Wilcoxon tests (Benjamini-Hochberg [BH]-adjusted p-values are all less than 0.0001). (h) Multi-

488 dimensional scaling plot of log-CPM (gene counts per million gene counts) for all

489 RNA-seq samples.

490

491

492

493

494

495

496

497

498

499

500

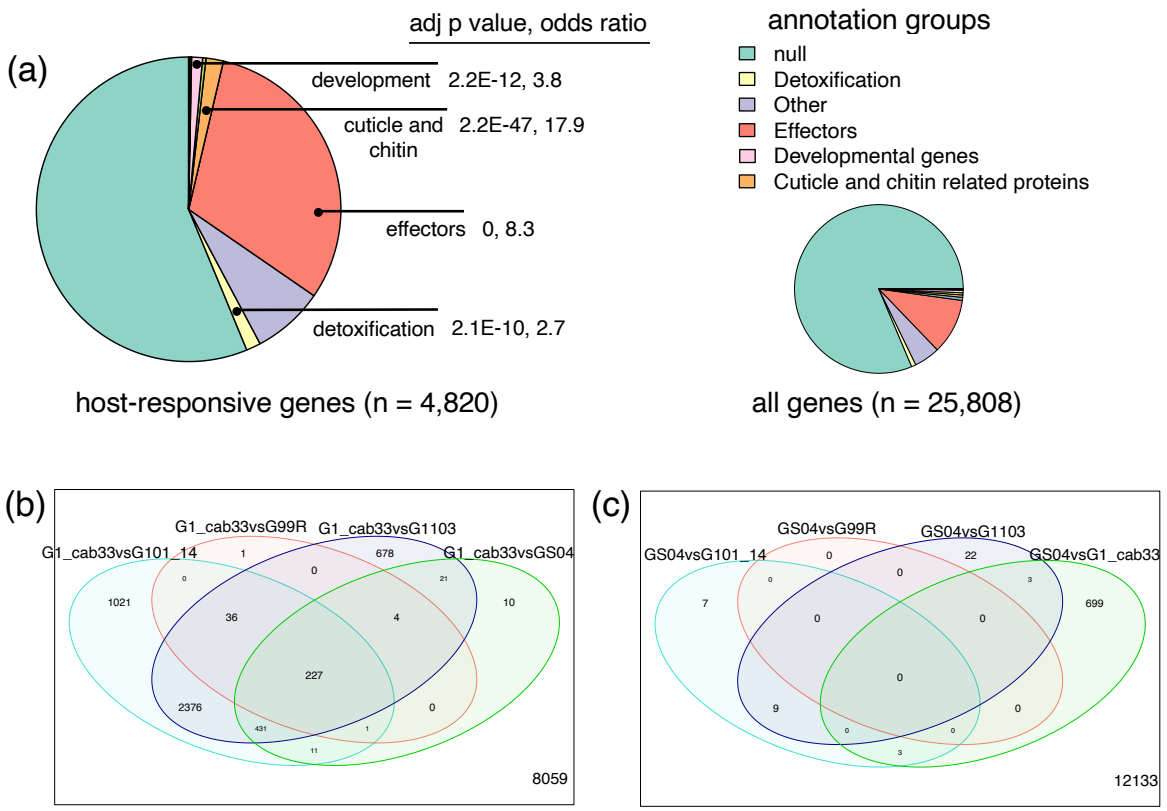


Figure 2: (a) Pie chart of all host-repsonsive *D. vitifoliae* genes and assigned functional groups; Fisher's exact test results are displayed for functional groups that are signficantly enriched among host-responsive genes relative to all genes. (b-c) Venn diagrams of differentially expressed *D. vitifoliae* genes contrasting *V. vinifera* cv Cab33 (b) or SO4 (c) with all other hosts genotypes.

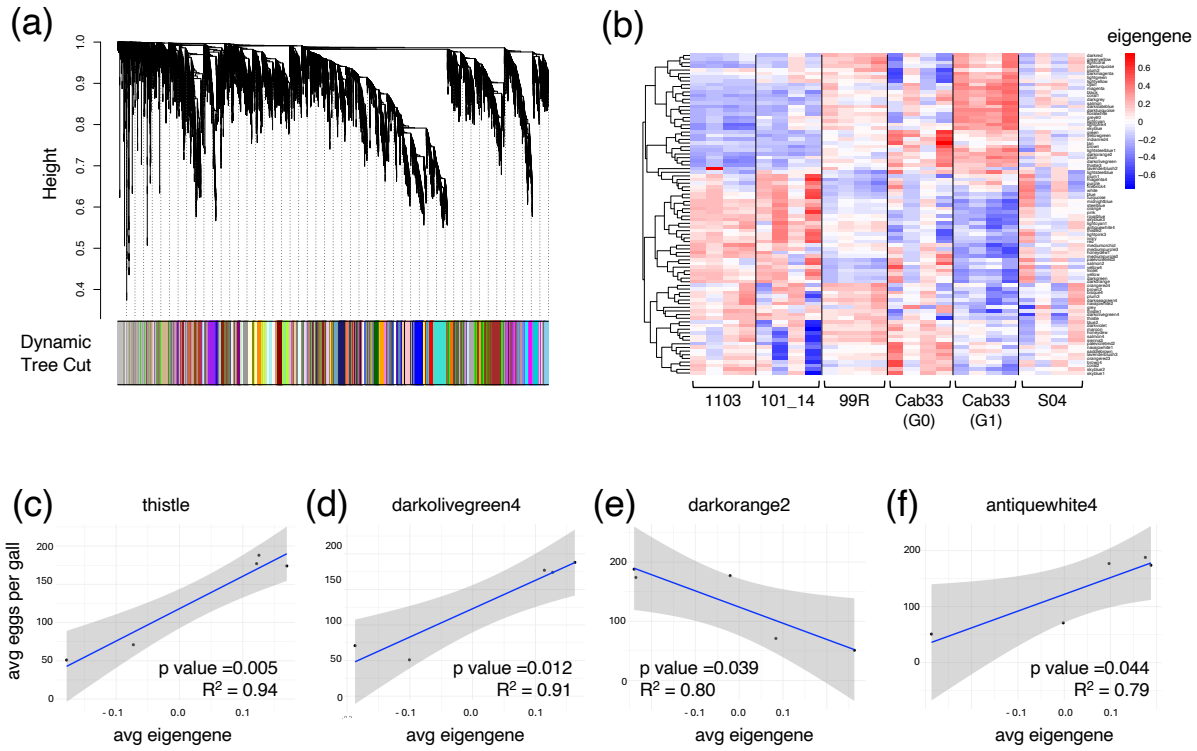


Figure 3: (a) Gene dendrogram for *D. vitifoliae* signed gene co-expression network and module assignment. (b) Heatmap of module eigengenes for all 88 modules clustered based on Euclidean distance. (c-f) Scatter plot and results of linear regression tests for four co-expression modules that correlate in expression with *D. vitifoliae* fitness.

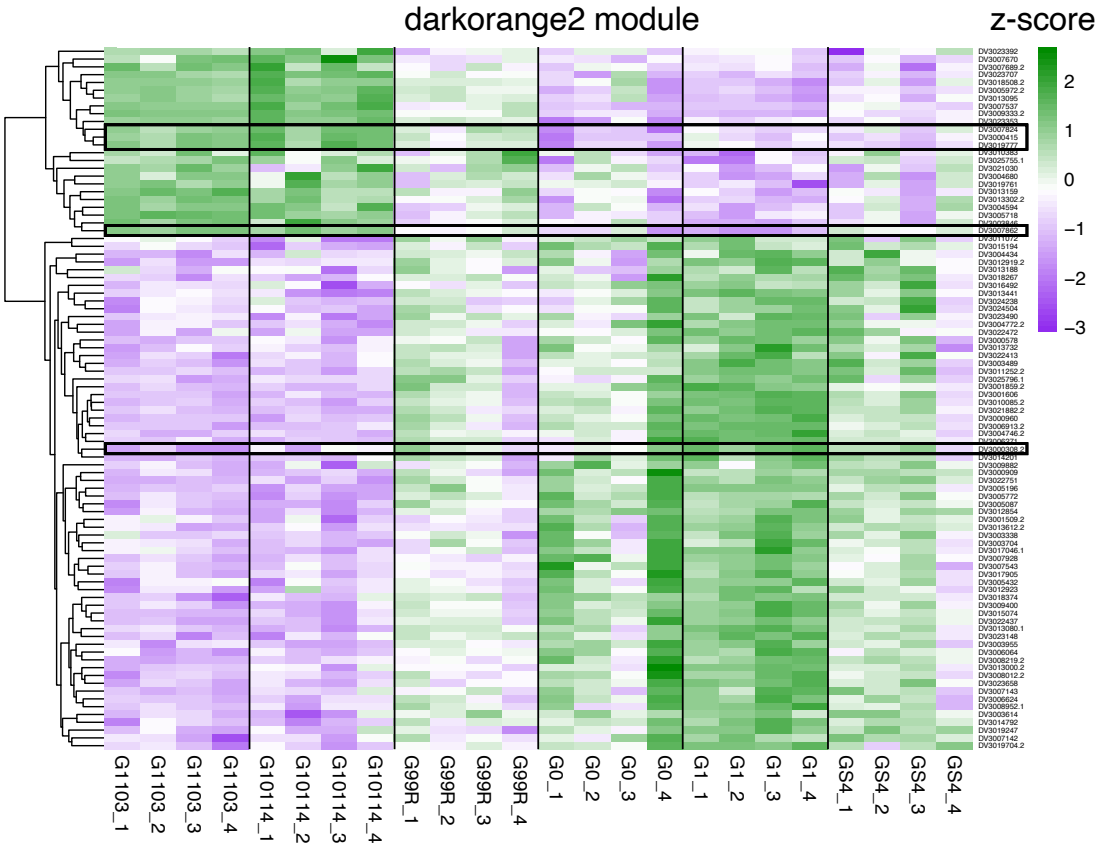
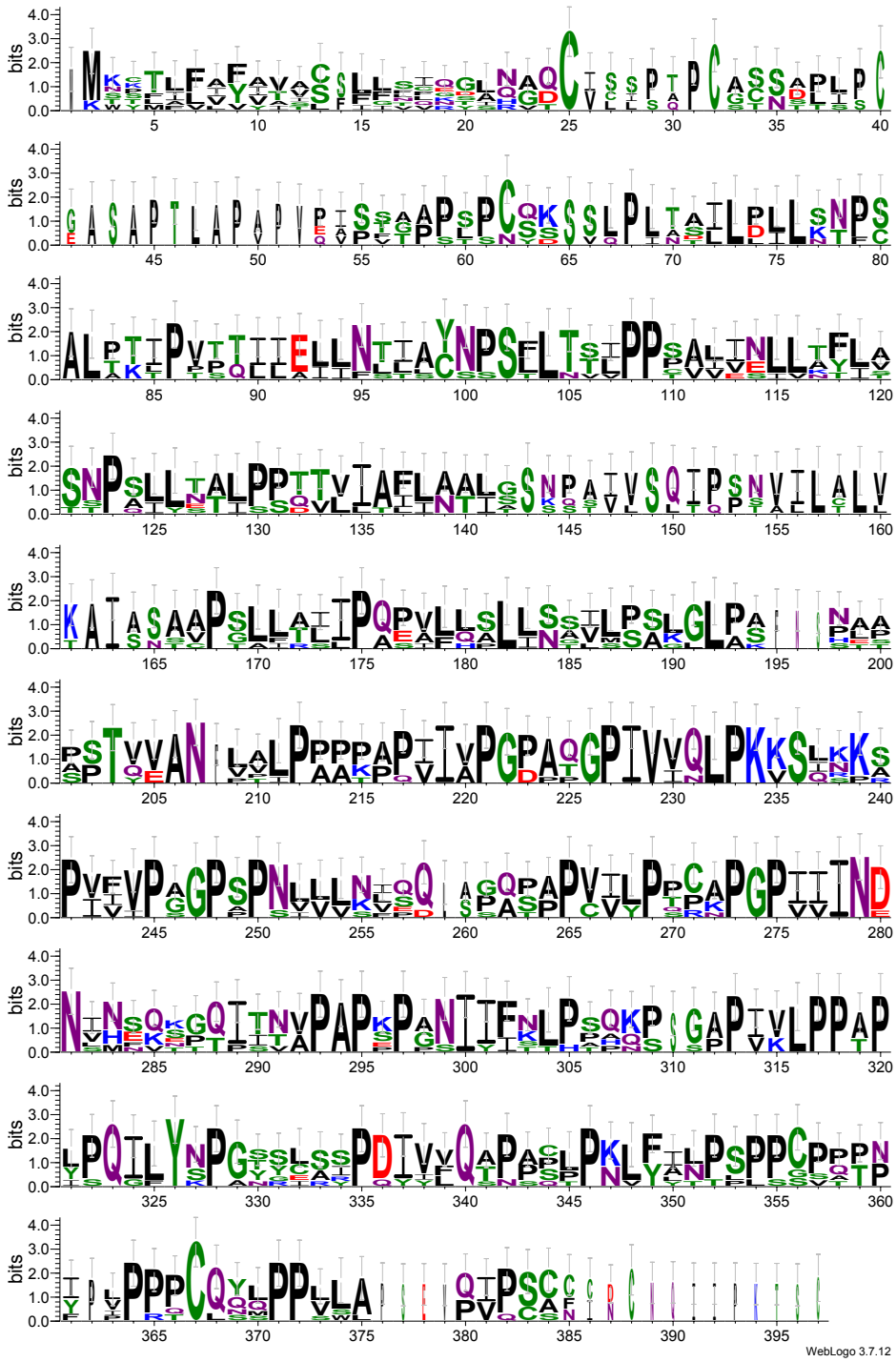


Figure 4: Heatmap of normalized gene expression for genes in the darkorange2 module. Black rectangles highlight genes with glycoside hydrolase KEGG annotation.



516

517 **Figure 5**: Sequence logo plot for the seven predicted effector proteins at the ARGL

518 locus.

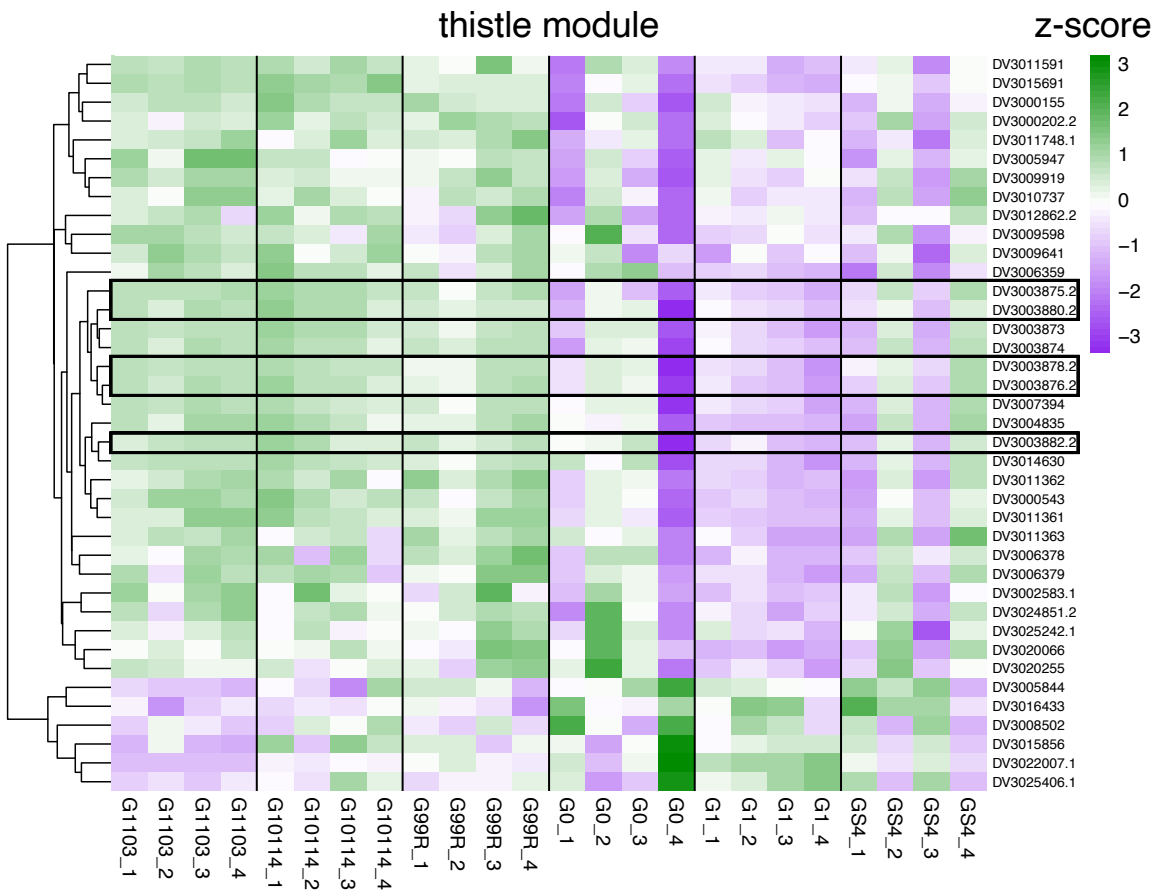


Figure 6: Heatmap of normalized gene expression for genes in the thistle module. Black rectangles highlight ARGL effectors.