Spread of a single superclone drives insecticide resistance in *Acyrthosiphon kondoi* across an invasive range

Joshua A. Thia[1,*], Benjamin J. Hunt, Shuchao Wang[2], Bartlomiej J Troczka[2], Evatt Chirgwin[3], Courtney J. Brown[1], Rumi Sakamoto[1], Monica Stelmach[1], Kelly Richardson[1], Leonhard S. Arinanto[1], Ashritha P. S. Dorai[1], Chinmayee Joglekar[1], Qiong Yang[1], Marielle Babineau[3], Chris Bass[2], Paul A. Umina[1,3], Ary A. Hoffmann[1]

[1] School of BioSciences, The University of Melbourne, Melbourne, Victoria, Australia

[2] Centre for Ecology and Conservation, University of Exeter, Penryn, Cornwall, UK

[3] Cesar Australia, Brunswick, Victoria, Australia

Name	**ORCID**
Joshua A Thia	https://orcid.org/0000-0001-9084-0959
Benjamin J Hunt	https://orcid.org/0000-0003-3825-9723
Shuchao Wang	https://orcid.org/0009-0003-1025-9155

Bartlomiej J Troczka	https://orcid.org/0000-0001-9177-2441
Evatt Chirgwin	https://orcid.org/0000-0002-8242-3837
Courtney J Brown	https://orcid.org/0000-0003-1515-1947
Monica Stelmach	https://orcid.org/0009-0006-7671-1775
Kelly Richardson	https://orcid.org/0000-0001-7186-1260
Chinmayee Joglekar	https://orcid.org/0009-0001-5345-9128
Ashritha PS Dorai	https://orcid.org/0009-0004-2928-039X
Qiong Yang	https://orcid.org/0000-0003-3336-5703
Marielle Babineau	https://orcid.org/0009-0008-9524-7765
Chris Bass	https://orcid.org/0000-0002-2590-1492
Paul A Umina	https://orcid.org/0000-0002-1835-3571
Ary A Hoffmann	https://orcid.org/0000-0001-9497-7645

Abstract

Populations under similar selection pressures may adapt via parallel evolution or dispersal of advantageous alleles. Here, we investigated insecticide resistance in the invasive blue-green aphid, *Acyrthosiphon kondoi*, which reproduces clonally in Australia and has rapidly developed resistance across geographic locations. Using genomic, transcriptomic, and experimental approaches, we explored the evolutionary origins and molecular mechanisms of resistance. We developed the first nuclear genome assembly for *A. kondoi* (443.8 Mb, 28,405 annotated genes, BUSCO score 97.5%) and a partial mitochondrial assembly (11,598 bp). All resistant strains shared a common ancestor, supporting the spread of a resistant 'superclone' lineage that is distinct from susceptible strains. Resistance was associated with over-expression of an esterase gene that was homologous to E4/FE4 esterases in other aphid pests that are linked to resistance. Functional experiments in *Drosophila melanogaster* confirmed a causal role of this E4-like esterase in resistance to organophosphates, carbamates, and pyrethroids. These findings highlight how clonal dispersal and insecticide overuse can transform local adaptation into a widespread pest management issue. Our results suggest a parallel macroevolutionary response to insecticide selection in *A. kondoi* and other aphid species at the gene family level, but with a distinct regulatory mechanism in *A. kondoi*. Given the rapid spread of the resistant superclone, alternative management strategies, including

34 expanded chemical control options and enhanced biological control, are urgently

35 needed to mitigate this growing pest problem.

36

37

38 # Introduction

39 Understanding variation in the rise and spread of adaptive traits and their underlying
40 genes is a key issue in evolutionary biology (Cohan and Hoffmann 1989; Korona et al.
41 1994; Colosimo et al. 2005). Populations or species exposed to similar selective pressures
42 may evolve through parallel changes (independent evolutionary events). Parallelism
43 might occur only at the phenotypic level (Roda et al. 2013; Magalhaes et al. 2020; Yang
44 et al. 2020; Sprengelmeyer and Pool 2021), or it may extend to the genotypic level,
45 where the same allelic changes are involved. The latter is more likely if populations share
46 standing genetic variation or if there are genetic constraints on evolving adaptive
47 phenotypes (Colosimo et al. 2005; Yeaman et al. 2018; Magalhaes et al. 2020; Montejo-
48 Kovacevich et al. 2022). Gene flow can also spread adaptive alleles across areas with
49 similar selective pressures (Endersby-Harshman et al. 2020; Tepolt et al. 2021; Zhang et
50 al. 2023). Indeed, gene flow can assist adaptation to new environmental stressors by
51 introducing adaptive variation into local populations (Chevillon et al. 1999; Bontrager
52 and Angert 2019). Disentangling these processes sheds light on the repeatability of
53 evolution and the origins of adaptive variation in local populations.

54 Clonal organisms are often considered to have lower adaptive capacity because they lack
55 recombination to produce adaptive allelic combinations (Muller 1932; Barraclough et al.
56 2003). However, when clonal diversity arises from mutation accumulation or rare sexual
57 reproduction, it is possible for selection to drive clonal lineages with adaptive alleles to

58 high frequency, often at the expense of other lineages (Muller 1932; Gerrish and Lenski
59 1998; Barraclough et al. 2003). Clonal organisms therefore provide a unique context for
60 tracking the origin and trajectory of adaptive traits and alleles, as these are directly
61 linked to specific (clonal) genetic backgrounds. For example, in the cystic fibrosis-
62 associated bacterium *Pseudomonas aeruginosa*, recurrent selective sweeps have been
63 associated with parallel genetic, transcriptional, and phenotypic changes for adaptive
64 pathogenicity among clonal lineages (Huse et al. 2010; Feliziani et al. 2014; Caballero et
65 al. 2015). In the parasitic protozoan, *Toxoplasma gondii*, dominance of three clonal
66 lineages in North America and Europe has been attributed to strong selection on key
67 virulence alleles associated with each of these lineages (Howe and Sibley 1995; Khan et
68 al. 2009). Studies on *Daphnia* species have also shown clear, predictable patterns of
69 changes in clone frequencies due to variation in competitive abilities of clonal lineages
70 under different environmental conditions (Korpelainen 1986; Steiner et al. 2016; Lyberger
71 and Schoener 2023). And in yeast, tracking of clonal lineages has illustrated how the
72 accumulation of beneficial mutations within clones leads to the dominance of highly fit
73 genotypes (Nguyen Ba et al. 2019).

74 Insecticide usage in agriculture provides an interesting context for studying adaptation
75 to similar selective environments. Pest species are often exposed to the same insecticide
76 across different populations as part of chemical control strategies (Walsh et al. 2018;
77 Endersby-Harshman et al. 2020; Singh et al. 2021; Thia, Korhonen, et al. 2023). Because

78 insecticides exert their toxicity through very specific protein targets and molecular
79 interactions, there is strong precedence for joint phenotypic and genotypic parallelism in
80 insecticide resistance traits. This has been demonstrated repeatedly, leading to parallel
81 and convergent evolution of some mutations in these target proteins that prevent
82 insecticide inhibition – so called 'target-site mutations' leading to amino acid
83 substitutions (Fournier 2005; Davies et al. 2007). Such parallelism can occur both at the
84 microevolutionary scale among populations (Walsh et al. 2018; Thia, Korhonen, et al.
85 2023), and at the macroevolutionary scale, leading to convergence among species
86 (Feyereisen et al. 2015). However, alternate mechanisms might also evolve, such as shifts
87 in the expression or efficacy of detoxification proteins, and polygenic resistance can lead
88 to non-parallel genetic mechanisms of insecticide resistance among populations (Field
89 and Devonshire 1998; Crossley et al. 2017; Yang et al. 2020; Thia, Korhonen, et al. 2023).

90 Although clonal reproduction is uncommon in animal populations, it tends to occur at a
91 relatively higher frequency among invertebrates in agricultural environments (Hoffmann
92 et al. 2008). Aphids (Order: Hemiptera) are a major agricultural arthropod pest: they are
93 sap-sucking insects that damage plants through direct feeding and virus transmission
94 (Blackman and Eastop 2017). Aphid populations are often composed of holocyclic
95 (sexual or asexual, host-alternating) clones, particularly in temperate regions with cold
96 winter temperatures, but they often become anholocyclic (asexual, nonhost-alternating)
97 in more benign climates (Rispe et al. 1998; Rispe and Pierre 1998; Williams and Dixon

98 2017). Clonal reproduction can dramatically shape the evolutionary trajectory of aphid
99 populations by favouring the dominance of 'superclones'. These are clonal lineages that
100 have acquired highly advantageous traits (through genetic mutation or endosymbiont
101 acquisition) that allow them to spread across broad geographic extents (Gilabert et al.
102 2015; Godefroid et al. 2024; Mahieu et al. 2024). For agriculturally important aphid
103 species, insecticide selection can facilitate the spread of superclones carrying insecticide
104 resistance mutations. For example, in the peach-potato aphid, *Myzus persicae*, although
105 different genetic lineages carrying resistance have evolved around the world (Singh et al.
106 2021), local sweeps of specific superclones have occurred in places like Australia (Umina
107 et al. 2014; de Little et al. 2017; Ward et al. 2024). Similarly, in the wheat aphid, *Sitobion*
108 *avenae*, a single insecticide resistant clonal lineage has become dominant in the UK,
109 reducing local clonal diversity (Morales-Hojas et al. 2020).

110 In this study, we investigated the recently evolved case of insecticide resistance in the
111 blue green aphid, *Acyrthosiphon kondoi*, in its invasive Australian range. This aphid is
112 native to Asia but has become an agricultural pest across the world, including the USA,
113 South America, Africa, and Australia (Ryalls et al. 2013). It is a pest of alfalfa and
114 leguminous grains crops such as lentils and lupins (Humphries et al. 2012; Ryalls et al.
115 2013). In Australia, and in other parts of its invasive range, *A. kondoi* reproduction is
116 mostly asexual (Ryalls et al. 2013), with sexual reproduction often occurring in its native
117 range, though data are limited (Kawada 1992). Historically, there are few reports of

118 insecticide resistance in *A. kondoi*. One of the only cases is described in unpublished
119 work from the USA (Natwick et al. 2014). Since then, there have been no documented
120 cases globally until the recent discovery of organophosphate, carbamate, and pyrethroid
121 resistance in Australian *A. kondoi* populations in 2020 (Chirgwin et al. 2023). This has led
122 to control failures and subsequent economic damage to Australian alfalfa growers
123 (Chirgwin et al. 2023). Two outstanding issues are whether the resistance has a single
124 evolutionary origin in Australian *A. kondoi*, and what genes contribute to this adaptive
125 trait.

126 Here, we consider the evolved resistance in Australian *A. kondoi* to study the
127 evolutionary origins of adaptation and its genetic basis. We began with a genomics
128 approach to address two initial objectives: (1) to determine whether resistant strains are
129 of the same or different clonal background; and (2) to identify putative genes that might
130 underpin resistance. Our finding of a candidate detoxification gene, an E4-like esterase,
131 led us to address two additional objectives: (3) to validate the differential expression
132 patterns of the candidate E4-like esterase among resistant and susceptible strains; and (4)
133 to validate the potential insecticide detoxification function provided by the candidate
134 E4-like esterase using transgenic *Drosophila*.

135

136 ## Results

137 ## Aphid strains

138 Our study leveraged a suite of field collections of clonal strains between 2020 and 2024,
139 as well as a laboratory clonal strain that had been isolated in 1999 before resistance was
140 reported in the field. The clonal strains came from a variety of plant hosts and were
141 broadly distributed across the southern regions of Australia (Table 1 and Figure 1a). We
142 first developed a set of eight core clonal isofemale strains using artificial selection with
143 the organophosphate insecticide, chlorpyrifos. These clonal isofemale strains formed the
144 core of our main experiments. Susceptible isofemale strains included Kellerberrin,
145 Eudunda, Jung and Moulamein, and resistant isofemale strains included Keith, Temora,
146 Bangham and Coombe. Success of selection was validated with follow-up screens using
147 chlorpyrifos (Supplementary Table S1). Previous work has demonstrated that resistant
148 strains of *A. kondoi* exhibit general resistance to carbamates, organophosphates and
149 pyrethroids (Chirgwin et al. 2023), so we did not screen our core strains against other
150 insecticides. We later broadened our geographic sampling by sampling single individuals
151 from field strains. The phenotypes for these field strains had been assessed in a separate
152 study (Evatt Chirgwin, unpublished data), but we did not specifically impose the same
153 artificial selection treatment on them as we had with our core strains.

154 ## Genome assembly and annotation

155 To investigate the evolution of insecticide resistance in *A. kondoi*, we generated the first
156 nuclear reference genome assembly (NCBI Genome Accession JAUMJG000000000) and

157 partial mitochondrial reference genome assembly (NCBI Nucleotide Accession PV793158)
158 for this species using the Kellerberrin strain. We assembled these genomic references
159 with a combination of whole-genome Illumina short-reads and whole-genome
160 Nanopore long-reads, and annotated them using transcriptome Illumina short-reads.

161 **Nuclear genome.** Our final nuclear genome assembly was 443.8 Mb, comprising 615
162 contigs and 37 scaffolds, an L50 of 70 sequences, an N50 of 1.8 Mb, an L90 of 300
163 sequences, and an N90 of 0.3 Mb. The mean assembled sequence length was 0.7 Mb,
164 with a minimum of 0.006 Mb and a maximum of 9.7 Mb. Repetitive content comprised
165 30% of the assembly. We annotated 28,198 genes. The completeness of our assembly
166 was high based on BUSCO scores calculated for arthropod single-copy orthologues
167 (genome/protein sequences): 97.9%/97.5% of arthropod single-copy orthologous genes
168 were present, with 91.3%/91.7% present as single copies and 6.6%/5.8% present as
169 duplicate copies. There were 0.7%/0.9% fragmented genes and 1.4%/1.6% missing genes.
170 A self-to-self alignment produced a generally clean dotplot, indicating that our assembly
171 had limited redundancy (Supplementary Figure S1a).

172 Our *A. kondoi* nuclear genome assembly has similar characteristics to genomes
173 developed for the congeneric pea aphid, *A. pisum*, available on GenBank, which is the
174 most closely related species with a chromosome-level genome assembly. The *A. pisum*
175 GCF_000142985.2 scaffold-level assembly is 541.7 Mb, comprising 60,595 contigs and
176 23,924 scaffolds, with 20,940 annotated genes. The *A. pisum* GCF_005508785.2

177 chromosome-level assembly is 533.6 Mb, comprising 66,716 contigs and 21,227
178 scaffolds arranged into 4 chromosomes, with 20,307 annotated genes. We used
179 alignments of our *A. kondoi* contigs against the *A. pisum* chromosome-level assembly
180 (GCF_005508785.2) to generate pseudo-chromosomes for *A. kondoi*, which were then
181 remapped these back to the *A. pisum* chromosomes. This analysis suggested that there
182 may be considerable overlap in homologous sequences between our *A. kondoi* assembly
183 and the *A. pisum* chromosome-level assembly (Supplementary Figure S1b). There may
184 also be a reasonable amount of synteny between these congeneric aphids
185 (Supplementary Figure S2). However, additional data (e.g., Hi-C sequencing) would
186 provide greater power to test synteny.

187 **Mitochondrial genome.** Our mitogenome reference was a partial sequence comprising
188 11,598 bp. This reference contained most of the protein-coding (eight genes) and tRNA
189 genes (14 genes), and both rRNA genes (rRNA-L and rRNA-S). Missing from our
190 mitogenome were the protein-coding genes *atp*6, *atp*8, *cox*3, *nad*3, and *nad*5, and the
191 tRNA genes, tRNA-Ala, tRNA-Arg, tRNA-Asn, tRNA-Glu, tRNA-Gly, tRNA-His, tRNA-Phe,
192 tRNA-Phe, tRNA-Trp, tRNA-Tyr, and a second copy of tRNA-Ser.

193 Differential gene expression

194 We compared the transcriptomes of resistant (Keith and Temora) and susceptible
195 (Kellerberrin and Eudunda) strains to identify candidate over-expressed genes linked to
196 resistance. Three major patterns emerged from these transcriptomic analyses. First, the

197 resistant Keith and Temora strains exhibited highly similar gene expression patterns.
198 Second, gene expression patterns between the resistant strains were clearly distinct from
199 those of the two susceptible Kellerberrin and Eudunda strains. Third, the Kellerberrin and
200 Eudunda strains exhibited distinct gene expression patterns relative to each other. These
201 patterns were evident in the clear separation of resistant from susceptible strains on PC
202 axis 1 (41.89% of variation), and in the separation of the susceptible strains from each
203 other on PC axis 2 (21.94% of variation) (Figure 1b).

204 Our primary goal for the transcriptome analyses was to identify genes that were over-
205 expressed in the resistant strains relative to the susceptible strains. Because resistant and
206 susceptible strains represented distinct clonal genetic background (see below),
207 interpreting expressional differences requires care to not confound basal genetic
208 differences with insecticide adaptation. However, as a first step, we used the
209 transcriptome data to identify over-expressed detoxification genes as potential
210 candidate insecticide resistance genes for further experimental validation.

211 After controlling for false positives, we identified 198 candidate genes that were
212 consistently over-expressed in all four pairwise combinations of resistant versus
213 susceptible strains (Supplementary Figure S3; Appendix 1). GO (gene ontology)
214 enrichment analysis did not find any terms that were over-represented in these top over-
215 expressed genes. Among the top over-expressed genes, seven were associated with
216 potential detoxification functions: an E4-like esterase (gene ID g15286 with a 201-fold

217 change); an ABCB transporter (gene ID g18463 with 4-fold change); two cytochrome
218 P450s (gene IDs g15285 and g11751, with 6-fold change and 13-fold change,
219 respectively); and three UDP-glycosyltransferases (gene IDs g2017, g14292, and g6298
220 with 4-fold, 3-fold, and 5-fold change, respectively).

221 Of greatest interest was the E4-like esterase gene for two reasons. Firstly, this was clearly
222 top over-expressed gene in all pairwise comparisons between either the Temora or Keith
223 strains and the susceptible Kellerberrin and Eudunda strains. Secondly, the protein
224 sequence of this *A. kondoi* E4-like esterase showed ~80% similarity to the *M. persicae*
225 E4/FE4 esterase, which is known to contribute to organophosphate, carbamate and
226 pyrethroid resistance in that species through over-expression that is linked to increased
227 copy number (Field and Devonshire 1998; Field and Foster 2002). Similar esterases have
228 also been shown to be over-expressed in other resistant aphid species (Ono et al. 1999;
229 Müller et al. 2023). This E4-like esterase therefore became the focus of further
230 experimental investigations.

231 Clonal genetic background

232 We inferred evolutionary histories among clonal strains by calculating genetic distances
233 using genome-wide biallelic SNPs. For the Keith, Temora, Kellerberrin and Eudunda
234 strains, SNPs were derived from transcriptomes. For the Bangham, Coombe, Moulamein,
235 and Jung strains, SNPs were derived from whole-genome sequencing data. Our
236 comparison of genetic distances among all core eight strains indicated that the resistant

237 strains formed a monophyletic lineage that was distinct from all susceptible strains
238 regardless of geographic origin (Figure 1c).

239 Within the resistant and susceptible lineages, genetic differences were comparable
240 across strains for 4,161 biallelic SNP loci. Comparing across resistant strains, the mean
241 allelic differences ranged from 9 to 84.75 alleles, with a mean of 60.31 alleles. Comparing
242 across susceptible strains, the mean allelic differences ranged from 9 to 141.25 alleles,
243 with a mean of 84.82 alleles. When comparing within strains, the mean allelic differences
244 for the four strains with transcriptomic data (Eudunda, Kellerberrin, Temora and Keith)
245 ranged from 11.62 to 17.88 alleles, with a mean of 13.97 alleles. We could not provide
246 the same within-strain assessment for the four strains with whole-genome data
247 (Bangham, Coombe, Jung, Moulamein) because we only sequenced one pooled sample
248 per strain. Nevertheless, across the resistant and susceptible lineages, allelic differences
249 between strains ranged from 3,657 to 3,740 alleles, with a mean of 3,699 alleles. Hence,
250 differences *within* broader resistant vs susceptible clonal genetic backgrounds were
251 dramatically reduced compared to differences *between* them.

252 We noticed that samples genotyped with whole-genome sequencing were always
253 nearest neighbours in our genetic distance trees and tended to sit paraphyletic to
254 samples genotyped with transcriptome sequencing. It is known that allele-specific
255 expression, splicing, or editing can cause issues with heterozygote detection from
256 transcriptome data (Castel et al. 2015; De Wit et al. 2015). However, in our study, there

257 were no clear reductions in SNP-level heterozygosity of transcriptome sequenced
258 samples relative to whole-genome sequenced samples (Supplementary Table S2). If
259 technical differences were at play, then a few key loci may have caused clustering of
260 samples by the sequencing approach, but this pattern was not evident in the results.

261 We further examined the genetic divide between resistant vs susceptible *A. kondoi* by
262 genotyping mitochondrial SNP loci. In principle, mitochondrial markers should not be
263 affected by allele-specific expression because there is only a single allele within a clonal
264 strain. We mapped our transcriptome and whole-genome sequencing data to our
265 mitogenome assembly and identified candidate sites that were fixed between the
266 resistant and susceptible strains. We then developed sets of primers to amplify focal
267 regions in the *cytb*, *cox*1, and *cox*2 genes, each of which contained a diagnostic SNP for
268 resistance status. These primers were used to genotype individuals from additional
269 clonal strains (not included in our genomic datasets) via Sanger sequencing of PCR
270 amplicons. Combining all mitochondrial data, clear genotypic differences were observed
271 between resistant vs susceptible lineages, irrespective of sequencing approach,
272 emphasising a strict association between the genetic background and the resistance
273 phenotype (Figure 2, left panel).

274 E4-like esterase genomic copy number, expression, and sequence 275 variation

276 The identified *A. kondoi* E4-like esterase (a type B carboxylesterase) was an intriguing
277 candidate for further investigation given its homology to esterases that confer
278 insecticide resistance in other aphids. We used qPCR to validate the over-expression of
279 the E4-like esterase in resistant *A. kondoi* and test whether this over-expression was
280 achieved through increased copy number of this gene, as seen in *M. persicae* (Field and
281 Devonshire 1998; Field and Foster 2002). A custom set of qPCR primers was designed for
282 the candidate E4-like esterase as our focal gene to be compared against actin as the
283 reference gene – the efficiency of these primers was tested before conducting our
284 experiments (Supplementary Figure S4).

285 We conducted two experiments that compared relative amounts of RNA vs DNA within
286 and between resistant vs susceptible strains: in Experiment 1, samples were comprised of
287 RNA and DNA extracted from single aphids, and in Experiment 2, samples were
288 comprised of two pooled aphids. As expected, RNA (cDNA) samples exhibited $-\Delta C_P$
289 values that were generally positive for resistant strains and negative for susceptible
290 strains (Figure 3), supporting the hypothesis of over-expression in the resistant strains.
291 The relative fold-change estimated between resistant and susceptible strains using the
292 $2^{-\Delta\Delta C_P}$method (Livak and Schmittgen 2001) (Supplementary Figure S5 and
293 Supplementary Table S3) showed that in Experiment 1, the median fold-change ranged
294 from 127- to 184-fold (mean of 155-fold), and in Experiment 2, the median fold-change
295 ranged from 29- to 124-fold (mean of 75-fold). In contrast, the $-\Delta C_P$ values for DNA

296 samples were ~0 across strains (Figure 3), suggesting that both genes are present at
297 equivalent genomic copy numbers and are amplified at the same rate in qPCR,
298 irrespective of resistance status. This therefore does not support the hypothesis that
299 over-expression of the E4-like esterase is achieved through increased copy number, so
300 some other regulatory mechanism must therefore be at play.

301 Closer examination of sequence variation within the E4-like esterase revealed distinct
302 genotypic differences between resistance and susceptible *A. kondoi* strain, irrespective of
303 whether they were genotyped with transcriptome, whole-genome, or Sanger sequencing.
304 A total of five non-synonymous mutations were observed in the candidate E4-like
305 esterase: Ala28Val, Ala61Thr, Met150Leu, Ala527Thr, and Pro586Leu. Resistant strains
306 were homozygous for the alternate allele at codons 28, 61, 150, 527 and 586, and
307 homozygous for the reference allele at codon 452 (Figure 2, right panel). Susceptible
308 strains had identical genotypes, which comprised homozygous reference alleles for
309 codons 28, 61 and 527, and heterozygous genotypes at codons 150 and 586 (Figure 2,
310 right panel).

311 To assess the possible functional effects of these non-synonymous mutations, we
312 aligned the *A. kondoi* E4-like esterase against the *M. persicae* E4 esterase homolog and
313 the *Torpedo californica* acetylcholinesterase as a model for type B carboxylesterases
314 (Figure 4). We also produced a structural model of the *A. kondoi* E4-like esterase, which
315 was of overall good quality (Figure 5): the ERRAT score was 93.91, and the

316 Ramanchandran plot value was 86%, with low-confidence regions occurring in the
317 disordered signal peptide sequence toward the N-terminus (Supplementary Figure S6).
318 Our annotation of the multi-sequence alignment and structural model suggested that
319 none of the non-synonymous mutations occur in, or near, the active site or interfere with
320 the amino acids forming the catalytic triad involved in substrate hydrolysis (Ser221,
321 Glu348, His461, *T. californica* numbering). Five of the non-synonymous mutations are
322 located close to the surface of the protein, and mostly on the opposite side to the active
323 site cleft (Figure 5); the exception was the Ala28Val mutation, located in the signal
324 peptide. It is therefore unlikely that the observed non-synonymous mutations would
325 dramatically alter the substrate profile of the enzyme.

326 E4-like esterase transgenesis

327 Experiments using transgenic *Drosophila melanogaster* were used to functionally validate
328 the role of the *A. kondoi* E4-like esterase in resistance. We successfully established three
329 transgenic strains of *Drosophila* expressing different version of the E4-like esterase,
330 which we called, H1, H2, and H3. H1 was identical to the susceptible reference genome
331 sequence (Figure 2, right panel). H2 contained the mutations Met150Leu and Pro586 Leu,
332 which are heterozygotes in susceptible *A. kondoi* and homozygous in resistant *A. kondoi*
333 (Figure 2, right panel). H3 contained three mutations, Ala28Val, Ala61Thr, and Ala527Thr,
334 absent in the susceptible *A. kondoi*, and homozygous in the resistant *A. kondoi* (Figure 2,
335 right panel).

336 Using insecticide bioassays, we demonstrated that ectopic expression of *A. kondoi* E4-
337 like esterase in *D. melanogaster* conferred significant resistance to three insecticides,
338 relative to a control strain of the same genetic background but lacking this gene (Table 2
339 and Figure 6). The level of resistance conferred by E4-like esterase in *D. melanogaster*
340 was modest, ranging from a 1.8-fold to 4.9-fold increase depending on the insecticide
341 and haplotype tested (Supplementary Table S4). For logistical reasons, we were unable
342 to compare all haplotypes in the same experiment, which limited a direct three-way
343 comparison (see Materials and Methods). However, we could directly compare H1
344 (susceptible-like) and H3 (resistant-like), which were run in the same experiment.

345 The H1 and H3 haplotypes had diametrically opposed genotypes at codons 28, 61 and
346 527 between resistant and susceptible *A. kondoi* strains (Figure 2, right panel). We found
347 that H1 exhibited a higher LD50 relative to H3 across all insecticides, but this was only
348 significant for chlorpyrifos (where confidence intervals did not overlap) (Table 2). This
349 pattern is counter to what we might expect if mutations in the resistant-like H3
350 haplotype were important for the detoxification of insecticides in *A. kondoi*. These results
351 further support the mechanism of resistance of the *A. kondoi* E4-like esterase involving
352 over-expression and not protein-coding changes in this gene.

353 Target-site mutations

354 Insecticides typically exert their toxicity by targeting specific genes. Organophosphates
355 and carbamates target the acetylcholinesterase (*ace*), and pyrethroids target the para-

356 like voltage-gated sodium channel (*para*). We therefore examined SNPs in the exonic
357 regions of these genes to assess whether any might provide target-site resistance to
358 their associated insecticides. Our genome contained three genes with *ace/ace*-like
359 annotations (g46, g11959, and g15835), but none exhibited non-synonymous mutations.
360 Our genome contained two *para* gene annotations (g11612 and g11614), but we only
361 observed three non-synonymous mutations in one *para* gene, g11614, which included
362 Thr104(1201)Ala, His359(1450)Asn, and His359(1450)Gln (note that codon numbers in
363 parentheses are the position from canonical *para* reference sequence from *Musca*
364 *domestica*) (Supplementary Figure S7). Comparison of our two *A. kondoi para* genes with
365 the pest aphids *A. pisum* and *M. persicae* indicated that gene g11614 is the *para* subunit
366 2, whereas gene g11612 is the *para* subunit 1 (Supplementary Figure S7). The structure
367 of these two genes in the genome of *A. kondoi* is very similar to that in *A. pisum* and to
368 that described in *M. persicae* (Amey et al. 2015): both genes occur in a proximate
369 genomic region but are on different strands (Supplementary Figure S8). While
370 susceptible strains are homozygous for the reference mutations in g11614 *para* subunit
371 2, all resistant strains are heterozygous (Supplementary Figure S9). These mutations have
372 not been observed in other pyrethroid-resistant arthropods (Dong et al. 2014). Moreover,
373 documented pyrethroid target-site mutations in aphids appear to evolve in subunit 1
374 (Amey et al. 2015; Zuo et al. 2016; Valmorbida et al. 2022), not subunit 2. We thus take
375 the conservative interpretation that mutations observed in our study are unlikely to be

376 involved in pyrethroid resistance, and that there is no evidence for target-site resistance
377 in *A. kondoi* generally.

378

379 ## Discussion

380 The evolutionary origins and repeatability of adaptation have long intrigued evolutionary
381 biologists (Cohan and Hoffmann 1989; Korona et al. 1994; Colosimo et al. 2005; Roda et
382 al. 2013; Tepolt et al. 2021). Here, we developed the first reference nuclear genome and
383 partial mitogenome for the aphid, *Acyrthosiphon kondoi,* an economically important
384 global pest. We used these new genomic resources to study the evolutionary origins and
385 molecular mechanisms of insecticide resistance in its invasive Australian range.
386 Insecticides are the major way in which *A. kondoi* is managed in Australian agriculture,
387 with organophosphates, carbamates, and synthetic pyrethroids widely used in alfalfa
388 growing regions. This prolonged selection pressure has led to the evolution of resistance,
389 first detected after farmers reported chemical control failures in a small number of field
390 populations in 2018 (Chirgwin et al. 2023). Although not registered to control *A. kondoi*
391 in most cropping situations and thought to be infrequently used against this species in
392 Australia, resistance was also detected to multiple synthetic pyrethroids, namely alpha-
393 cypermethrin and gamma-cyhalothrin (Chirgwin et al. 2023). Insecticide resistance is a
394 trait that is often broadly beneficial in pest organisms due to the use of agrichemicals
395 across large areas (Chevillon et al. 1999; Endersby-Harshman et al. 2020; Zhang et al.

396 2023). Moreover, clonality in asexual aphids can provide the added benefit of rapidly
397 spreading resistant lineages in local populations and facilitating recovery following major
398 selection events (Muller 1932; Crow and Kimura 1965; Vorburger 2006; Loxdale 2008;
399 Loxdale et al. 2010).

400 Our genomic investigation shows that all resistant clonal strains shared a single
401 evolutionary origin and are highly differentiated from all other susceptible clonal strains
402 examined here. The resistant lineage exhibits resistances to all three insecticide classes
403 (organophosphates, carbamates, and pyrethroids) used to control it in Australia,
404 undoubtedly providing a selective advantage for the aphid within the context of
405 Australia's alfalfa and legume crops where there are limited chemical control options.
406 Resistant strains analysed in this study were collected hundreds of kilometres apart. This
407 geographic spread of a single resistant clonal lineage of *A. kondoi* is characteristic of
408 superclones in other aphid species. In Australian populations of *Myzus persicae*, more
409 than 300 unique clones have been characterised using microsatellites, yet there is very
410 limited genetic variation among aphids collected from agricultural fields from different
411 areas due to selective sweeps of resistant superclones (Umina et al. 2014; de Little and
412 Umina 2017; Thia, Zhan, et al. 2023). Similarly, a resistant clone of the English grain aphid,
413 *Sitobion avenae*, has displaced sexually reproducing lineages in the UK, resulting in
414 reduced genetic variation in the field (Morales-Hojas et al. 2020). An important question
415 arising from this work is how long resistance will persist and (or) continue to spread in

416 the field. This will depend on whether selection pressures can be relaxed to reduce
417 resistance by exploiting costs of resistance, as observed in other arthropod pests (Foster
418 et al. 1999; Bass et al. 2014; Zhang et al. 2015; Maino et al. 2018) However, if the costs of
419 resistance are minimal, resistant *A. kondoi* clones may continue to persist even if
420 insecticide usage is reduced.

421 The most common mechanisms of evolved insecticide resistance involve target-site
422 changes and the evolution of detoxification mechanisms. In this work, we did not find
423 any clear evidence that resistance in *A. kondoi* is linked to target-site mechanisms. We
424 identified no non-synonymous mutations in any annotated *ace* or *ace*-like genes. The
425 only three non-synonymous mutations detected were in the g11614 *para* gene, but
426 these are not associated with resistance in other arthropods. Instead, our data provide
427 strong evidence that insecticide resistance in *A. kondoi* has evolved through
428 detoxification mechanisms. Using transcriptomic analyses, we identified a type B
429 carboxylesterase, an E4-like esterase, as a candidate gene for resistance. Relative to the
430 susceptible strains, resistant strains exhibited over-expression of this E4-like esterase in
431 the range of 29-fold to 184-fold in qPCR assays and >200-fold in transcriptomic analyses.

432 The *A. kondoi* E4-like esterase is an intriguing candidate for resistance given is its
433 similarity to the E4/FE4 esterases in *M. persicae*, which are known to be major detoxifiers
434 of organophosphate, carbamate, pyrethroid insecticides (Field and Devonshire 1998;
435 Field and Foster 2002). Homologs of this esterase family have also been identified in

436 other aphid species with evolved resistance. In the wheat aphid, *Schizaphis graminum*, an
437 E4/FE4 esterase homolog has been linked to organophophate resistance (Ono et al.
438 1999). In the pea aphid, *A. pisum*, two recently described pyrethroid resistant strains
439 appear to over-express an FE4 esterase homolog, although it is still unclear whether this
440 esterase contributes to resistance (Müller et al. 2023). In both *M. persicae* and *S.*
441 *graminum*, over-expression of these esterases is achieved through gene amplification,
442 with additional regulation through methylation patterns (Hick et al. 1996; Field et al.
443 1999; Ono et al. 1999; Field 2000). Whilst there appears to be an intriguing
444 macroevolutionary parallelism at the gene family level between *M. persicae*, *S. graminum*
445 and *A. kondoi*, the regulatory mechanism in *A. kondoi* appears to be different in that it is
446 not associated with gene amplification. Exploring the role of promoter sequence or
447 epigenetic variation is a possible avenue for future research.

448 Functional analysis of the *A. kondoi* E4-like esterase demonstrated that this enzyme is a
449 general detoxifier of insecticides. In our bioassays, transgenic *Drosophila* ectopically
450 expressing the *A. kondoi* E4-like esterase were significantly more resistant to alpha-
451 cypermethrin (a pyrethroid), chlorpyrifos (an organophosphate), and pirimicarb (a
452 carbamate). This highlights the ability for this esterase to provide cross-resistance across
453 different chemical groups, effectively acting as a gene with pleiotropic effects. Our
454 current data does not support the role of protein-coding mutations in the E4-like
455 esterase contributing to resistance in *A. kondoi*. None of the six non-synonymous

456 mutations observed in this study occurred in the active site or were proximal to the
457 catalytic triad (Ser238, Glu363, His487, *A. kondoi* numbering) that mediates hydrolysis
458 (Cygler et al. 1993; Hatfield et al. 2016). Although it is possible that mutations distal to an
459 active site can impact catalytic activity (Amichot et al. 2004), we did not see differences in
460 our transgenic *Drosophila* experiments. The resistant-like H3 *Drosophila* strain carrying
461 Ala28Val, Ala61Thr, and Ala527Thr, did not exhibit greater resistance than the
462 susceptible-like H1 strain carrying no mutations. A caveat of this work was that we were
463 unable to compare the H2 *Drosophila* strain carrying the Met150Leu and Pro586Leu
464 mutations to the H1 and H3 strain, and we were also unable to generate a strain
465 containing the Lys452Gln mutation. However, the Met150Leu and Pro586Leu are
466 heterozygous in susceptible *A. kondoi*, and resistant *A. kondoi* are homozygous for the
467 reference Lys452 allele. Hence, Met150Leu, Lys452Gln and Pro586Leu are weaker
468 candidates for resistance. There is scope for additional experimental work and protein
469 modelling to further test the hypothesis that protein-coding differences in the E4-like
470 esterase contribute to resistant phenotypes. But for now, we suggest that over-
471 expression is the main mechanism through which this gene contributes to resistance in *A.*
472 *kondoi*.

473 We note that whilst resistance in our mutant *Drosophila* strains was significant, it was
474 also modest, with resistance ratios ranging from 1.81 to 4.88, relative to the non-mutant
475 *Drosophila* strain. This in a sense mirrors the modest to moderate resistance ratios

476 observed in *A. kondoi*, which has been reported in the range of 6.80 to 20.26 depending
477 on the resistant strain and insecticide tested (Chirgwin et al. 2023). Although our data
478 supports the E4-like esterase as an important gene contributing to insecticide resistance,
479 there are likely other mechanisms that contribute to the phenotypic variance in
480 resistance between the *A. kondoi* genetic lineages. In our differential expression analyses,
481 for example, cytochrome P450s, UDP-glucuronosyltransferases, and an ABCB transporter
482 were also among the top over-expressed genes. Exploring the roles of these other genes
483 will be the subject for future work.

484 Our reference nuclear and (partial) mitochondrial genomes for *A. kondoi,* together with
485 the detailed transcriptomic analysis, provides an important contribution to the limited
486 genomic resources available for this economically significant aphid species. The high
487 BUSCO score (97.5%) for our nuclear genome indicates that this is a high-quality
488 assembly, with 615 contigs and a total length of 443.8 Mb. Although we were only able
489 to assembly a partial mitogenome, this still represents the longest mitochondrial
490 sequences from *A. kondoi* to date, capturing most of the protein coding and tRNA genes,
491 and both the rRNA genes. These genomic resources allowed us to uncover key insights
492 into the evolution and genetic mechanisms underlying insecticide resistance in *A. kondoi*
493 and assisted in developing diagnostic markers for the resistant genetic lineage. Such
494 diagnostic markers will facilitate the collection of spatiotemporal data on resistance,
495 providing insights into the dynamics of the resistant clonal lineage.

496 ## Concluding remarks

497 In this work, we disentangled the hypotheses of parallel evolution vs the spread of
498 insecticide resistance in *A. kondoi*, an important agricultural pest. Through genomic and
499 transcriptomic analyses, we revealed that a major component of insecticide resistance in
500 *A. kondoi* is due to detoxification by an over-expressed E4-like esterase. This esterase
501 shares homology with other resistance-conferring esterases in different aphid species.
502 Our work also highlights the rapid spread of a single resistant superclonal lineage,
503 reflecting the influence of human-induced selection pressures on pest populations.
504 While our findings suggest a parallel macroevolutionary response to insecticide
505 exposure in *A. kondoi* and other aphid species, they also suggest a unique regulatory
506 mechanism of the *A. kondoi* E4-like esterase, warranting further investigation. The cross-
507 resistance provided by the E4-like esterase (and other potential detoxification genes)
508 poses a serious challenge to management because multiple insecticide classes may
509 reinforce selection on the same genetic mechanism(s). Given the rapid spread of
510 resistance in Australia, there is a need to develop alternative management strategies for
511 *A. kondoi*, including expanding chemical control options and enhancing the use of
512 biological control.

513

514 ## Materials and Methods

515 ## Aphid strains

516 Eight clonal isofemale strains formed the focal strains for our experiments. These strains
517 were generated with artificial selection using leaf-dip insecticide bioassays (Umina et al.
518 2014). Nymphs from Keith, Temora, Bangham and Coombe were exposed to alfalfa
519 leaves dipped in 15 mg a.i./L chlorpyrifos (Lorsban® 500EC, Corteva) for 72 hours. This
520 concentration was previously shown to effectively separate resistant aphids from
521 insecticide-susceptible individuals (Chirgwin et al. 2023). The chlorpyrifos solution was
522 prepared by diluting 15 mg a.i./L in distilled water. Leaves were dipped in the insecticide
523 solution for 10 seconds and then plated onto 50 mm Petri dishes containing 1% agar.
524 Three dishes were prepared for each strain, each containing three large alfalfa leaves and
525 10 three-day-old nymphs. After 72 hours, surviving aphids were separated and
526 individually plated on 30 mm Petri dishes with a single untreated alfalfa leaf on agar.
527 Approximately eight days later, individuals that survived and produced nymphs were
528 retained in separate Petri dishes. After three generations, the healthiest surviving
529 replicate was picked to start an isofemale strain. Such strains were also generated for the
530 susceptible Kellerberrin, Eudunda, Jung, and Moulamein strains, but these were not
531 exposed to artificial selection. A follow up bioassay using the same discriminating dose
532 of chlorpyrifos (15 mg a.i./L) confirmed the resistance status of these isofemale strains
533 (Supplementary Table S1).

534 We later sampled additional field strains to broaden the geographic scope of our study,
535 which were from Albany, Avalon, Kaniva, Canowindra, Coolah, Cummins, Kerang, and
536 Manoora. Whilst we know the resistance status of these field strains (Evatt Chirgwin,
537 unpublished data), we did not subject them to artificial selection and screening
538 procedures in the same way as for our core experimental strains. Single individuals were
539 sampled from these field strains for genetic analysis (see below). Genomic DNA from
540 these individuals was extracted using a Qiagen EZ2 Tissue Kit.

541 Genome assembly and annotation

542 **Nuclear genome.** We developed the first reference for the *A. kondoi* nuclear genome to
543 investigate the genomic basis of insecticide resistance. The susceptible (laboratory)
544 Kellerberrin strain was used for the assembly. We performed two high molecular weight
545 DNA extractions from 22 and 24 pooled adult aphids, respectively, using the Nanobind®
546 Tissue Kit following a modified 'UHMW DNA Extraction – Dounce Homogenizer' protocol.
547 However, we replaced the Dounce Homogeniser with a simpler homogenisation step,
548 crushing aphids in a 1.5 mL Eppendorf tube with a pestle chilled on ice. The tissue
549 homogenate, lysis buffer, and proteinase K were incubated overnight at 50°C. All other
550 steps were as per the protocol. The DNA was sent to Novogene (Singapore) for library
551 preparation, long-read sequencing using the Oxford Nanopore Technology (ONT)
552 PromethION platform, and short-read Illumina paired end (150 bp) sequencing on the
553 NovaSeq platform.

554 We also sequenced a transcriptome from 10 pooled 14-day-old apterous aphids from
555 the Kellerberrin strain. To synchronize aphids, 15 adults were placed onto a 100 mm Petri
556 dish with alfalfa leaves in agar. These aphids were given an initial 24 hours to settle, after
557 which any offspring produced were removed. After a further 24 hours, the adults were
558 removed and any offspring produced were left behind on plates. Offspring were re-
559 plated every 3 to 4 days, and any showing signs of developing into alates were removed.
560 After 14 days, apterous aphids were pooled, frozen in liquid nitrogen, and sent to AGRF
561 (Australian Genome Research Facility, Melbourne) for RNA extraction, library preparation,
562 and short-read Illumina sequencing (150 bp) on the NovaSeq platform. In total, we
563 obtained 7,076,013 ONT long-reads, 191,891,532 paired reads from the Illumina whole-
564 genome sequencing, and 11,532,979 paired reads from Illumina transcriptome
565 sequencing. Illumina reads were quality trimmed using FASTP v0.23.2 (Chen et al. 2018).

566 Whole-genome short reads were used to estimate the genome size using a combination
567 of the KMC v3.1.2 (Kokot et al. 2017) and GENOMESCOPE v2 (Vurture et al. 2017) programs.
568 Our preliminary assessment of genome size for *A. kondoi* was ~452 Mb, with a genomic
569 heterozygosity rate of ~0.87%, based on GENOMESCOPE. High-quality ONT long-reads
570 were assembled with FLYE v2.9 (Kolmogorov et al. 2019), followed by three iterative
571 polishing steps with whole-genome short reads using PILON v1.23 (Walker et al. 2014).
572 ONT long reads were mapped back onto this assembly with MINIMAP2 v2.17 (Li 2016) and
573 duplicated regions were purged using PURGE_HAPLOTIGS v1.0.4 (Roach et al. 2018).

574 Repetitive regions were annotated using RED v2018.09.10 (Girgis 2015). Repeats were
575 annotated using REPEATMODELER v2.0.4 (Flynn et al. 2020).

576 To annotate genic regions, RNA short reads were mapped to the curated assembly with
577 HISAT2 v2.1.0 (Kim et al. 2019). All short-read alignments were filtered for quality (Q20)
578 and duplicate reads and sorted using SAMTOOLS v1.17 (Li et al. 2009; Danecek et al. 2021).
579 Gene annotation was performed with the BRAKER v.2.1.6 pipeline. The output AUGUSTUS
580 (Stanke et al. 2006) gene annotations were retained for downstream analysis. Parsing of
581 annotations and extraction of the longest protein coding transcript was performed
582 through a combination of GFFREAD v0.12.7 (Pertea and Pertea 2020), AGAT v1.1.0 (Dainat
583 2022), and SEQKIT v2.0.0 (Shen et al. 2016). We assigned function to gene annotations
584 using a combination of BLASTP v0.23.2 (Altschul et al. 1990) and BLAST2GO (Conesa et al.
585 2005; Conesa and Götz 2008). We downloaded the entire 'nr' database on the 10th of
586 June 2023. We used BLASTP to match the amino acid sequences of the longest transcripts
587 against all arthropod sequences in the 'nr' database. Gene ontology (GO) terms were
588 assigned using INTERPROSCAN v5.59-91.0 (Zdobnov and Apweiler 2001) and revaluated
589 using BLAST2GO. Assembly statistics were generated using QUAST v5.1.0 (Gurevich et al.
590 2013) and the completeness of single copy ortholog genes was assessed using BUSCO
591 v3.1.0 (Simão et al. 2015).

592 We further assessed our nuclear assembly for redundancy using a self-to-self alignment
593 with MINIMAP2. We also assessed completeness by aligning our *A. kondoi* assembly

594 against the *A. pisum* chromosome-level assembly (NCBI Accession GCF_005508785.2).
595 We first used the program RAGTAG (Alonge et al. 2022) to generate pseudo-
596 chromosomes for *A. kondoi* using the *A. pisum* reference genome as a guide. The
597 pseudo-chromosomes were mapped back to the *A. pisum* reference genome with
598 MINIMAP2. We visualised dotplots and synteny between the *A. kondoi* pseudo-
599 chromosomes and *A. pisum* chromosomes with R's PAFR v0.0.2 package (Winter 2020).

600 **Mitochondrial genome.** We assembled a partial mitogenome reference for *A. kondoi*.
601 We first mapped Illumina-short-reads from the Kellerberrin strain to our nuclear
602 reference genome using BOWTIE2 and extracted the non-mapped reads with SAMTOOLS.
603 The non-mapped reads were used to assemble a mitogenome reference with the
604 program MITOFINDER v1.4.2 (Allio et al. 2020), using the *A. pisum* mitogenome as a guide
605 (GenBank Accession NC_011594.1). We did attempt to perform an assembly from ONT
606 long-reads but found that these yielded poor assemblies relative to the short-reads
607 (unpublished data).

608 Differential gene expression

609 We examined patterns of differential expression among the Keith, Temora, Kellerberrin
610 and Eudunda strains to identify genes contributing to insecticide resistance in *A. kondoi*.
611 Transcriptomes were derived from pools of 12-day-old apterous aphids. As above, we
612 plated 15 adults onto 100 mm Petri dishes containing alfalfa leaves in agar alfalfa, with a
613 24-hour settling and culling period, and offspring after 48 hours were removed. These

614 offspring were replated onto 30 mm Petri dishes containing a single alfalfa leaf on agar
615 at densities of five to eight aphids per dish. After 8 days, aphids were moved to a 50 mm
616 Petri dish containing two alfalfa leaves on agar. On day 12, apterous aphids from each
617 strain were removed and snap frozen in liquid nitrogen. Four pools of aphids were used,
618 with each pool comprising approximately 20 to 25 aphids (approximately 10 to 15 mg of
619 tissue). RNA extraction was performed by OzOmics (Melbourne, Australia) using a
620 Monarch® Total RNA Miniprep Kit (New England BioLabs). Library preparation was
621 performed by AGRF (Melbourne) with 150 bp paired-end sequencing on their NovaSeq
622 platform. We successfully derived transcriptomes from all pools except one pool from
623 the Eudunda strain. Hence, n = 3 for Eudunda, and n= 4 for all other strains. For the
624 successfully sequenced transcriptomes, we obtained 35,232,161 to 52,797,301 raw reads,
625 with a mean of 40,452,782 reads.

626 Raw transcriptomic reads were trimmed using FASTP. Trimmed transcript reads were then
627 mapped to our assembled reference genome with HISAT2. Mapped reads were filtered to
628 retain high-quality mapped reads (MAPQ ≥ 30) and were sorted using SAMTOOLS.
629 Expression was quantified using STRINGTIE v2.2.1, and the STRINGTIE's PREPDE.PY script was
630 used to summarise data into a matrix of transcript abundances for all samples. This
631 transcript abundance matrix was imported into R for differential expression analysis
632 among strains. Global patterns of gene expression were visualised by analysing all strains

633 together with DESEQ2 v1.40.2 (Love et al. 2014) and performing a PCA on the transcript
634 abundances.

635 Differentially expressed genes between resistant and susceptible strains were identified
636 using a pairwise approach. This involved tests on all possible pairs of resistant and
637 susceptible strains (four total combinations), which allowed us to identify genes that
638 were on average differentially expressed irrespective of the strains compared. Differential
639 expression was determined with functions from the packages DESEQ2 v1.40.2 (Love et al.
640 2014) and APEGLM v1.22.1 (Zhu et al. 2019) in R. Our pairwise models predicted
641 expression as a function of 'Strain' as a categorical factor with two levels (one for each
642 strain). Within each pair, genes were ranked based on their *p*-value. We also separately
643 generated the top 500 over- and under-expressed genes per pair.

644 Next, we identified candidate genes for resistance. Because our study was primarily
645 interested in identifying detoxification genes, which are typically over-expressed in
646 resistant strains, we focused further analyses on the top over-expressed genes.
647 Candidate genes needed to satisfy two criteria: (1) they were in the top 500 genes in a
648 pairwise comparison, and (2) they were found across all four pairwise comparisons.
649 Mean ranks were generated based on the mean *p*-value across pairs.

650 Tests for GO term enrichment were performed on significantly over-expressed genes.
651 Any gene that was significantly over-expressed in all four pairwise comparisons was

652 considered a true outlier, and their GO terms used in enrichment analyses. These
653 analyses involved Fisher's exact test in R with the FISHER.TEST function, and *p*-values
654 adjusted for multiple testing with a Bonferroni correction.

655 Clonal genetic background

656 **Nuclear genome.** We tested whether resistant strains of *A. kondoi* were derived from
657 the same clonal genetic background. For these analyses, we combined transcriptomic
658 sequencing data from the Keith, Temora, Kellerberrin and Eudunda strains (above) with
659 additional whole-genome sequencing data from the Bangham, Coombe, Jung and
660 Moulamein strains. Whole-genome data was obtained by extracting genomic DNA from
661 6 pooled adult aphids using a Monarch® Genomic DNA Purification Kit (New England
662 BioLabs), one pool per strain. Whole-genome library preparation was performed by
663 Novogene (Hong Kong) with 150 bp paired-end sequencing on their NovaSeq platform.
664 We obtained 36,680,245 to 41,806,538 raw reads, with a mean of 40,405,084 reads.

665 Like the transcript reads, whole-genome reads were trimmed with FASTP, mapped with
666 HISAT2, and filtered, sorted and deduplicated with SAMTOOLS. We then called variants in
667 exonic regions by combining the mapped transcript and whole-genome reads using
668 BCFTOOLS (Li 2011; Danecek et al. 2021). We then reduced our dataset to only contain SNP
669 (single nucleotide polymorphism) loci with quality scores ≥30 using BCFTOOLS. Variants
670 were imported into R using GENOMALICIOUS v 0.7.8 (Thia and Riginos 2019) and further

671 filtered to only include SNPs with depth ≥15 reads, 0% missing data, and biallelic sites.
672 The final dataset contained 4,161 biallelic SNPs.

673 Pairwise genetic distances between samples were quantified as the number of allelic
674 differences across all biallelic loci. The pairwise genetic distance matrix was used to
675 derive a UPGMA clustering tree with the PHANGORN v2.12.1 package (Schliep 2011), with
676 negative branch lengths adjusted using the BAT v2.9.6 package (Cardoso et al. 2015).
677 Bootstrapping was undertaken with the APE v 5.8 package (Paradis et al. 2004). The
678 resulting tree was visualised with GGTREE v3.12.0 (Yu et al. 2017) and TREEIO v1.28.0 (Wang
679 et al. 2020).

680 **Mitochondrial genome.** We analysed mitogenomic variation as supporting data for the
681 clonal genetic background of our *A. kondoi* strains. We first mapped transcriptome and
682 whole-genome sequencing reads for our core strains against our mitogenome reference
683 with HISAT2 and filtered alignments with SAMTOOLS to remove reads with quality scores
684 <30. Variants were called with BCFTOOLS and were filtered to remove those with quality
685 scores <30. Variants were imported into R and analysed with a combination of DATA.TABLE
686 (Dowle and Srinivasan 2019), TIDYVERSE (Wickham et al. 2019), and GENOMALICIOUS. From
687 these variants, we identified three diagnostic mitochondrial SNPs that distinguished
688 resistant and susceptible strains: positions 2,597 (in *cytb*), 9,857 (in *cox*1), and 11,363 (in
689 *cox*2).

690 We developed primers to amplify mitogenomic regions containing the three diagnostic
691 SNPs using GENEIOUS PRIME (Supplementary Table S5). These primers were used to
692 genotype representative individuals from the field strains. PCR amplification was
693 performed in 25 uL volume comprising 1.25 uL $MgCl_2$ (50 mM), 2.5 uL ThermoPol
694 standard buffer (Mg free), 1.25 uL each of forward and reverse primer (10 mM), 2 uL
695 dNTPs (2.5 mM), 0.4 uL IMMOLASE Taq polymerase, and 1.5 uL template DNA. PCR
696 conditions were denaturation at 95°C for 10 mins, 45 cycles of 94 °C for 1 min, 56 °C for
697 1 min and 72 °C for 1 min, with a final extension at 72 °C for 7 mins. Amplicons were sent
698 to Macrogen (South Korea) for Sanger sequencing in both forward and reverse
699 directions. Raw sequences were quality checked and formed into consensus sequences
700 in GENEIOUS PRIME. We also used GENEIOUS PRIME to map consensus sequences to our
701 mitogenome reference and these alignments were used to score genotypes at our
702 diagnostic SNPs.

703 E4-like esterase genomic copy number, expression, and sequence 704 variation

705 **Copy number and expression.** Our top candidate over-expressed gene for resistance
706 was an E4-like esterase, which was over-expressed in resistant strains relative to
707 susceptible strains (see Results). We used qPCR to validate the expression levels of the
708 E4-like esterase, and to test whether elevated genomic copy number was driving over-
709 expression. We performed two separate experiments with aphids for qPCR being
710 obtained using the same protocols as those for obtaining transcriptomes (see Section,

711 'Differential gene expression'). In Experiment 1, we measured individual levels (one aphid
712 per sample) of genomic copy number and expression for the Eudunda, Kellerberrin, Keith
713 and Temora strains. These individual measures also allowed us to test for expression
714 differences between morphs (winged alate versus non-winged apterous aphids). For
715 each strain, we analysed 24 samples of individual aphids. In Experiment 2, we measured
716 genomic copy number and expression in pooled samples (two aphids per sample) for
717 the Jung, Moulamein, Bangham and Coombe, Eudunda, Kellerberrin, Keith. Aphids for
718 both experiments were obtained using the same protocols as those for obtaining
719 transcriptomes (see Section, 'Differential gene expression'). The Temora strain was not
720 included in Experiment 2 due to a lack of nymphs during the age-matching step. In
721 Experiment 2, we analysed five to six samples (two aphids per sample) because our
722 findings from Experiment 1 indicated that individual levels of expression were consistent
723 within strains and that there were no differences among morphs (see Results).

724 We developed a workflow that allowed us to simultaneously extract both the DNA and
725 RNA from the same sample. Aphid samples (single aphids, or a pool of two aphids) were
726 frozen in liquid nitrogen in a 1.5 mL Eppendorf tube and stored for up to two weeks at –
727 80°C. These were homogenized on ice with a small plastic pestle. The used pestle was
728 then placed into a second Eppendorf tube containing 150 µL of 5% Chelex® buffer and
729 proteinase K (20mg/mL, Meridian, CSA-01254; Cincinnati, OH, USA) and swirled in this

730 buffer mix to suspend the residual aphid homogenate – this was used for DNA
731 extraction. We followed the protocol in Yang et al. (2023) for Chelex® DNA extraction.

732 Simultaneously, 200 µL of RNA protection reagent from the Monarch® Total RNA
733 Miniprep Kit (New England BioSciences, Notting Hill, VIC, Australia; #T2010) was added
734 directly to the homogenized aphid tissue in the original sample tube; this was used for
735 RNA extraction. We performed RNA extraction following the Monarch® Total RNA
736 Miniprep Kit protocol. Extracted RNA was converted into cDNA using the ThermoFisher
737 Scientific High-Capacity cDNA Reverse Transcription Kit (Parkville, VIC, Australia;
738 #4368814) as per manufacturer's instructions.

739 We used qPCR to quantify the relative expression of the candidate E4-like esterase
740 across clonal strains of *A. kondoi*. DNA extractions were diluted 1:3 and cDNA samples
741 were diluted 1:10 with PCR water. A set of primers for the target gene was designed in
742 GENEIOUS PRIME for the E4 esterase: Akond_E4_qPCR_F, CCTCGTCTTGGGTAACACTCA, and
743 Akond_E4_qPCR_R, CATTTTCGCCGTTTGGTCCA. The aphid actin gene provided a
744 reference gene with primers (Yang et al. 2023): actin_aphid_F1,
745 GTGATGGTGTATCTCACACTGTC; and actin_aphid_R1, AGCAGTGGTGGTGAAACTG. Our
746 qPCR reactions were performed according to Lee et al. (2012) with a Roche LightCycler
747 480. Briefly, 10-minute incubation at 95°C was used to heat activate the quick start
748 IMMOLASE (Meridian, BIO-21047; Cincinnati, OH, USA), which was followed by 45 cycles
749 of 95°C for 10 seconds, 54°C for 20 seconds, and 72°C for 15 seconds. The cycle ended

750 with a high-resolution melt of 95°C for 1 minute (ramp rate = 4.8°C/second), 40°C for 1
751 minute (ramp rate = 2.5°C), and then an increase to 65°C (ramp rate = 4.8°C/second).
752 Fluorescence data was collected continuously from 65°C to 95°C (ramp rate =
753 0.02°C/second, 25 acquisitions per °C). Prior to our main experiment, we validated the
754 amplification efficiency of primers using the Kellerberrin strain (Supplementary Figure
755 S4).

756 A mean C_P (crossing-point) value for each aphid was calculated over two replicate qPCR
757 reactions. We calculated the $-\Delta C_P$ as –(E4-like esterase C_P – actin C_P) for each aphid. The
758 $2^{-\Delta\Delta C_P}$ method (Livak and Schmittgen 2001) was used to estimate the relative expression
759 (fold-change) between resistant and susceptible strains. We calculated the $2^{-\Delta\Delta C_P}$ score
760 between each pair of resistant vs susceptible strains within each experiment. Within a
761 strain comparison, we calculated $2^{-\Delta\Delta C_P}$ between all pairs of aphids from both strains.
762 This allowed us to obtain a median value with 2.5% and 97.5% percentiles for the
763 estimated $2^{-\Delta\Delta C_P}$.

764 ***Sequence variation.*** We characterized the distribution of segregating polymorphism in
765 the exonic regions of our candidate E4-like esterase. For the eight core strains, we used
766 SNPs that were not filtered for missing data and determined the protein coding effects
767 through custom R code, combining the R packages BIOSTRINGS, DATA.TABLE (Dowle and
768 Srinivasan 2019), TIDYVERSE (Wickham et al. 2019), SEQINR (Charif and Lobry 2007), and
769 GENOMALICIOUS. We identified six non-synonymous SNPs that produced diagnostic

770 genotypic differences between resistant and susceptible strains. Like our analysis of
771 diagnostic mitochondrial SNPs (see above section 'Clonal genetic background'), we
772 developed a set of primers for the diagnostic non-synonymous E4-like esterase SNPs
773 using GENEIOUS PRIME (see Supplementary Table S5). We then used these primers to
774 genotype representative individuals from the field strains, using the same PCR protocol
775 and analysis pipeline as derived from the diagnostic mitochondrial SNPs (see above
776 section 'Clonal genetic background').

777 We performed protein sequence analysis to infer possible functional effects of the non-
778 synonymous mutations. First, we aligned our *A. kondoi* E4-like esterase protein sequence
779 against the *Myzus persicae* E4 esterase (UniProt Accession P35501) and the *Torpedo*
780 *californica* acetylcholinesterase (UniProt Accession P04058) using GENEIOUS PRIME. We
781 included the *T. californica* acetylcholinesterase because it is a model for type B
782 carboxylesterases. Our alignments were used to understand whether the *A. kondoi* non-
783 synonymous mutations occurred in or near the active site. Annotation of the active site
784 was performed with the InterProScan server (Zdobnov and Apweiler 2001) as
785 implemented on the EMBL-EBI server (https://www.ebi.ac.uk/interpro/search/sequence/).
786 Second, we performed protein structural modelling of the *A. kondoi* E4-like esterase
787 using the ALPHAFOLD 3 server (https://alphafoldserver.com/) (Abramson et al. 2024).
788 Quality of the predicted structure was assessed using ERRAT (Colovos and Yeates 1993)

789 and PROCHECK (Laskowski et al. 1996) as implemented on the SAVES v6.1 server

790 (https://saves.mbi.ucla.edu/).

791 E4-like esterase transgenesis

792 Ectopic expression of our candidate *A. kondoi* E4-like esterase in *Drosophila*

793 *melanogaster* (herein, '*Drosophila*') was used to validate the role of this gene in

794 insecticide resistance. Our analysis of sequence variation in the candidate E4-like

795 esterase revealed non-synonymous differences between the susceptible and resistance *A.*

796 *kondoi* strains (see Results). We were able to successfully obtain three strains of *D.*

797 *melanogaster* carrying different E4-like esterase haplotypes, which we call H1, H2, and

798 H3. H1 was identical to the susceptible reference genome sequence. H2 contained the

799 mutations, Met150Leu and Pro586 Leu, which are heterozygotes in susceptible *A. kondoi*

800 and homozygous in resistant *A. kondoi*. H3 contained three mutations, Ala28Val,

801 Ala61Thr, and Ala527Thr, which are absent in the susceptible *A. kondoi*, and homozygous

802 in the resistant *A. kondoi*.

803 The E4-like esterase was synthesised (Twist Bioscience) and cloned into the pUASTattB

804 plasmid between the EcoRI and XbaI sites. (GenBank: EF362409.1). The construct was

805 then transformed into the *Drosophila* germline, which carried an attP docking site on

806 chromosome 2 (attP40) and the phiC31 integrase gene under the control of the vasa

807 regulatory region on the X chromosome (y w M (eGFP, vas-int, dmRFP)ZH-2A: P

808 [CaryP]attP40). Transgenic strains were balanced, and the integration of the transgenes
809 confirmed using PCR and sequencing.

810 To conduct the bioassay, virgin females of the Act5C-GAL4 strain ("y[1] w[∗]; P(Act5C-
811 GAL4-w)E1/CyO,""1;2") (Bloomington Stock Center) were crossed with UAS-E4-like males
812 to allow mating and oviposition. We exposed adult female flies to size concentrations of
813 technical grade pirimicarb, chlorpyrifos and alpha-cypermethrin. The insecticide
814 solutions were diluted in water and overlaid on 1.5% agar (1% sucrose, 0.5% acetic acid)
815 in standard fly vials, then allowed to dry at room temperature overnight. Twenty adult
816 females (2 to 7 days post-eclosion) were added to each vial and mortality assessed after
817 48 hours (alpha-cypermethrin) or 72 hours (pirimicarb and chlorpyrifos). Five replicates
818 were conducted for each concentration. A *Drosophila* strain of the same genetic
819 background as the two transgenic strains but lacking a transgene (progeny of crosses of
820 the Act5C-GAL4 and UAS-no transgene strains) was included in all bioassays as a
821 reference control. Control mortality was assessed using vials containing agar but lacking
822 insecticide.

823 Our bioassays were performed as two experiments. We initially acquired the H2 mutant
824 line before H1 and H3. In our first bioassay, Experiment 1, we contrasted the H2 mutant
825 against the Control line. We were later able to acquire the H1 and H3 mutant lines.
826 Unfortunately, in the time between experiments, the H2 mutant line was lost. In our
827 second bioassay, Experiment 2, we contrasted the H1 and H3 mutants against the

828 Control line. We acknowledge that running all three mutant lines in the same bioassay
829 would have been ideal, but due to the circumstances, this was not possible. However, we
830 note that Experiment 2 did contain lines carrying a susceptible version (H1) and a
831 resistant version (H3) of the E4-like esterase, which provided a functional contrast of
832 haplotypes carried in susceptible and resistant *A. kondoi*.

833 Dose-response data was analysed in R. For each insecticide and *Drosophila* strain
834 combination, we fitted a quasibinomial model with mortality as a function of the log10-
835 transformed insecticide concentration, using the GLM function from the STATS package.
836 We derived χ^2 goodness-of-fit test statistics and *p*-values (Brown 1978) using custom
837 code. LD50-values were obtained with the DOSE.P function from the MASS package.
838 Resistance ratios for the mutant strain were calculated as the mutant LD50-value divided
839 by the control LD50-value for each insecticide. Models were fit separately for each
840 experiment.

841 Target-site mutations

842 We used our combined transcriptome and whole-genome data to characterise allelic
843 variation in target genes for organophosphate and carbamate insecticides (the
844 acetylcholinesterase gene, *ace*) and pyrethroid insecticides (the *para*-like voltage-gated
845 sodium channel, *para*). Our assembly of the *A. kondoi* genome identified three genes
846 with *ace* annotations, with gene IDs, g11959, g46, and g15835, and lengths of 822, 706
847 and 664 amino acids, respectively. There were two genes with *para* annotations, with

848 gene IDs, g11612 and g11614, and lengths of 1,160 and 958 amino acids, respectively.
849 Sequence analysis was performed in R. To determine the relative codon positions of the
850 target gene proteins, we aligned the *A. kondoi* sequences to their respective canonical
851 reference species: *ace* sequences to *Torpedo californica* (X03439) and *para* sequences to
852 *Musca domestica* (NP_001273814). Protein alignments were performed using the MSA
853 v1.36.1 package (Bodenhofer et al. 2015).

854 We next determined whether non-synonymous mutations in our target genes might be
855 associated with insecticide resistance. Here, we used SNPs that were not filtered for
856 missing data and considered all possible alleles (not just biallelic loci) that occurred in
857 the exonic regions of the annotated *ace* and *para* genes. Protein coding annotations
858 were made with SNPEFF (Cingolani et al. 2012). These were then imported into R, where
859 we extracted information about relative codon positions from our *A. kondoi* alignments
860 to canonical reference proteins and matched these to our protein coding annotations
861 using GENOMALICIOUS. We were most interested in any non-synonymous mutations in the
862 primary *A. kondoi ace* and *para* gene, especially if they had been previously reported in
863 the literature as conferring resistance.

864 We only found mutations in one of the *para* genes (see Results). To better contextualise
865 these mutations, we analytically validated the functional roles of these annotated *para*
866 genes. Aphids have evolved a heterodimeric *para* sodium channel, controlled by two
867 genes, subunit 1 (domains I and II) and subunit 2 (domains III and IV) (Amey et al. 2015).

868 We therefore compared our *A. kondoi para* genes to two other aphids, the congeneric *A.*
869 *pisum* and *M. persicae*. We used the *Myzus persicae* protein sequences of *para* subunits 1
870 and 2 (NCBI Protein Accessions CBI71141.1 and CBI71142.1, respectively) (Amey et al.
871 2015) to query the *Acythrosiphon pisum* genome (NCBI Genome Accession
872 GCF_005508785.1) in NCBI's BLASTP platform. We then recovered two respective top hits,
873 with NCBI Protein Accession XP_029346126.1 (Gene ID 100158802) and XP_001949648.2
874 (Gene ID 100164620). In R, we aligned these *M. persicae* and *A. pisum* protein sequences
875 with our *A. kondoi* sequences and that of the canonical *M. domestica* reference sequence,
876 using MSA, visualizing the alignment with GENOMALICIOUS.

877

878 ## Acknowledgements

879 We thank Alex Gill from the University of Melbourne for his advice and support on
880 culturing *A. kondoi* in the lab. This work was supported by funding through the
881 University of Melbourne's 'Early Carer Researcher Global Collaborations Award' (JAT),
882 and investment from the Grains Research and Development Corporation (UOM1906-
883 002RTX; UOM2404-006RT) and Hort Innovation (ST23002) through the 'Australian Grain
884 and Horticulture Pest Innovation Program' (AAH and PAU) with additional support
885 provided by the University of Melbourne and Cesar Australia. AgriFutures Australia
886 provided additional financial support for sample collection (EC and PAU). The authors
887 would like to acknowledge the use of the University of Exeter's Advanced Research

888 Computing facilities in carrying out this work. For the purpose of open access, the
889 authors have applied a 'Creative Commons Attribution (CC BY) licence to any Author
890 Accepted Manuscript version arising from this submission.

891

892 ## Contributions

893 **University of Melbourne.** JAT, PAU, and AAH conceptualised the project. JAT
894 developed the isofemale strains of *Acythosiphon kondoi* used in this study. JAT prepared
895 all samples for transcriptomic and genomic sequencing. JAT co-led the nuclear genome
896 assembly and differential expression analysis. JAT led the mitogenome assembly, analysis
897 of clonal genetic backgrounds, experimental design and analysis of E4-like esterase
898 expression and copy number validation, analysis of target-site mutations, analysis of
899 diagnostic mitogenome and E4-like esterase SNPs, and analysis of dose-responses of
900 transgenic *Drosophila*. CJB led the development of molecular protocols for validating E4-
901 like esterase expression and copy number. RS led the molecular protocols for
902 genotyping diagnostic mitochondrial and E4-like esterase SNPs and contributed to their
903 analysis. Culturing and (or) molecular work associated with validating expression and
904 copy number was contributed to by CJB, LAS, MS, KR, CJ, APSD, and QY. The original
905 version of this manuscript was written by JAT, with contributions from AAH and PAU,
906 and additional comments made by RS, CJB, LAS, MS, KR, CJ, APSD, and QY. AAH, JAT and
907 PAU and acquired funding and performed project administration.

908 **University of Exeter.** BJH co-led the genome assembly and differential expression
909 analysis. SW and BT designed and executed the chemical bioassays of transgenic
910 *Drosophila*. BT led the protein structural analysis. Contributions to original manuscript
911 was made by CB, with additional comments made by BJH, BT, and SW. CB contributed
912 funds for sequencing and hosted JAT as a visiting fellow at the University of Exeter.

913 **Cesar Australia.** EC performed the original field collections of *A. kondoi*, identified the
914 discriminating doses used to separate resistant from susceptible clones. MB contributed
915 to molecular work associated with validating expression and copy number. Additional
916 comments to the manuscript were made by EC and MB. EC and PAU acquired funds for
917 this work.

918

919 ## Data availability

920 The nuclear reference genome for *Acrythosiphon kondoi* that was generated in this study
921 has been deposited into GenBank Genomes (JAUMJG000000000), and the partial
922 mitochondrial genome has been submitted to GenBank Nucleotides (PV793158).
923 Genomic and transcriptomic reads associated with this work have been submitted as a
924 GenBank BioProject PRJNA989388. The GenBank SRA accessions for sequencing reads,
925 and Nucleotide accessions for Sanger sequences, can be found in Appendix 2. Custom
926 gene annotations and counts of repetitive element classes that accompany our nuclear

927 reference genome can be found in Appendix 3. Data and scripts associated with this
928 work are archived through the University of Melbourne's FigShare server (Thia 2025):
929 doi.org/10.26188/29466431.

930

931 ## References

932 Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L,
933 Ballard AJ, Bambrick J. 2024. Accurate structure prediction of biomolecular
934 interactions with AlphaFold 3. *Nature* 630:493–500.

935 Allio R, Schomaker-Bastos A, Romiguier J, Prosdocimi F, Nabholz B, Delsuc F. 2020.
936 MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target
937 enrichment phylogenomics. *Mol Ecol Resour* 20:892–905.

938 Alonge M, Lebeigle L, Kirsche M, Jenike K, Ou S, Aganezov S, Wang X, Lippman ZB,
939 Schatz MC, Soyk S. 2022. Automated assembly scaffolding using RagTag elevates a
940 new tomato system for high-throughput genome editing. *Genome Biol* [Internet]
941 23:258. Available from: https://doi.org/10.1186/s13059-022-02823-7

942 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search
943 tool. *J Mol Biol* 215:403–410.

944 Amey JS, O'Reilly AO, Burton MJ, Puinean AM, Mellor IR, Duce IR, Field LM, Wallace BA,
945 Williamson MS, Davies TGE. 2015. An evolutionarily-unique heterodimeric voltage-
946 gated cation channel found in aphids. *FEBS Lett* 589:598–607.

947 Amichot M, Tarés S, Brun-Barale A, Arthaud L, Bride JM, Bergé JB. 2004. Point mutations
948 associated with insecticide resistance in the Drosophila cytochrome P450 Cyp6a2
949 enable DDT metabolism. *Eur J Biochem* 271:1250–1257.

950 Barraclough TG, Birky W, Burt A. 2003. Diversification in sexual and asexual organisms.
951 *Evolution (N Y)* 57:2166–2172.

952 Bass C, Puinean AM, Zimmer CT, Denholm I, Field LM, Foster SP, Gutbrod O, Nauen R,
953 Slater R, Williamson MS. 2014. The evolution of insecticide resistance in the peach
954 potato aphid, *Myzus persicae*. *Insect Biochem Mol Biol* 51:41–51.

955 Blackman RL, Eastop VF. 2017. Taxonomic issues. In: Van Emden HF, Harrington R,
956 editors. Aphids as Crop Pests. Wallingford, Oxfordshire, UK, and Cambridge,
957 Massachusetts, USA: CABI Publishing. p. 1–29.

958 Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. 2015. msa: an R package for
959 multiple sequence alignment. *Bioinformatics* 31:3997–3999.

960 Bontrager M, Angert AL. 2019. Gene flow improves fitness at a range edge under climate
961 change. *Evol Lett* 3:55–68.

962 Brown CC. 1978. The Statistical Analysis of Dose-Effect Relationships. In: Principles of
963 Ecotoxicology. London: John Wiley and Sons. p. 115–148.

964 Caballero JD, Clark ST, Coburn B, Zhang Y, Wang PW, Donaldson SL, Elizabeth Tullis D,
965 Yau YCW, Waters VJ, Hwang DM, et al. 2015. Selective sweeps and parallel
966 pathoadaptation drive pseudomonas aeruginosa evolution in the cystic fibrosis lung.
967 *mBio* 6:e00981-15.

968 Cardoso P, Rigal F, Carvalho JC. 2015. BAT–Biodiversity Assessment Tools, an R package
969 for the measurement and estimation of alpha and beta taxon, phylogenetic and
970 functional diversity. *Methods Ecol Evol* 6:232–236.

971 Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T. 2015. Tools and
972 best practices for data processing in allelic expression analysis. *Genome Biol* 195:1–
973 12.

974 Charif D, Lobry JR. 2007. SeqinR 1.0-2: a contributed package to the R project for
975 statistical computing devoted to biological sequences retrieval and analysis. In:
976 Structural approaches to sequence evolution: Molecules, networks, populations.
977 Springer. p. 207–232.

978 Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
979 *Bioinformatics* 34:i884–i890.

980 Chevillon C, Raymond M, Guillemaud T, Lenormand T, Pasteur N. 1999. Population
981 genetics of insecticide resistance in the mosquito. *Biological Journal of the Linnean*
982 *Society* 68:147–157.

983 Chirgwin E, Thia JA, Copping K, Umina PA. 2023. Discovery of insecticide resistance in
984 field-collected populations of the aphid pest, *Acyrthosiphon kondoi* Shinji. *Pest*
985 *Manag Sci* 80:1338–1347.

986 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM.
987 2012. A program for annotating and predicting the effects of single nucleotide
988 polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain
989 w1118; iso-2; iso-3. *Fly (Austin)* 6:80–92.

990 Cohan FM, Hoffmann AA. 1989. Uniform selection as a diversifying force in evolution:
991 evidence from Drosophila. *Am Nat* 134:613–637.

992 Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz
993 J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread Parallel Evolution in
994 Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science (1979)* 307:1928–
995 1933.

996 Colovos C, Yeates TO. 1993. Verification of protein structures: patterns of nonbonded
997 atomic interactions. *Protein Science* 2:1511–1519.

998 Conesa A, Götz S. 2008. Blast2GO: a comprehensive suite for functional analysis in plant
999 genomics. *Int J Plant Genomics* 2008.

1000 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a
1001 universal tool for annotation, visualization and analysis in functional genomics
1002 research. *Bioinformatics* 21:3674–3676.

1003 Crossley MS, Chen YH, Groves RL, Schoville SD. 2017. Landscape genomics of Colorado
1004 potato beetle provides evidence of polygenic adaptation to insecticides. *Mol*
1005 *Ecol*:6284–6300.

1006 Crow JF, Kimura M. 1965. Evolution in sexual and asexual organisms. *Am Nat*:439–450.

1007 Cygler M, Schrag JD, Sussman JL, Harel M, Silman I, Gentry MK, Doctor BP. 1993.
1008 Relationship between sequence conservation and three-dimensional structure in a
1009 large family of esterases, lipases, and related proteins. *Protein Science* 2:366–382.

1010 Dainat J. 2022. Another Gtf/Gff Analysis Toolkit (AGAT): Resolve interoperability issues
1011 and accomplish more with your annotations. In: Plant and Animal Genome XXIX
1012 Conference. Available from: https://github.com/NBISweden/AGAT

1013 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
1014 McCarthy SA, Davies RM. 2021. Twelve years of SAMtools and BCFtools. *Gigascience*
1015 10:giab008.

1016 Davies TGE, Field LM, Usherwood PNR, Williamson MS. 2007. DDT, pyrethrins,
1017 pyrethroids and insect sodium channels. *IUBMB Life* 59:151–162.

1018 Dong K, Du Y, Rinkevich F, Nomura Y, Xu P, Wang L, Silver K, Zhorov BS. 2014. Molecular
1019 biology of insect sodium channels and pyrethroid resistance. *Insect Biochem Mol Biol*
1020 [Internet] 50:1–17. Available from: http://dx.doi.org/10.1016/j.ibmb.2014.03.012

1021 Dowle M, Srinivasan A. 2019. data.table: Extension of "data.frame."

1022 Endersby-Harshman NM, Schmidt TL, Chung J, van Rooyen A, Weeks AR, Hoffmann AA.
1023 2020. Heterogeneous genetic invasions of three insecticide resistance mutations in
1024 Indo-Pacific populations of Aedes aegypti (L.). *Mol Ecol* 29:1628–1641.

1025 Feliziani S, Marvig RL, Luján AM, Moyano AJ, Di Rienzo JA, Krogh Johansen H, Molin S,
1026 Smania AM. 2014. Coexistence and Within-Host Evolution of Diversified Lineages of
1027 Hypermutable *Pseudomonas aeruginosa* in Long-term Cystic Fibrosis Infections.
1028 *PLoS Genet* 10.

1029 Feyereisen R, Dermauw W, Van Leeuwen T. 2015. Genotype to phenotype, the molecular
1030 and physiological dimensions of resistance in arthropods. *Pestic Biochem Physiol*
1031 121:61–77.

1032 Field LM. 2000. Methylation and expression of amplified esterase genes in the aphid
1033 *Myzus persicae* (Sulzer). *Biochemical Journal* 349:863–868.

1034 Field LM, Blackman RL, Tyler-Smith C, Devonshire AL. 1999. Relationship between
1035 amount of esterase and gene copy number in insecticide-resistant *Myzus persicae*
1036 (Sulzer). *Biochemical Journal* 339:737–742.

1037 Field LM, Devonshire AL. 1998. Evidence that the E4 and FE4 esterase genes responsible
1038 for insecticide resistance in the aphid *Myzus persicae* (Sulzer) are part of a gene
1039 family. *Biochemical Journal* 330:169–173.

1040 Field LM, Foster SP. 2002. Amplified esterase genes and their relationship with other
1041 insecticide resistance mechanisms in English field populations of the aphid, *Myzus*
1042 *persicae* (Sulzer). *Pest Manag Sci* 58:889–894.

1043 Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020.
1044 RepeatModeler2 for automated genomic discovery of transposable element families.
1045 *Proceedings of the National Academy of Sciences* 117:9451–9457.

1046 Foster SP, Woodcock CM, Williamson MS, Devonshire AL, Denholm I, Thompson R. 1999.
1047 Reduced alarm response by peach-potato aphids, *Myzus persicae* (Hemiptera:
1048 Aphididae), with knock-down resistance to insecticides (kdr) may impose a fitness
1049 cost through increased vulnerability to natural enemies. *Bull Entomol Res* 89:133–
1050 138.

1051 Fournier D. 2005. Mutations of acetylcholinesterase which confer insecticide resistance in
1052 insect populations. *Chem Biol Interact* 157–158:257–261.

1053 Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual
1054 population. *Genetica* 102:127–144.

1055 Gilabert A, Dedryver CA, Stoeckel S, Plantegenest M, Simon JC. 2015. Longitudinal clines
1056 in the frequency distribution of "super-clones" in an aphid crop pest. *Bull Entomol*
1057 *Res* 105:694–703.

1058 Girgis HZ. 2015. Red: an intelligent, rapid, accurate tool for detecting repeats de-novo on
1059 the genomic scale. *BMC Bioinformatics* [Internet] 16:227. Available from:
1060 https://doi.org/10.1186/s12859-015-0654-5

1061 Godefroid M, Meynard CN, Clamens AL, Popkin M, Jousselin E. 2024. On the role of niche
1062 specialization in the geographic distribution of aphid asexual lineages: a case study
1063 using the leaf-curl plum aphid superclones. *Oikos* 2024.

1064 Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for
1065 genome assemblies. *Bioinformatics* 29:1072–1075.

1066 Hatfield MJ, Umans RA, Hyatt JL, Edwards CC, Wierdl M, Tsurkan L, Taylor MR, Potter PM.
1067 2016. Carboxylesterases: General detoxifying enzymes. *Chem Biol Interact* 259:327–
1068 331.

1069 Hick CA, Field LM, Devonshire AL. 1996. Changes in the methylation of amplified
1070 esterase DNA during loss and reselection of insecticide resistance in peach-potato
1071 aphids, Myzus persicae. *Insect Biochem Mol Biol* 26:41–47.

1072 Hoffmann AA, Reynolds KT, Nash MA, Weeks AR. 2008. A high incidence of
1073 parthenogenesis in agricultural pests. *Proceedings of the Royal Society B: Biological*
1074 *Sciences* 275:2473–2481.

1075 Howe DK, Sibley LD. 1995. Toxoplasma gondii Comprises Three Clonal Lineages:
1076 Correlation of Parasite Genotype with Human Disease. *J Infect Dis* 172:1561–1566.

1077 Humphries AW, Peck DM, Robinson SS, Rowe T, Oldach K. 2012. A new biotype of
1078 bluegreen aphid (*Acyrthosiphon kondoi* Shinji) found in south-eastern Australia
1079 overcomes resistance in a broad range of pasture legumes. *Crop Pasture Sci* 63:893–
1080 901.

1081 Huse HK, Kwon T, Zlosnik JEA. 2010. Parallel Evolution in Pseudomonas aeruginosa over
1082 39,000 Generations In Vivo. *mBio* 1:e00199-10.

1083 Kawada K. 1992. Production of sexuals in the aphid Acyrthosiphon kondoi in Japan
1084 (Homoptera: Aphidinea: Aphididae). *Entomologia Generalis* 17:115–119.

1085 Khan A, Taylor S, Ajioka JW, Rosenthal BM, Sibley LD. 2009. Selection at a single locus
1086 leads to widespread expansion of *Toxoplasma gondii* lineages that are virulent in
1087 mice. *PLoS Genet* 5:1–14.

1088 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment
1089 and genotyping with HISAT2 and HISAT-genotype. *Nature*
1090 *Biotechnologyotechnology* 37:907–915.

1091 Kokot M, Długosz M, Deorowicz S. 2017. KMC 3: counting and manipulating k-mer
1092 statistics. *Bioinformatics* 33:2759–2761.

1093 Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads
1094 using repeat graphs. *Nat Biotechnol* 37:540–546.

1095 Korona R, Nakatsu CH, Forney LJ, Lenski RE. 1994. Evidence for multiple adaptive peaks
1096 from populations of bacteria evolving in a structured habitat. *Proceedings of the*
1097 *National Academy of Sciences* 91:9037–9041.

1098 Korpelainen H. 1986. Competition between clones: An experimental study in a natural
1099 population of *Daphnia magna*. *Hereditas* 105:29–35.

1100 Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. 1996. AQUA and
1101 PROCHECK-NMR: programs for checking the quality of protein structures solved by
1102 NMR. *J Biomol NMR* 8:477–486.

1103 Lee SF, White VL, Weeks AR, Hoffmann AA, Endersby NM. 2012. High-throughput PCR
1104 assays to monitor Wolbachia infection in the dengue mosquito (*Aedes aegypti*) and
1105 *Drosophila simulans*. *Appl Environ Microbiol* 78:4740–4743.

1106 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association
1107 mapping and population genetical parameter estimation from sequencing data.
1108 *Bioinformatics* 27:2987–2993.

1109 Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long
1110 sequences. *Bioinformatics* 32:2103–2110.

1111 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
1112 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–
1113 2079.

1114 de Little SC, Edwards O, van Rooyen AR, Weeks A, Umina PA. 2017. Discovery of
1115 metabolic resistance to neonicotinoids in green peach aphids (*Myzus persicae*) in
1116 Australia. *Pest Manag Sci* 73:1611–1617.

1117 de Little SC, Umina PA. 2017. Susceptibility of Australian *Myzus persicae* (Hemiptera:
1118 Aphididae) to Three Recently Registered Insecticides: Spirotetramat, Cyantraniliprole,
1119 and Sulfoxaflor. *J Econ Entomol* 110:1764–1769.

1120 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
1121 quantitative PCR and the 2<super>−ΔΔCT</super>method. *Methods* 25:402–408.

1122 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion
1123 for RNA-seq data with DESeq2. *Genome Biol* 15:1–21.

1124 Loxdale HD. 2008. The nature and reality of the aphid clone: Genetic variation,
1125 adaptation and evolution. *Agric For Entomol* 10:81–90.

1126 Loxdale HD, Massonnet B, Weisser WW. 2010. Why are there so few aphid clones? *Bull*
1127 *Entomol Res* 100:613–622.

1128 Lyberger K, Schoener TW. 2023. Differential genotype response to increased resource
1129 abundance helps explain parallel evolution of *Daphnia* populations in the wild. *Ecol*
1130 *Evol* 13:1–9.

1131 Magalhaes IS, Whiting JR, D'Agostino D, Hohenlohe PA, Mahmud M, Bell MA, Skúlason S,
1132 MacColl ADC. 2020. Intercontinental genomic parallelism in multiple three-spined
1133 stickleback adaptive radiations. *Nat Ecol Evol*.

1134 Mahieu L, González-González A, Rubio-Meléndez ME, Moya-Hernández M, Francis F,
1135 Ramírez CC. 2024. An Aphid Pest Superclone Benefits From a Facultative Bacterial
1136 Endosymbiont in a Host-Dependent Manner, Leading to Reproductive and
1137 Proteomic Changes. *Arch Insect Biochem Physiol* 117:e22154.

1138 Maino JL, Umina PA, Hoffmann AA. 2018. Climate contributes to the evolution of
1139 pesticide resistance. *Global Ecology and Biogeography* 27:223–232.

1140 Montejo-Kovacevich G, Meier JI, Bacquet CN, Warren IA, Chan YF, Kucka M, Salazar C,
1141 Rueda-M N, Montgomery SH, McMillan WO, et al. 2022. Repeated genetic
1142 adaptation to altitude in two tropical butterflies. *Nat Commun* 13:1–16.

1143 Morales-Hojas R, Chen J, Sun J, Alvira F, Xiaoling I. 2020. Contrasting population
1144 structure and demographic history of cereal aphids in different environmental and
1145 agricultural landscapes. *Ecol Evol* 10:9647–9662.

1146 Muller HJ. 1932. Some genetic aspects of sex. *Am Nat* 66:118–138.

1147 Müller V, Buer B, Lueke B, Mazzoni E, Pym A, Bass C, Nauen R. 2023. Molecular
1148 characterization of pyrethroid resistance in field-collected populations of the pea
1149 aphid, *Acyrthosiphon pisum*. *Entomol. Gen* 43:627–637.

1150 Natwick E, Brlow V, Godfrey L, Goddell P, Long R. 2014. Blue alfalfa aphid: Old pest with
1151 new management challenges. In: California Alfalfa, Forage, and Grain Symposium.
1152 Long Beach, California, USA.

1153 Nguyen Ba AN, Cvijović I, Rojas Echenique JI, Lawrence KR, Rego-Costa A, Liu X, Levy SF,
1154 Desai MM. 2019. High-resolution lineage tracking reveals travelling wave of
1155 adaptation in laboratory yeast. *Nature* 575:494–499.

1156 Ono M, Swanson JJ, Field LM, Devonshire AL, Siegfried BD. 1999. Amplification and
1157 methylation of an esterase gene associated with insecticide-resistance in greenbugs,
1158 *Schizaphis graminum*(Rondani) (Homoptera: Aphididae). Available from:
1159 www.elsevier.com/locate/ibmb

1160 Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R
1161 language. *Bioinformatics* 20:289–290.

1162 Pertea G, Pertea M. 2020. GFF utilities: GffRead and GffCompare. *F1000Res* 9.

1163 Rispe C, Pierre J -S., Simon J -C., Gouyon P -H. 1998. Models of sexual and asexual
1164 coexistence in aphids based on constraints. *J Evol Biol* 11:685–701.

1165 Rispe C, Pierre J-S. 1998. Coexistence between cyclical parthenogens, obligate
1166 parthenogens, and intermediates in a fluctuating environment. *J Theor Biol* 195:97–
1167 110.

1168 Roach MJ, Schmidt S, Borneman AR. 2018. Purge Haplotigs: Synteny Reduction for Third-
1169 gen Diploid Genome Assemblies. *BMC Bioinformatics* 19:1–10.

1170 Roda F, Liu H, Wilkinson MJ, Walter GM, James ME, Bernal DM, Melo MC, Lowe A,
1171 Rieseberg LH, Prentis P, et al. 2013. Convergence and divergence during the
1172 adaptation to similar environments by an Australian groundsel. *Evolution (N Y)*
1173 67:2515–2529.

1174 Ryalls JMW, Riegler M, Moore BD, Johnson SN. 2013. Biology and trophic interactions of
1175 lucerne aphids. *Agric For Entomol* 15:335–350.

1176 Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27:592–593.

1177 Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q
1178 file manipulation. *PLoS One* 11:e0163962.

1179 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO:
1180 Assessing genome assembly and annotation completeness with single-copy
1181 orthologs. *Bioinformatics* 31:3210–3212.

1182 Singh KS, Cordeiro EMG, Troczka BJ, Pym A, Mackisack J, Mathers TC, Duarte A, Legeai F,
1183 Robin S, Bielza P, et al. 2021. Global patterns in genomic diversity underpinning the
1184 evolution of insecticide resistance in the aphid crop pest *Myzus persicae*. *Commun*
1185 *Biol*:1–16.

1186 Sprengelmeyer QD, Pool JE. 2021. Ethanol resistance in Drosophila melanogaster has
1187 increased in parallel cold-adapted populations and shows a variable genetic
1188 architecture within and between populations. *Ecol Evol* 11:15364–15376.

1189 Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab
1190 initio prediction of alternative transcripts. *Nucleic Acids Res* 34:W435–W439.

1191 Steiner CF, Stockwell RD, Tadros M, Shaman L, Patel K, Khraizat L. 2016. Impacts of
1192 dispersal on rapid adaptation and dynamic stability of *Daphnia* in fluctuating
1193 environments. *Proceedings of the Royal Society B: Biological Sciences* 283.

1194 Tepolt CK, Grosholz E, de Rivera C, Ruiz G. 2021. Balanced polymorphism fuels rapid
1195 selection in an invasive crab despite high gene flow and low genetic diversity. *Mol*
1196 *Ecol*:1–15.

1197 Thia JA. 2025. Dataset: Spread of a single superclone drives insecticide resistance in
1198 Acyrthosiphon kondoi across an invasive range. *FigShare*.

1199 Thia JA, Korhonen PK, Young ND, Gasser RB, Umina PA, Yang Q, Edwards O, Tom W,
1200 Hoffmann AA. 2023. The redlegged earth mite draft genome provides new insights
1201 into pesticide resistance evolution and demography in its invasive Australian range.
1202 *J Evol Biol* 36:381–398.

1203 Thia JA, Riginos C. 2019. *genomalicious*: serving up a smorgasbord of R functions for
1204 population genomic analyses. *bioRxiv*:667337.

1205 Thia JA, Zhan D, Robinson K, Umina PA, Hoffmann AA, Yang Q. 2023. "Drifting" *Buchnera*
1206 genomes track the microevolutionary trajectories of their aphid hosts. *bioRxiv*.

1207 Umina PA, Edwards O, Carson P, Van Rooyen A, Anderson A. 2014. High levels of
1208 resistance to carbamate and pyrethroid chemicals widespread in Australian *Myzus*
1209 *persicae* (Hemiptera: Aphididae) populations. *J Econ Entomol* 107:1626–1638.

1210 Valmorbida I, Hohenstein JD, Coates BS, Bevilaqua JG, Menger J, Hodgson EW, Koch RL,
1211 O'Neal ME. 2022. Association of voltage-gated sodium channel mutations with
1212 field-evolved pyrethroid resistant phenotypes in soybean aphid and genetic markers
1213 for their detection. *Sci Rep* 12.

1214 Vorburger C. 2006. Temporal dynamics of genotypic diversity reveal strong clonal
1215 selection in the aphid *Myzus persicae*. *J Evol Biol* 19:97–107.

1216 Vurture GW, Sedlazeck FJ, Nattestad M, Underwood CJ, Fang H, Gurtowski J, Schatz MC.
1217 2017. GenomeScope: fast reference-free genome profiling from short reads.
1218 *Bioinformatics* 33:2202–2204.

1219 Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q,
1220 Wortman J, Young SK. 2014. Pilon: an integrated tool for comprehensive microbial
1221 variant detection and genome assembly improvement. *PLoS One* 9:e112963.

1222 Walsh TK, Joussen N, Tian K, Mcgaughran A, Anderson CJ, Qiu X, Ahn S-J, Bird L, Pavlidi
1223 N, Vontas J, et al. 2018. Multiple recombination events between two cytochrome
1224 P450 loci contribute to global pyrethroid resistance in *Helicoverpa armigera*. *PLoS*
1225 *One* 13:e0197760.

1226 Wang L-G, Lam TT-Y, Xu S, Dai Z, Zhou L, Feng T, Guo P, Dunn CW, Jones BR, Bradley T.
1227 2020. Treeio: an R package for phylogenetic tree input and output with richly
1228 annotated and associated data. *Mol Biol Evol* 37:599–603.

1229 Ward S, Jalali T, van Rooyen A, Reidy-Crofts J, Moore K, Edwards O, Umina PA. 2024. The
1230 evolving story of sulfoxaflor resistance in the green peach aphid, Myzus persicae
1231 (Sulzer). *Pest Manag Sci* 80:866–873.

1232 Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes
1233 A, Henry L, Hester J. 2019. Welcome to the Tidyverse. *J Open Source Softw* 4:1686.

1234 Williams I, Dixon A. 2017. Life Cycles and Polymorphism. In: Van Emden HF, Harrington R,
1235 editors. Aphids as Crop Pests. Wallingford, Oxfordshire, UK, and Cambridge,
1236 Massachusetts, USA: CABI Publishing. p. 69–81.

1237 Winter D. 2020. pafr: Read, Manipulate and Visualize "Pairwise Mapping Format" Data
1238 (v0.0.2). Available from: https://github.com/dwinter/pafr

1239 De Wit P, Pespeni MH, Palumbi SR. 2015. SNP genotyping and population genomics
1240 from expressed sequences - Current advances and future possibilities. *Mol Ecol*
1241 24:2310–2323.

1242 Yang Q, Gill A, Robinson KL, Umina PA, Ross PA, Zhan D, Brown C, Bell N, MacMahon A,
1243 Hoffmann AA. 2023. A diversity of endosymbionts across Australian aphids and their
1244 persistence in aphid cultures. *Environ Microbiol*:1988–2001.

1245 Yang Q, Umina PA, Rašić G, Bell N, Fang J, Lord A, Hoffmann AA. 2020. Origin of
1246 resistance to pyrethroids in the redlegged earth mite (*Halotydeus destructor*) in
1247 Australia: repeated local evolution and migration. *Pest Manag Sci*.

1248 Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. 2018. Quantifying how constraints
1249 limit the diversity of viable routes to adaptation. *PLoS Genet* 14.

1250 Yu G, Smith DK, Zhu H, Guan Y, Lam TT. 2017. ggtree: an R package for visualization and
1251 annotation of phylogenetic trees with their covariates and other associated data.
1252 *Methods Ecol Evol* 8:28–36.

1253 Zdobnov EM, Apweiler R. 2001. InterProScan–an integration platform for the signature-
1254 recognition methods in InterPro. *Bioinformatics* 17:847–848.

1255 Zhang L, Li Z, Peng Y, Liang X, Wilson K, Chipabika G, Karangwa P, Uzayisenga B, Mensah
1256 BA, Kachigamba DL, et al. 2023. Global genomic signature reveals the evolution of
1257 fall armyworm in the Eastern hemisphere. *Mol Ecol* 32:5463–5478.

1258 Zhang LJ, Jing YP, Li XH, Li CW, Bourguet D, Wu G. 2015. Temperature-sensitive fitness
1259 cost of insecticide resistance in Chinese populations of the diamondback moth
1260 *Plutella xylostella*. *Mol Ecol* 24:1611–1627.

1261 Zhu A, Ibrahim JG, Love MI. 2019. Heavy-tailed prior distributions for sequence count
1262 data: removing the noise and preserving large differences. *Bioinformatics* 35:2084–
1263 2092.

1264 Zuo Y, Peng X, Wang K, Lin F, Li Y, Chen M. 2016. Expression patterns, mutation
1265 detection and RNA interference of Rhopalosiphum padi voltage-gated sodium
1266 channel genes. *Sci Rep* 6.

1267

1268

1269

Figures and tables

Table 1. Clonal *Acyrthosiphon kondoi* strains examined in this study.

Strain	State	Latitude	Longitude	Year collected	Host plant	Resistant	Source
Albany	WA	–34.476	118.48	2023	Clover	No	Field strain, single individual
Avalon	VIC	–38.018	144.491	2024	Lucerne	No	Field strain, single individual
Eudunda	SA	–34.1413	139.0024	2022	Vetch	No	Selected isofemale strain
Jung	VIC	–36.5524	142.3643	2022	Lentils	No	Selected isofemale strain
Kaniva	VIC	–36.225	141.135	2023	Lentils	No	Field strain, single individual
Kellerberrin *	WA	–31.6336	117.7232	1999	Lucerne	No	Selected isofemale strain
Moulamein *	NSW	–35.084	144.0406	2022	Mixed pasture	No	Selected isofemale strain
Bangham	SA	–36.489	140.8792	2021	Subclover	Yes	Selected isofemale strain
Canowindra	NSW	–33.58	148.628	2020	Lucerne	Yes	Field strain, single individual
Coolah	NSW	–31.823	149.72	2024	Lucerne	Yes	Field strain, single individual
Coombe	SA	–35.9536	140.3239	2022	Lucerne	Yes	Field strain, single individual
Cummins	SA	–34.464	135.462	2022	Lucerne	Yes	Field strain, single individual
Keith	SA	–36.0983	140.3739	2020	Lucerne	Yes	Selected isofemale strain
Kerang	VIC	–35.963	143.656	2023	Lentils	Yes	Field strain, single individual
Manoora	SA	–34.015	138.883	2023	Lucerne	Yes	Field strain, single individual
Temora	NSW	–34.5612	147.6033	2020	Lucerne	Yes	Selected isofemale strain

* Note, location data for the Kellerberrin and Moulamein strain are approximates only based on the nearest township.

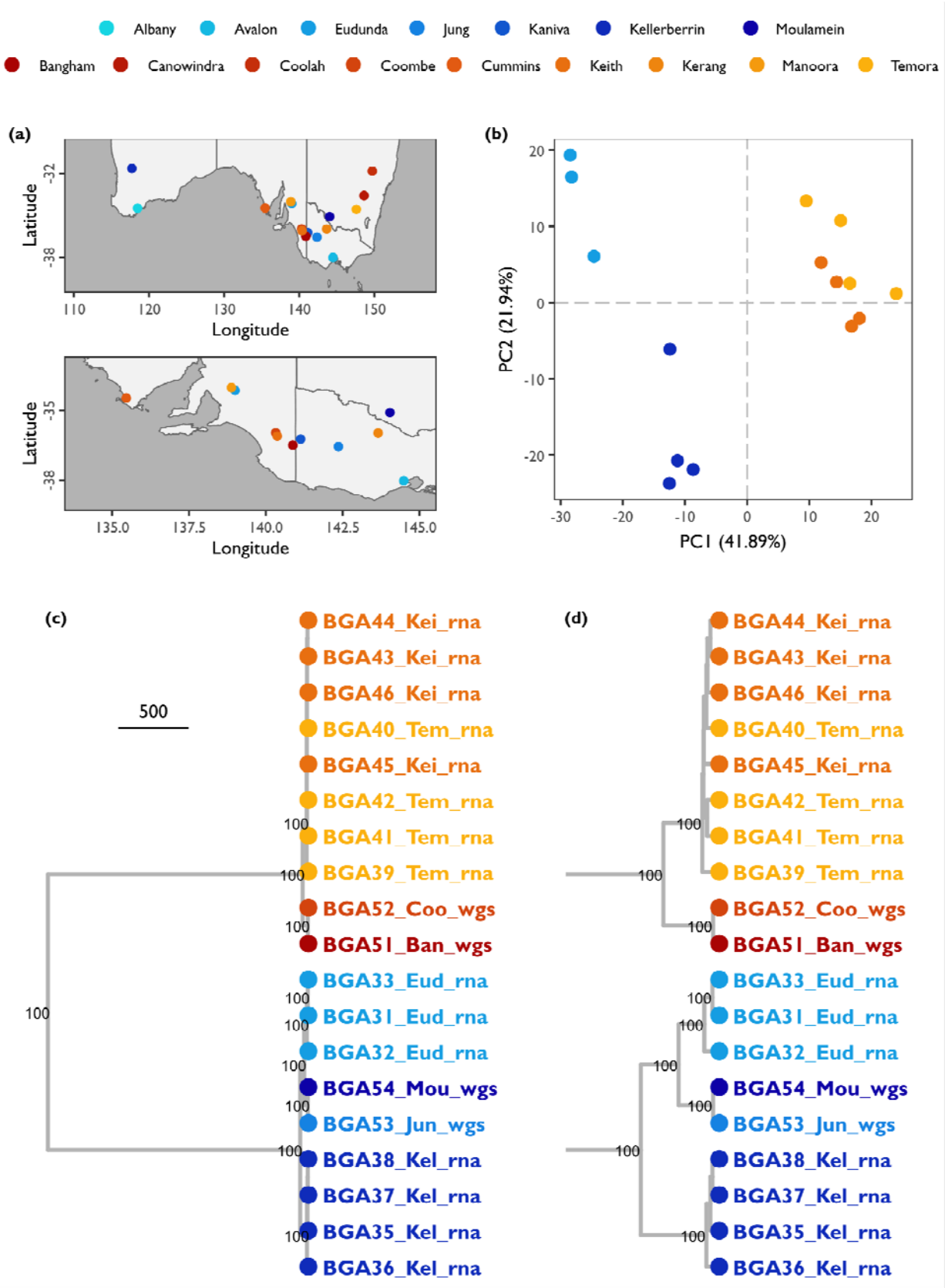

Figure 1. Transcriptomic and genetic comparisons of *Acyrthosiphon kondoi* strains examined in this study. The susceptible strains include Eudunda (Eud), Kellerberrin (Kel), Jung (Jun) and

Moulamein (Mou). The resistance strains include Bangham (Ban), Coombe (Coo), Keith (Kei) and Temora (Tem). (a) Source locations of each strain in Australia. The top panel provides a broader geographic view of eastern and western regions, whereas the bottom panel provides a finer geographic view of southern regions. (b) Transcriptomic profiles of the Eudunda, Kellerberrin, Keith and Temora strains projected into principal component space. Percentages in parentheses on the x-axis and y-axis capture the amount of variance explained by the first and second PC axis, respectively. (c) A UPGMA tree constructed from allelic distances among samples of each strain. Branch nodes are labelled with bootstrap values >80%. Samples follow the naming convention 'ID_Strain_Method', where 'ID' is a unique sample identifier, 'Strain' is one of the eight sampled strains, and 'Method' is one of 'rna' for transcriptome sequencing or 'wgs' for whole-genome sequencing. (a–c) Points are coloured by strain, see the legend.

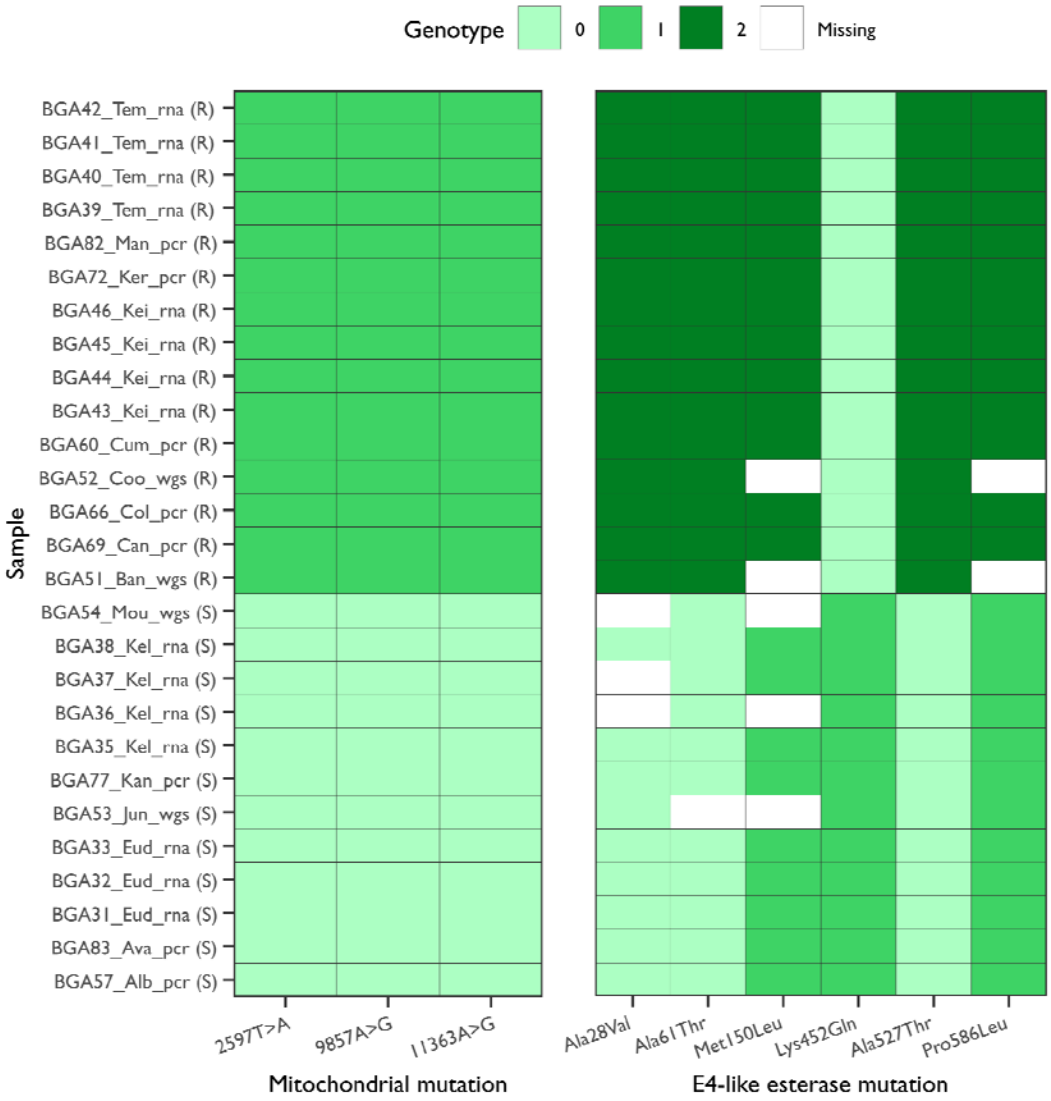

Figure 2. Diagnostic mutations in mitochondrial genes and the candidate E4-like esterase gene. Mutant (non-reference) alleles are arranged along the *x*-axis. Samples are arranged on the *y*-axis, labelled as 'ID_Strain_Method (Status)': 'ID' is a unique sample identifier; 'Strain' is strains name; 'Method' is one of 'rna' for transcriptome sequencing, 'wgs' for whole-genome sequencing, and 'pcr' for Sanger sequenced PCR amplicons; and 'Status' is one of 'S' for susceptible, or 'R' for resistant. Cell colours indicate the observed genotype (see legend).

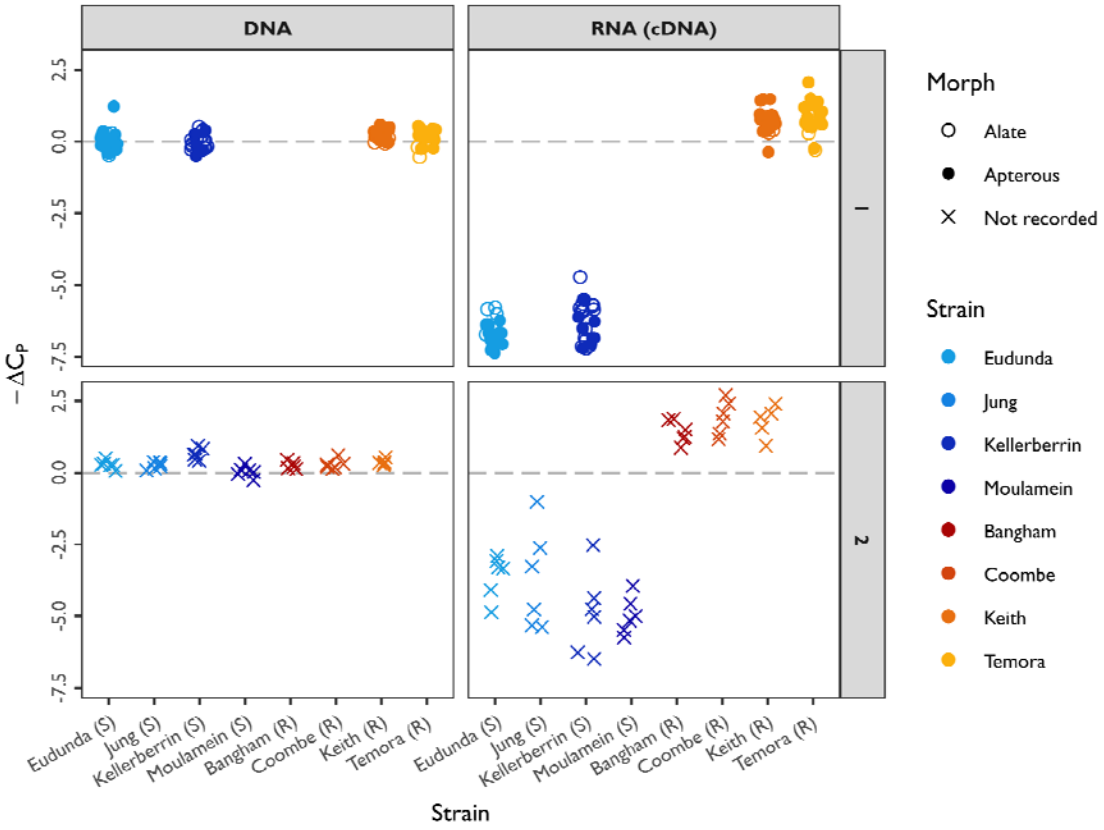

Figure 3. Among-strain comparisons of candidate E4-like esterase relative genomic copy number and expression in 12-day-old *Acyrthosiphon kondoi* measured with qPCR. Clonal strains are arranged on the *x*-axis. The $-\Delta C_P$ is on the *y*-axis, which represents the inverse of the ΔC_P, where the ΔC_P = E4-like esterase C_P – actin C_P. A positive $-\Delta C_P$ therefore indicates a higher amplification of the E4-like esterase (more molecules), and a negative $-\Delta C_P$ indicates a lower amplification of the E4-like esterase (fewer molecules), relative to the reference actin gene. Panels contain results from different experiments (rows: Experiment 1 and 2) and molecule types (columns: DNA and RNA/cDNA). The dashed grey line demarcates a $-\Delta C_P$ of 0, that is, no difference in the amplification of the E4-like esterase and actin genes. Points are coloured by strain, with shapes indicating the aphid morph (see legend). In Experiment 1: points represent individual-level measurements for the susceptible strains Eudunda and Kellerberrin, and the resistant strains Keith and Temora (morph type is indicated by shape, see legend). In Experiment

2: points represent pooled measurements (two aphids per sample) for the susceptible strains Eudunda, Jung, Kellerberrin, and Moulamein, and the resistant strains Bangham, Coombe, and Keith (morph type not recorded).

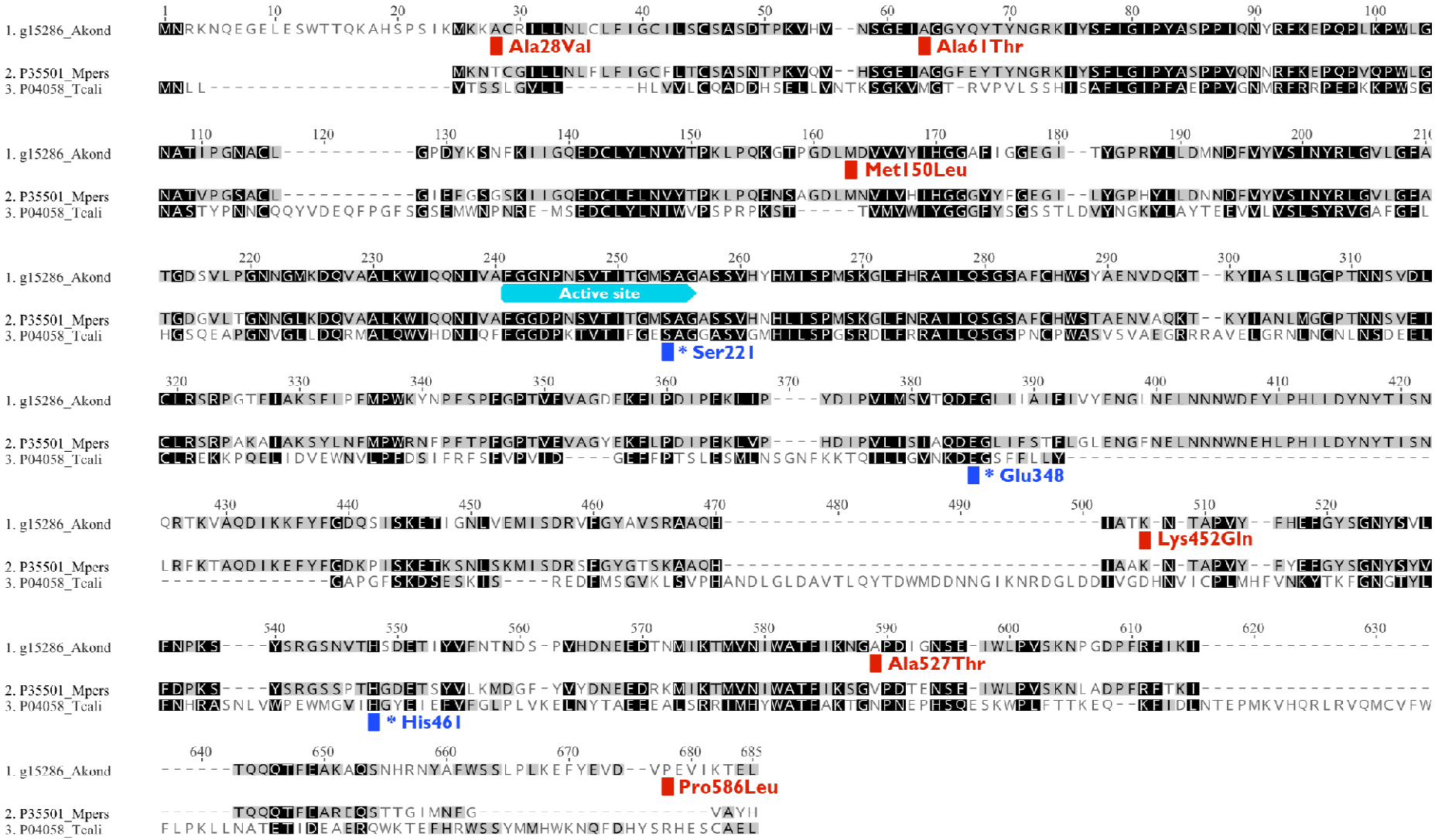
1. g15286_Akond
2. P35501_Mpers
3. P04058_Tcali
Ala28Val
Ala61Thr
Met150Leu
Active site
* Ser221
* Glu348
Lys452Gln
Ala527Thr
* His461
Pro586Leu

Figure 4. Protein alignment of the *Acrythosiphon kondoi* E4-like esterase to other type B carboxylesterases. The *A. kondoi* E4-like esterase (1) is compared against its homolog *Myzus persicae* E4 esterase (2), as well as the *Torpedo californica* acetylcholinesterase (3) as a classic example of type B carboxylesterase. Amino acids are shaded by their level of conservatism (darker shades are more conserved). Red annotations in the *A. kondoi* E4-like esterase sequence indicate positions where non-synonymous mutations are found (*A. kondoi* numbering). Cyan annotations indicate positions of the active site in the *A. kondoi* E4-like esterase sequence. Blue annotations mark the amino acids forming the catalytic triad in the *T. californica* acetylcholinesterase sequence (indicated with an asterisk, *T. californica* numbering). (See also Figure 5)

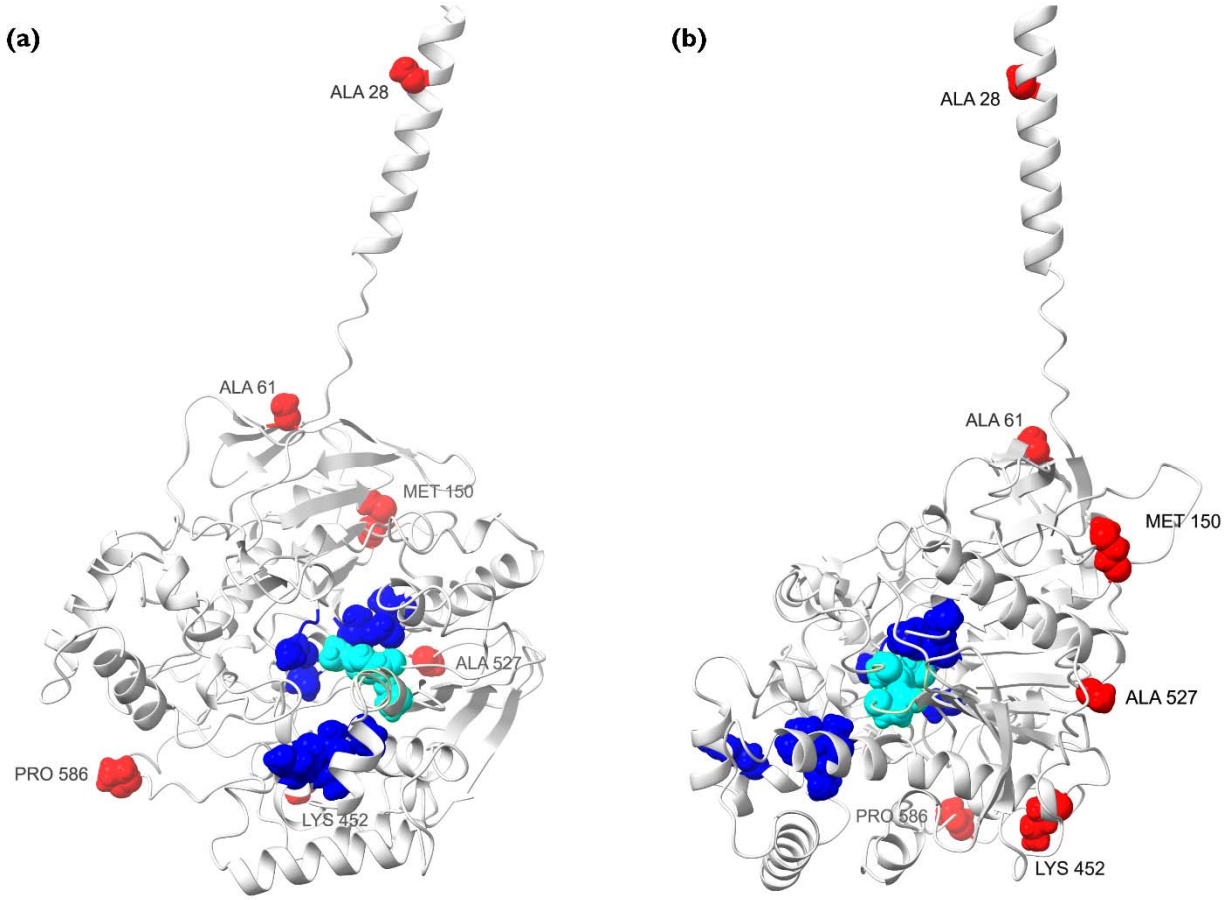

Figure 5. Structural modelling of the *A. kondoi* E4-like esterase protein. (a) Projection of the protein structure with the active site facing forward. The catalytic triad is highlighted in cyan, while the amino acids forming part of the substrate pocket are highlighted in blue. The positions of six non-synonymous mutations are highlighted in red and are labelled by their reference amino acid. (b) The same structure viewed from the side, showing that the positions of the non-synonymous mutations are located close to the surface on the opposite side of the active site. The exception is the Ale28 position, which is located on the signal peptide. (See also Figure 4)

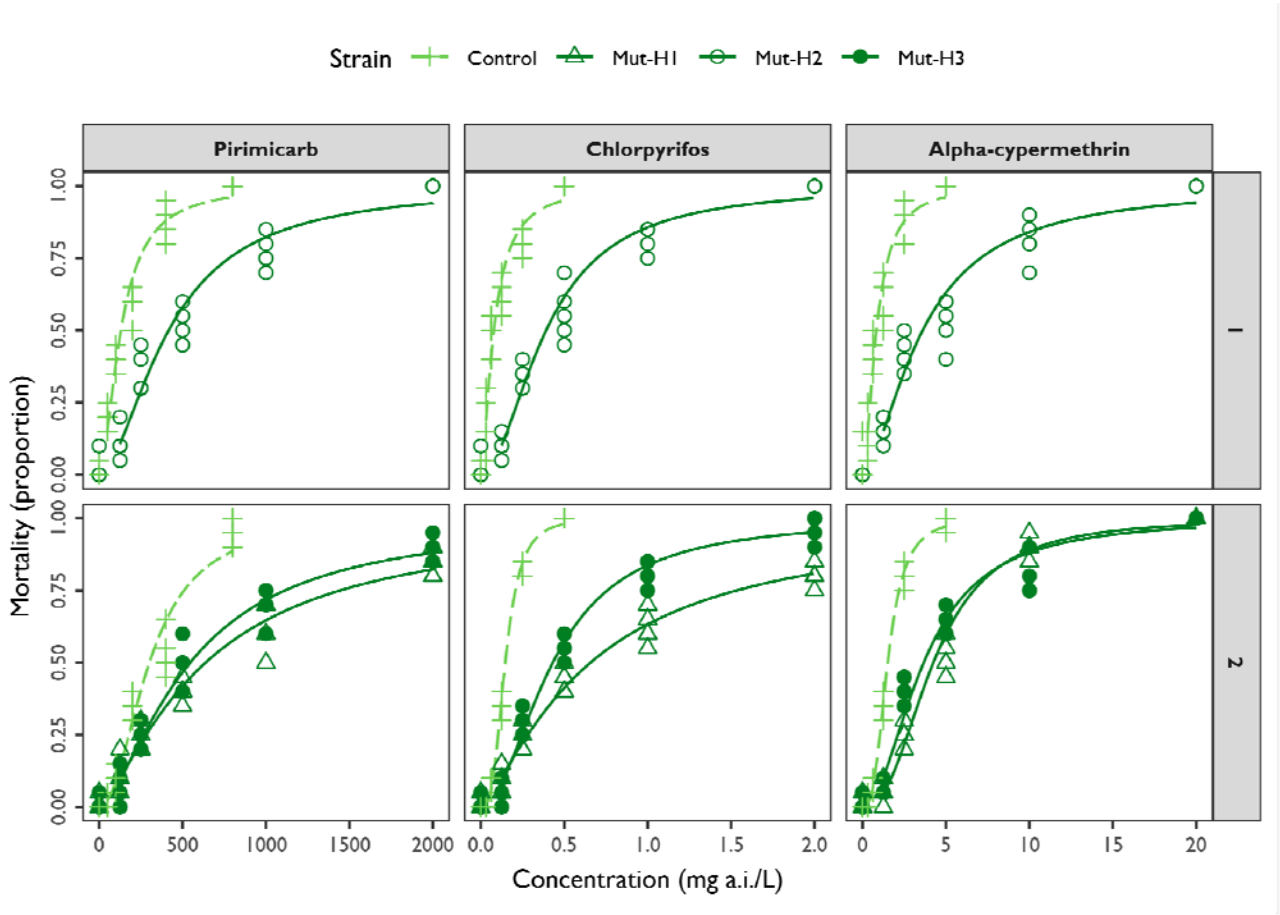

Figure 6. Dose-responses generated from insecticide bioassays on transgenic *Drosophila melanogaster* carrying our candidate E4-like esterase from *Acythosiphon kondoi*. The *x*-axis represents the concentration of insecticide, and the *y*-axis represents the proportion of mortality. Points represent individual replicates with predicted dose response curves (line type and colour depicts strains, see legend). Panels contain data for different experiments (rows) and insecticides tested (columns) (see also Table 2).

Supplementary information

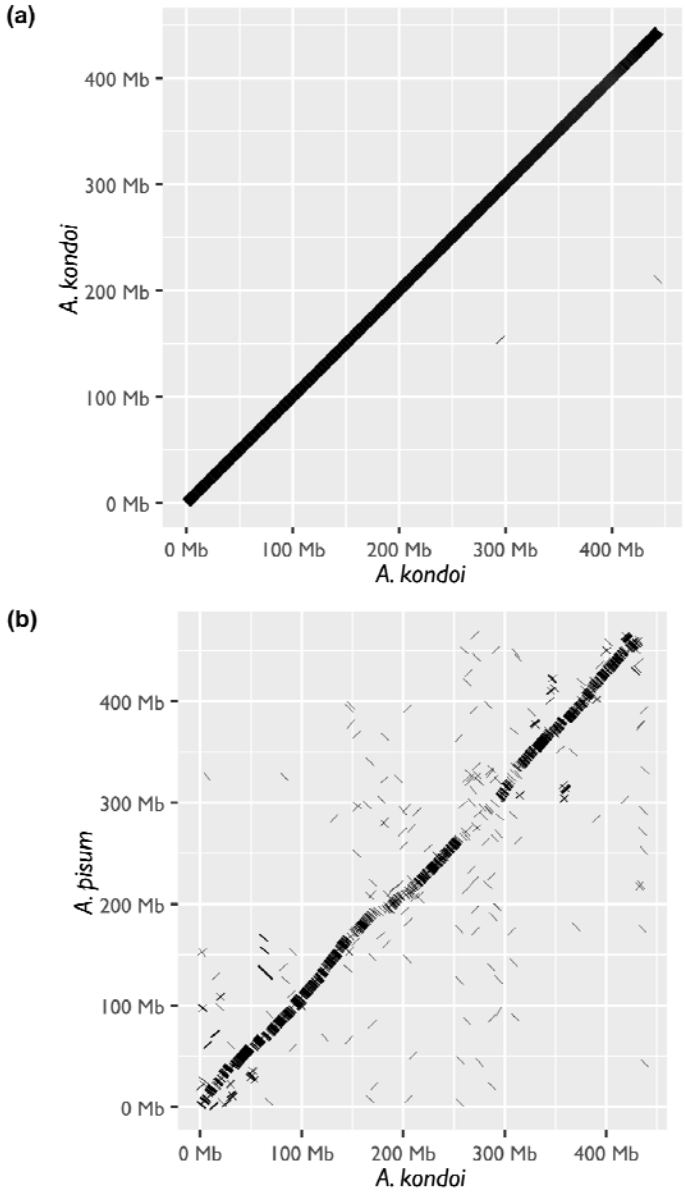

Figure S1. Dotplots of genomic alignments. The *x*-axis and *y*-axis indicate the concatenated genomic position (sequences lined up end-to-end). (a) Self-to-self alignment of our scaffold-level assembly of *Acythorsiphon kondoi*. The linear dotplot suggests that all sequences included in our final assembly are highly unique. (b) Alignment of *A. kondoi* pseudo-chromosomes to the *A. pisum* chromosome-level assembly (NCBI GCF_005508785.2). *Acythosiphon kondoi* pseudo-chromosomes were generated with *RagTag* in an initial

alignment against the *A. pisum* genome. Derived pseudo-chromosomes were then mapped back to the *A. pisum* genome. (See also Figure S2)

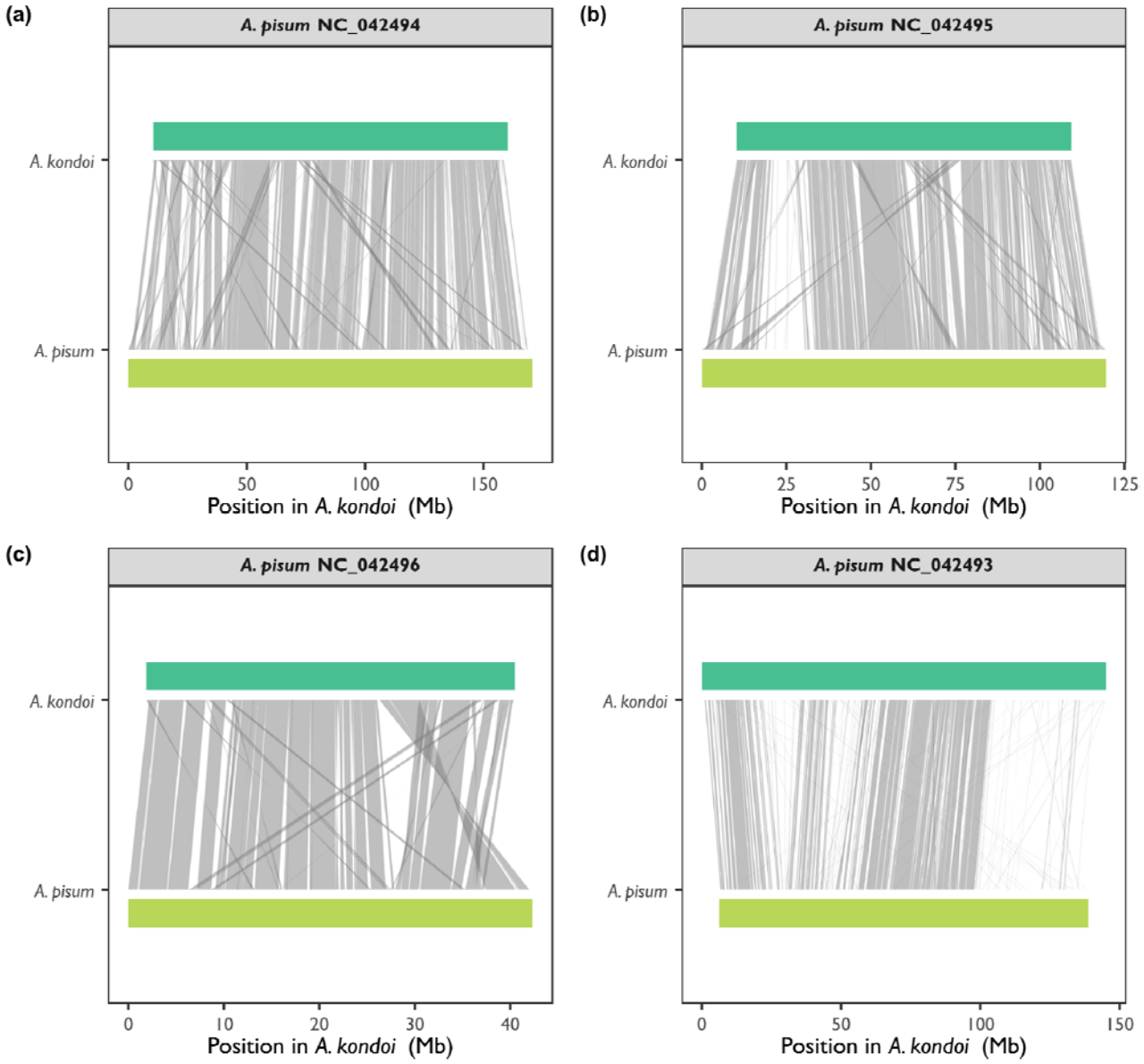

Figure S2. Realignment of *Acyrthoisphon kondoi* pseudo-chromosomes against the chromosome-level assembly of the *A. pisum* genome (NCBI GCF_005508785.2). The *x*-axis represents the position on the *A. kondoi* pseudo-chromosome. The *y*-axis delineates *A. kondoi* and *A. pisum* sequences. Grey bars illustrate alignments > 50,000 bp between the two sequences. Panels separate the three *A. pisum* chromosomes: (a) NC_042494, the A1 chromosome; (b) NC_042495, the A2 chromosome; (c) NC_042496, the A3 chromosome; and (d) NC_042493, the X chromosome. (See also Figure S1)

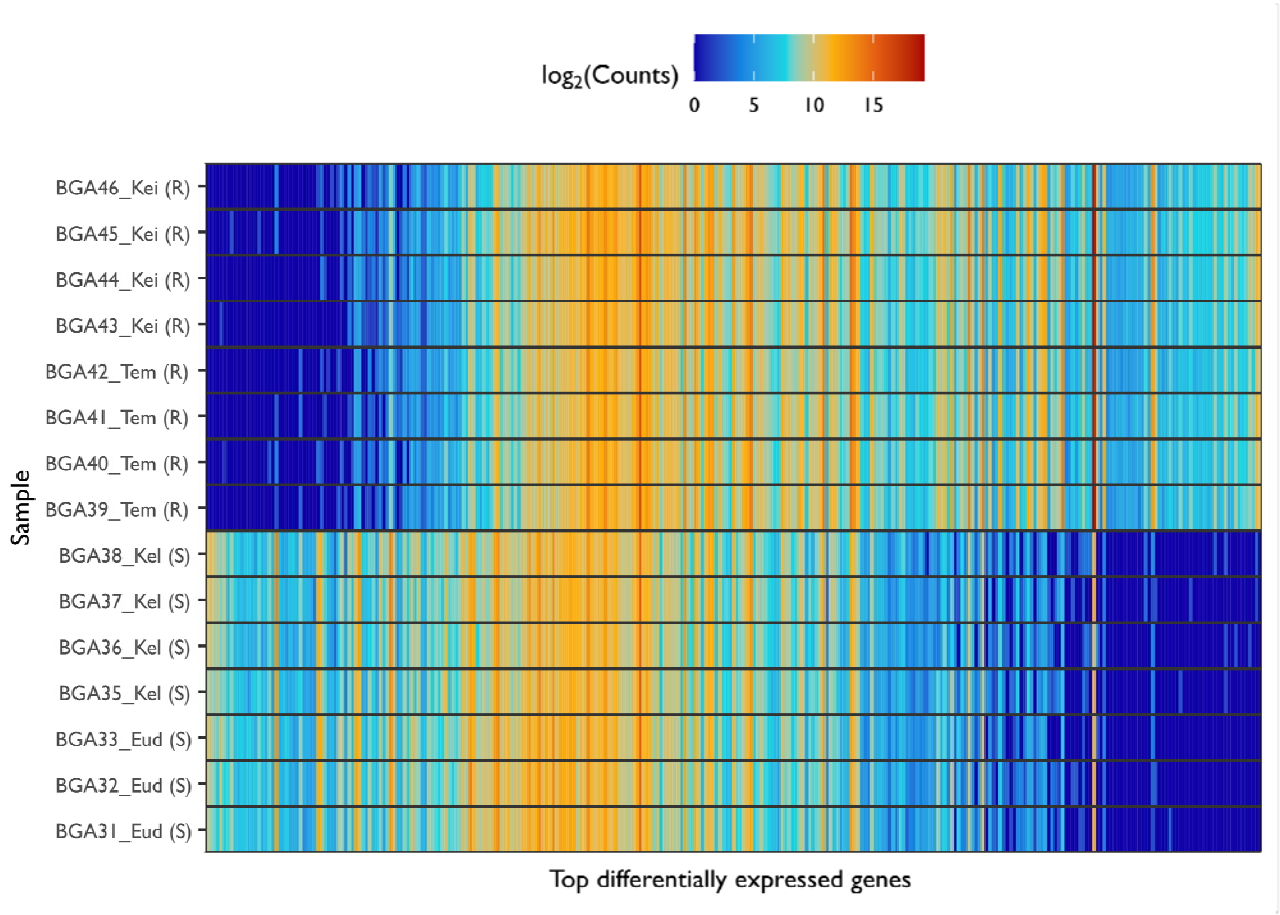

Figure S3. Read counts for the top differentially expressed genes (under- and over-expressed). Genes are on the *x*-axis, arranged from most under-expressed to most over-expressed in the resistant strains, relative to the susceptible strains. The *y*-axis represents different samples, labelled as 'ID_Strain (Status)': 'ID' is a unique sample identifier; 'Strain' is strains name; and 'Status' is one of 'S' for susceptible, or 'R' for resistant. Cell colours indicate the read counts (see legend).

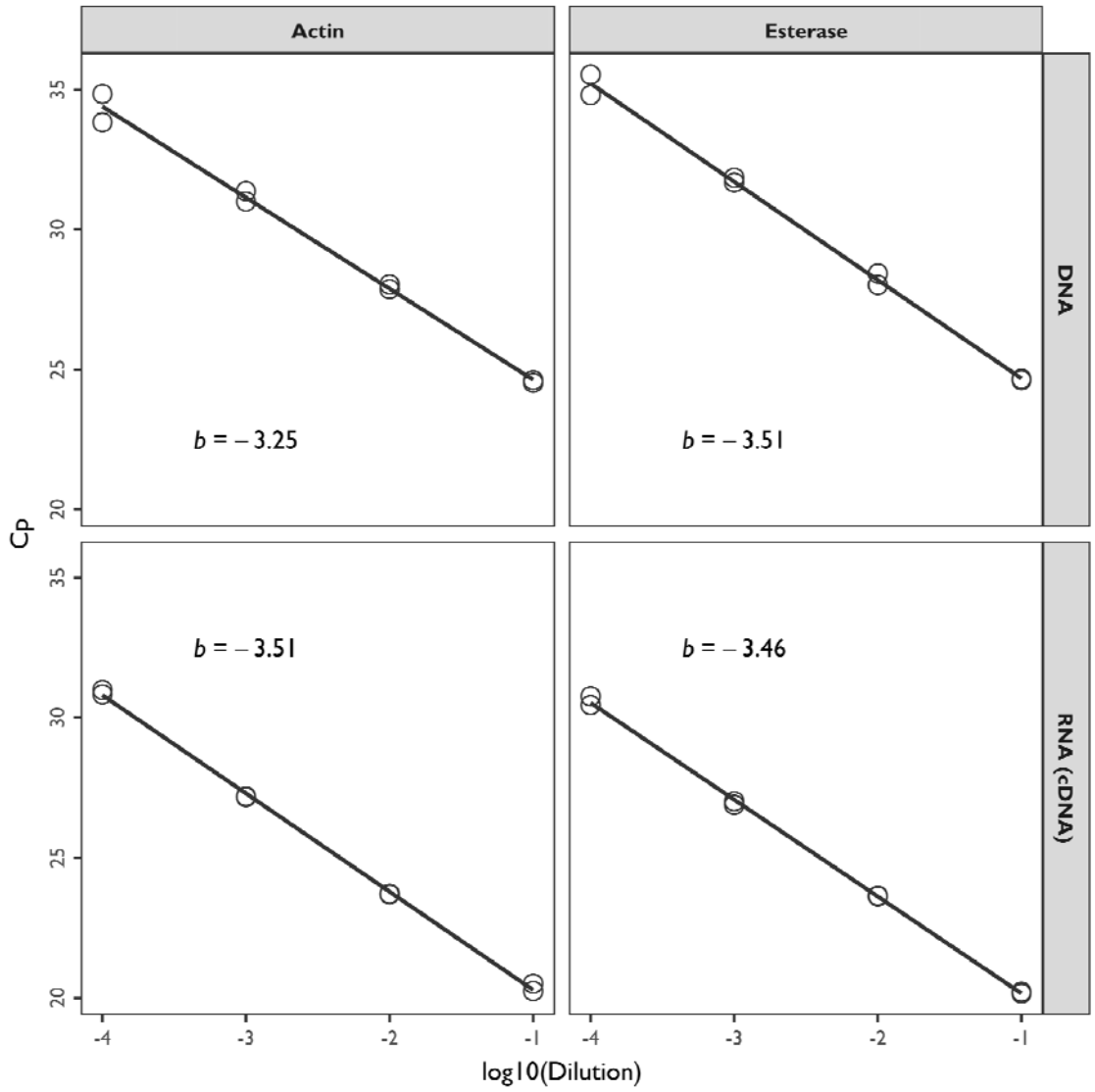

Figure S4. Primer efficiency test for estimating relative expression of the E4-like esterase. The *x*-axis represents the log10-transformed dilution factor of nucleic acid, whereas the *y*-axis represents the mean Cp value obtained in qPCR. Points represent different samples. Black lines indicate the line of best fit. The slope of the line, *b*, is also labelled on the plots: values between −3.3 and −3.6 are good. Panels separate results for the nucleic acid origin (in rows: DNA or RNA) and gene (in columns: actin or the esterase). These efficiency tests were performed on extractions prepared from the Kellerberrin strain.

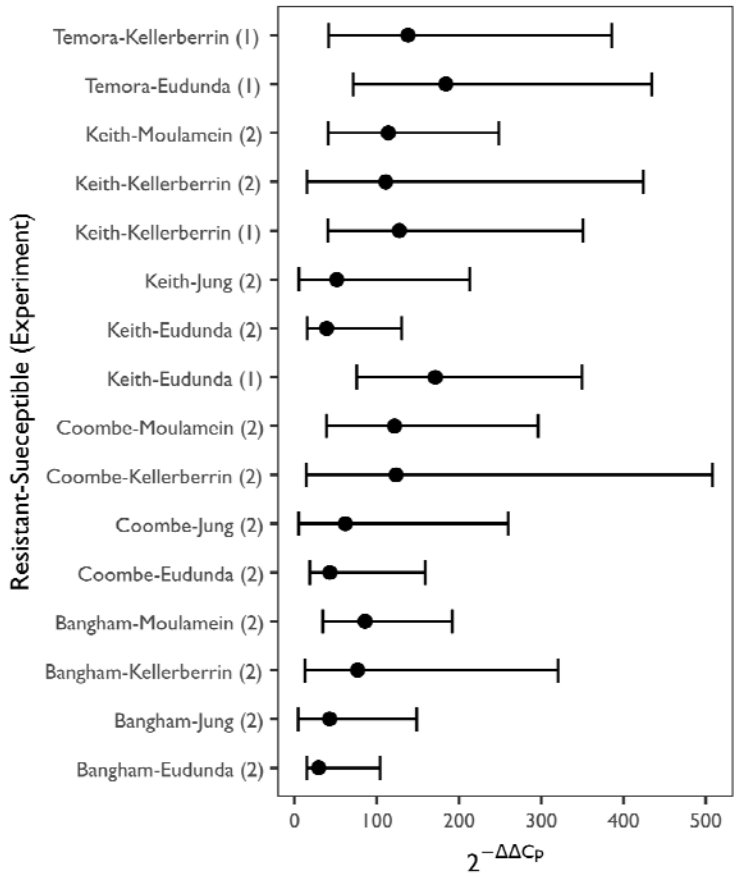

Figure S5. E4-like esterase fold-change expression differences between resistant vs. susceptible strains of *Acyrthosiphon kondoi*. The *x*-axis represents the measured fold-change difference, and the *y*-axis represents comparisons between different strains. Labels on the *y*-axis are presented as resistant vs susceptible with experimental number in parentheses. Points indicate median estimates of fold-change expression differences, with error bars indicating the 2.5% and 97.5% confidence intervals.

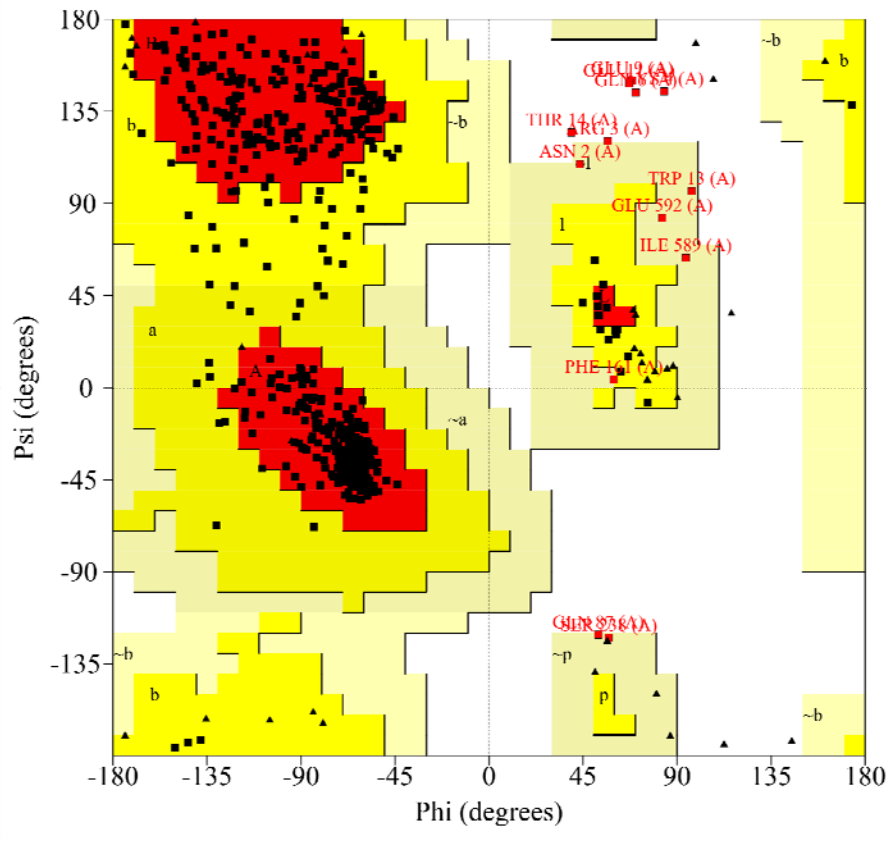

Figure S6. Ramachandran plot for our structural model of the *Acyrthosiphon kondoi* E4-like esterase. The *x*-axis represents the torsion angle around the N–Cα bond, and the *y*-axis represents the torsion angle around Cα–C bonds Points represent amino acids, with red points indicating conformations in disallowed regions. Red shading indicates favourable conformations, yellow shading indicates allowed conformations, and white shading indicates disallowed conformations.

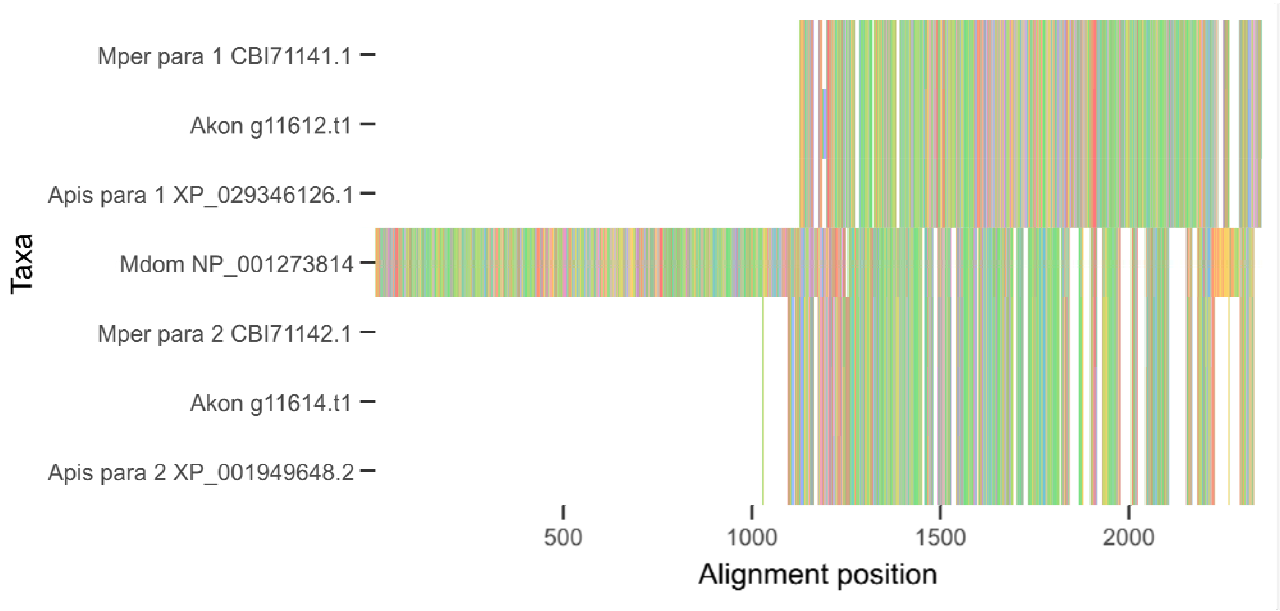

Figure S7. Alignment of *Acrythosiphon kondoi* (Akon) annotated *para* proteins to the canonical reference sequence, *Musca domestica* (Mdom), and to subunit 1 and 2 sequences found in the aphids *Myzus persicae* (Mper) and *A. pisum* (Apis). The alignment position is on the *x*-axis with species arranged on the *y*-axis. Colours represent the different amino acid classes.

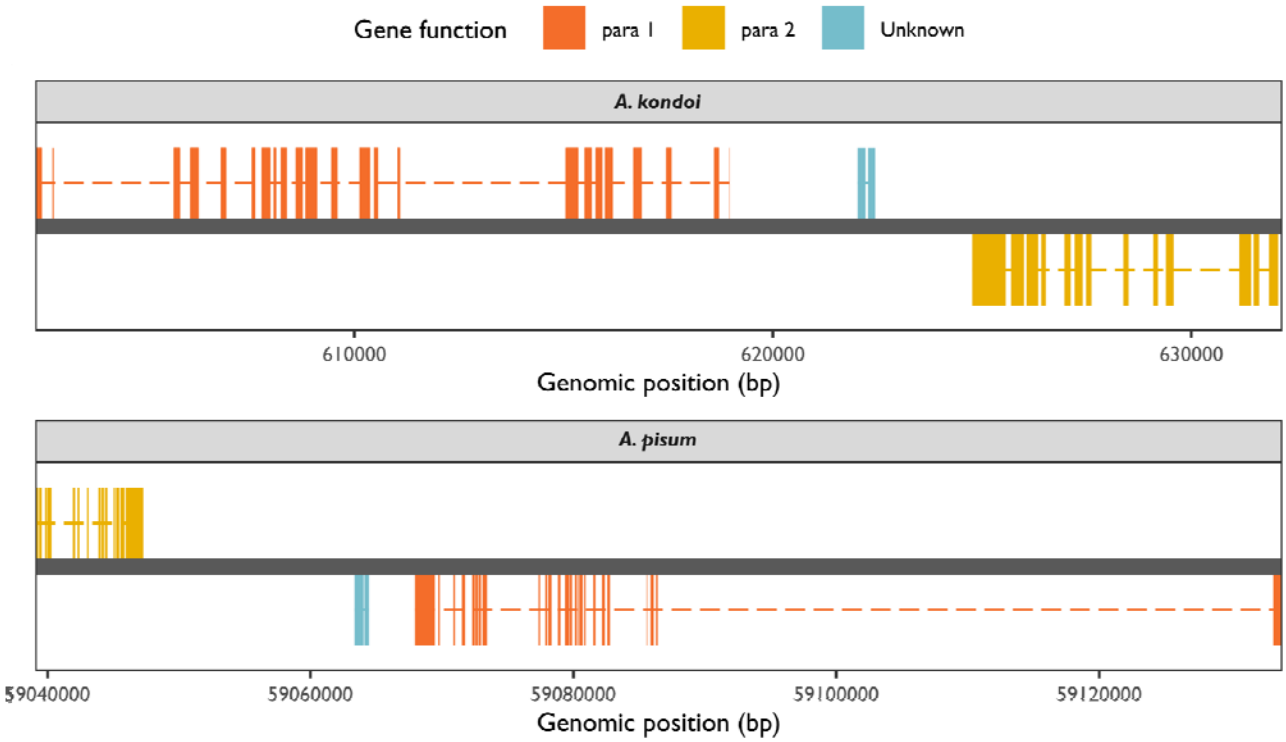

Figure S8. Genomic regions containing *para* genes in our reference assembly of *Acyrthosiphon kondoi* and in the reference assembly of *A. pisum* (NCBI Genome Accession GCF_005508785.1). The *x*-axis represents the genomic position in base pairs. The dark grey horizontal bar represents the genome. Coloured bars represent the exons of different genes, those on top represent genes on the positive strand, whereas those on the bottom represent genes on the negative strand. Dashed lines connect exons within the same gene. Exons are coloured by their gene function (see legend).

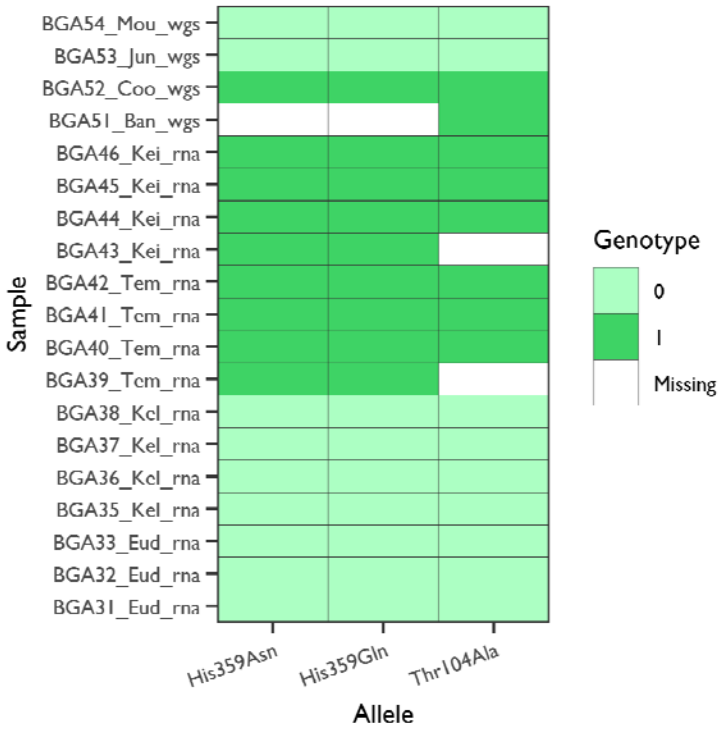

Figure S9. Non-synonymous mutations in the *para* g11614 gene annotated in the genome of *Acyrthosiphon kondoi*. Mutant alleles are arranged along the *x*-axis. Samples are arranged on the *y*-axis, labelled as 'ID_Strain_Method', where 'ID' is a unique sample identifier, 'Strain' is one of the eight clonal strains, and 'Method' is one of 'rna' for transcriptome sequencing or 'wgs' for whole-genome sequencing. Cell colours indicate the observed genotype (see legend). The susceptible strains include Eudunda (Eud), Kellerberrin (Kel), Jung (Jun) and Moulamein (Mou). The resistance strains include Bangham (Ban), Coombe (Coo), Keith (Kei) and Temora (Tem).

Table S1. Mortality of the core *Acyrthosiphon kondoi* isofemale clonal strains to a water control and treatment of 15 mg a.i./L of chlorpyrifos in a leaf dip bioassay (72 hours exposure).

Status	Strain	mg/L	Replicates	Aphids	Mean	CI 2.5%	CI 97.5%
Resistant	Bangham	0	6	30	0		
		15	6	30	0.333	0.19	0.516
	Coombe	0	6	30	0		
		15	6	30	0.433	0.271	0.612
	Keith	0	6	30	0		
		15	10	50	0.04	0.01	0.146
	Temora	0	6	30	0		
		15	10	50	0.06	0.019	0.17
Susceptible	Eudunda	0	6	30	0.033	0.005	0.202
		15	10	50	1		
	Jung	0	6	30	0		
		15	6	30	1		
	Kellerberrin	0	6	30	0		
		15	10	50	1		
	Moulamein	0	6	30	0		
		15	6	30	1		

Table S2. SNP heterozygosity for samples generated using transcriptome ('rna') and whole-genome ('wgs') sequencing at biallelic SNPs used to construct genetic distances.

Strain	Sample (R or S)	SNP heterozygosity
Eudunda	BGA31_Eud_rna (S)	0.514
Eudunda	BGA32_Eud_rna (S)	0.515
Eudunda	BGA33_Eud_rna (S)	0.514
Jung	BGA53_Jun_wgs (S)	0.527
Kellerberrin	BGA35_Kel_rna (S)	0.494
Kellerberrin	BGA36_Kel_rna (S)	0.495
Kellerberrin	BGA37_Kel_rna (S)	0.495
Kellerberrin	BGA38_Kel_rna (S)	0.496
Moulamein	BGA54_Mou_wgs (S)	0.527
Bangham	BGA51_Ban_wgs (R)	0.588
Coombe	BGA52_Coo_wgs (R)	0.589
Keith	BGA43_Kei_rna (R)	0.572
Keith	BGA44_Kei_rna (R)	0.572
Keith	BGA45_Kei_rna (R)	0.574
Keith	BGA46_Kei_rna (R)	0.574
Temora	BGA39_Tem_rna (R)	0.573
Temora	BGA40_Tem_rna (R)	0.574
Temora	BGA41_Tem_rna (R)	0.573
Temora	BGA42_Tem_rna (R)	0.573

Table S3. Fold-change difference in the expression of the candidate E4-like esterase between resistant and susceptible strains estimated from qPCR assays with the $2^{-\Delta\Delta C_P}$ method. Median values with 2.5% and 97.5% percentiles for the estimates are reported.

Experiment	Resistant	Susceptible	Median	2.5%	97.5%
1	Keith	Eudunda	171	76	350
	Keith	Kellerberrin	128	41	351
	Temora	Eudunda	184	71	434
	Temora	Kellerberrin	138	42	386
2	Bangham	Eudunda	29	15	104
	Bangham	Jung	43	5	149
	Bangham	Kellerberrin	77	13	321
	Bangham	Moulamein	86	34	192
	Coombe	Eudunda	43	19	159
	Coombe	Jung	62	5	260
	Coombe	Kellerberrin	124	15	508
	Coombe	Moulamein	122	39	296
	Keith	Eudunda	39	16	130
	Keith	Jung	52	5	213
	Keith	Kellerberrin	111	15	424
	Keith	Moulamein	114	41	249

Table S4. Results from insecticide bioassays on transgenic *Drosophila melanogaster* carrying our candidate E4-like esterase from *Acyrthosiphon kondoi*. Reported are the intercepts (b_0), slopes (b_1), χ^2 goodness-of-fit statistics and associated *P*-values, LD_{50} values (mg a.i./L) with 95% confidence intervals, and resistance ratios (RR) of mutant strains. (See also Figure 4)

Experiment	Strain	Insecticide	b_0	b_1	χ^2 (df)	$P(\chi^2)$	LD50 values (95% CIs)	RR
1	Mut-H2	Pirimicarb	–10.61	4.05	0.60 (3)	0.896	416.7 (369.71,469.66)	3.19
		Chlorpyrifos	1.79	4.39	0.41 (3)	0.939	0.39 (0.36,0.43)	4.88
		Alpha-cypermethrin	–2.09	3.76	0.65 (3)	0.884	3.6 (3.16,4.11)	4.29
	Control	Pirimicarb	–8.64	4.09	0.40 (3)	0.94	130.5 (118.01,144.31)	–
		Chlorpyrifos	4.06	3.67	0.51 (3)	0.916	0.08 (0.07,0.09)	–
		Alpha-cypermethrin	0.32	4.15	0.40 (3)	0.94	0.84 (0.74,0.94)	–
2	Mut-H1	Pirimicarb	–8.68	3.09	0.06 (3)	0.997	637.96 (578.62,703.39)	2.12
		Chlorpyrifos	0.54	2.99	0.06 (3)	0.996	0.66 (0.6,0.71)	4.40
		Alpha-cypermethrin	–3.51	5.60	0.30 (3)	0.961	4.24 (3.94,4.58)	2.81
	Mut-H3	Pirimicarb	–9.70	3.54	0.05 (3)	0.997	545.54 (497.4,598.34)	1.81
		Chlorpyrifos	1.64	4.49	0.10 (3)	0.991	0.43 (0.4,0.47)	2.87
		Alpha-cypermethrin	–2.60	4.60	0.66 (3)	0.883	3.67 (3.29,4.1)	2.43
	Control	Pirimicarb	–11.97	4.83	0.53 (3)	0.913	301.35 (270.88,335.26)	–
		Chlorpyrifos	6.12	7.34	0.23 (3)	0.972	0.15 (0.14,0.15)	–
		Alpha-cypermethrin	–1.21	6.78	0.09 (3)	0.993	1.51 (1.42,1.61)	–

1 **Table S5.** Primers for PCR and Macrogen sequencing of focal regions in the E4-like esterase and mitochondrial genes.

Region	Product (bp)	Primer name	Primer sequence
E4-like esterase exon 2	600	Akond_E4_X2_F	TGTCAGGCGGTTGTGAATGA
		Akond_E4_X2_R	GGTTGTGGTTCCTGCGAAAA
E4-like esterase exon 4	578	Akond_E4_X4_F	ATCACTGGGTGGCTGCTTTT
		Akond_E4_X4_R	ACACATTATTTACGCTGCACAT
E4-like esterase exon 7	589	Akond_E4_X7_F	GTCCTGTCGTGGACGGAAAA
		Akond_E4_X7_R	TCGTCGGAATGAGTTACGTCT
E4-like esterase exon 9	597	Akond_E4_X9_F	ACACGAATGATTCCCCTGTTCA
		Akond_E4_X9_R	CGTACTTTTTGCCCACTGAGC
Mitochondrial *cox1*	565	Akond_cox1_F	TGGAACAGGAACAGGATGAACT
		Akond_cox1_R	TGCTGATGTAAAGTAAGCTCGTG
Mitochondrial *cox2*	526	Akond_cox2_F	ATTTATTTGAACAACAACTCCTCCT
		Akond_cox2_R	AGAGATCATTACTTTGCTTTTCAGTCA
Mitochondrial *cytb*	562	Akond_cytb_F	TCCAACACCTGCCAGAATCA
		Akond_cytb_R	TGGGTTATTAGATCCCGTTATGTGT

2

3