Two different adaptive speciation mechanisms operate during adaptation to a novel hot environment

Sheng-Kai Hsu[1,2], Wei-Yun Lai[1,2], Johannes Novak[3], Felix Lehner[1], Ana Marija Jakšić[1,2,4], Elisabetta Versace[5], Christian Schlötterer[1*]

Affiliations:

[1] Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria.

[2] Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Vienna, Austria.

[3] Institute of Animal Nutrition and Functional Plant Compounds, Vetmeduni Vienna, Vienna, Austria.

[4] Current affiliation: École polytechnique fédérale de Lausanne, Lausanne, Switzerland.

[5] Department of Biological and Experimental Psychology, Queen Mary University of London, London, UK.

*Correspondence: christian.schloetterer@vetmeduni.ac.at; Tel.: +43-1-25077-4300.

Keywords: *Drosophila simulans*, experimental evolution, mutation order speciation, ecological speciation

Abstract

Background: Ecological speciation and mutation-order speciation are two different mechanisms of adaptation-driven speciation. Both mechanisms predict different patterns of reproductive isolation for replicate populations adapting to the same environment. With ecological speciation, barriers to gene flow emerge between populations from different environments, but not among replicate populations from the same environment. Mutation-order speciation predicts reproductive isolation among populations adapted to the same environment.

Results: We demonstrate that both speciation processes occurred within about 100 generations when replicate *Drosophila simulans* populations adapted to a novel, hot environment. Gene expression analysis identified the underlying molecular mechanisms. Premating ecological speciation is the byproduct of an altered lipid metabolism, which also changed the cuticular hydrocarbon (CHC) composition in hot-evolved flies. Postmating reproductive isolation supports mutation-order speciation most likely driven by co-evolution of reproduction-associated genes.

Conclusion: Adaptation processes can rapidly induce incipient speciation and different speciation mechanisms affect pre- and postmating reproductive isolation. We propose that the definition of mutation-order speciation should be expanded to account for polygenic processes from standing genetic variation.

Introduction

Most new species are formed by selection [1-3] as already suggested by Darwin [4]. Two different mechanisms by which selection can lead to speciation have been proposed, ecological speciation and mutation order speciation. During ecological speciation reproductive isolation occurs as a consequence of adaptation to different ecological stressors. Importantly, reproductive isolation is not the direct target of selection, but rather occurs because the selected genes either have pleiotropic effects on reproduction or are linked to another loci affecting reproduction [5]. Mutation-order speciation describes the random occurrence of different favorable, but incompatible, mutations in distinct populations adapting to the same environmental stressor [6, 7]. Many examples of ecological speciation are documented [e. g. 8, 9-12] and also mutation order speciation has received some empirical support [13].

Both speciation mechanisms make different predictions for populations evolving independently in the same environments. Ecological speciation predicts that reproductive isolation develops between populations adapted to different environments, but not among different populations in the same environment [2]. With mutation-order speciation, independent populations which are adapted to the same environment become incompatible because they use different sets of incompatible mutations. Mutation-order speciation can result from intra-genomic conflict as well as antagonistic co-evolution between the two sexes [2].

Experimental evolution provides an excellent empirical approach to study speciation processes, because environmental conditions can be well-controlled and replicate

69 populations are exposed to the same environmental stress. Nevertheless, only ecological
70 speciation has received strong empirical support by laboratory experiments [14-18] and in
71 most cases the selected traits were not identified.
72
73 Here, we take advantage of a highly replicated experimental evolution study, where 10
74 replicate populations derived from a single polymorphic *D. simulans* population have adapted
75 independently to a novel hot temperature regime in about 100 generations [19, 20]. Although
76 ambient temperature is one of the major ecological factors driving adaptation in natural
77 populations, it is not clear from previous work to what extent temperature has triggered
78 speciation processes [21]. We find support for ecological speciation, because premating
79 isolation occurred between ancestral and hot-evolved populations, but no difference was
80 found among the replicates, which evolved independently in the same novel hot
81 environment. Consistent with the predictions of mutation order speciation, we observed
82 postmating reproductive isolation between the evolved replicates. We propose that
83 ecological speciation is based on a modified lipid metabolism, which also causes shifts in CHC
84 composition, a well-understood mechanism of reproductive isolation in fruit flies [22-24].
85 Because mutation-order speciation can involve inter-genic antagonistic co-evolution [2], we
86 suggest that the high variance in evolutionary expression change of reproduction related
87 genes drives the observed post-mating reproductive isolation.
88
89

90 **Results**

91 After more than 100 generations of adaptation to a novel hot temperature regime, we
92 estimated reproductive isolation (Figure 1). As predicted for ecological speciation, we
93 observed significant mate discrimination between ancestral and hot-evolved populations but
94 not for replicate populations evolved independently to the same selection regime. While hot-
95 evolved males were more actively chasing females than ancestral males, the courting
96 behavior was significantly reduced when exposed to ancestral females. The same pattern was
97 seen for males from the ancestral population (Figure 2a, Kruskal-Wallis test, $p < 0.05$).
98 Reproductive isolation was further shown in an assortative mating experiment where flies
99 from two populations were mixed and mating partners could be chosen freely. Combining
100 ancestral and evolved populations resulted in strong positive assortative mating, while no
101 assortative mating could be detected among independently evolved populations (Figure 2b).
102 These patterns of premating reproductive isolation are fully consistent with the expectations
103 for ecological speciation: significant differences between flies adapted to different
104 environments, but no reproductive isolation between flies from different populations evolved
105 in the same environment.

106

107 Previously we showed that the lipid metabolism is significantly altered in the hot evolved
108 populations studied here [19]. Since the synthesis of cuticular hydrocarbons (CHCs), which
109 serve a central role in mate recognition [22-24], is controlled by genes involved in lipid
110 metabolism [25], we reasoned that the CHC composition may have been altered as a
111 byproduct of the change in lipid metabolism. We tested the divergence in CHC composition
112 between ancestral and evolved populations by measuring all hydrophobic compounds on the
113 outer skeleton of virgin flies of each sex with gas chromatography/MASS spectrometry

114 (GC/MS). We detected 18 major CHC compounds across both sexes, which differ in length of
115 carbon chain or numbers of double bonds [22, 25] (Table 1 and Supplementary Figure 1).
116 Principal component analysis (PCA) indicated that the relative abundance of the CHCs differed
117 significantly between sexes as well as between evolved and ancestral populations (Figure 3a).
118 The first PC, which explains 68.4% of total variance, reflects the differences between the two
119 sexes – males synthesize more unsaturated CHCs than females (Figure 3b). The variation
120 caused by adaptation to a hot laboratory environment is reflected by PC2 which accounts for
121 16.38% of the total variance (Figure 3a). No significant interactions between sexual
122 dimorphism and evolution were observed for these traits (Table 1), indicating parallel
123 evolution of CHC composition in both sexes. Evolved flies synthesize more unsaturated CHCs
124 with longer carbon chains (Figure 3b and c). Interestingly, similar modifications in CHC
125 compositions have been documented for multiple *Drosophila* species in latitudinal clines [26],
126 implying that the same causal link with temperature adaptation is also present in natural
127 populations.
128
129 We used gene expression data of the focal populations [19] to elucidate the basis of the CHC
130 composition changes. Among the 469 genes with a significant increase in expression across
131 replicate evolved populations, we found a prominent enrichment for genes involved in "very
132 long-chain fatty acid biosynthetic process" in both sexes. 27 out of the 36 (33 are detected in
133 our data) previously reported CHC synthesis-related genes in *Drosophila melanogaster* [25]
134 showed significant expression changes in at least one sex during adaptation in the replicated
135 *D. simulans* populations (Fisher's exact test, odds ratio = 3.33, p-value < 0.01; Supplementary
136 Figure 2). 16 of these genes changed their expression in the same direction in both sexes.
137 *Desaturases* (*desat 1* and *desat 2*) and *elongases* (*CG9458*, *CG18609* and *CG30008*) were

138 consistently up-regulated in all replicate evolved populations (Figure 3c and Supplementary
139 Figure 2). This consistent transcriptional modification of long-chain fatty acid/CHC
140 biosynthesis during the evolution can explain the significant changes in CHC composition in
141 the evolved flies.
142
143 Reproductive isolation can occur before and after mating. While the significant premating
144 isolation between ancestral and hot evolved populations supports ecological speciation, it is
145 not informative about mutation-order speciation. Because most of the mechanisms proposed
146 for mutation-order speciation operate after mating, we also tested postmating reproductive
147 isolation among evolved replicates. F1 flies from crosses between replicates of the evolved
148 populations produced 8.3% fewer viable offspring than F1 individuals from the same
149 replicates of the evolved populations (Figure 4b, Wilcoxon's test, $p = 0.031$). Considering the
150 absence of pre-mating isolation (Figure 2b), the reduced hybrid fitness indicates that another
151 (post-mating; fertilization- and/or viability-affecting) process occurs between the ancestral
152 and evolved populations and leads to reproductive isolation.
153
154 We used RNA-Seq to explore the functional basis of this post-mating reproductive isolation.
155 Reasoning that putative causal genes may diverge in gene expression, we searched for genes
156 differentially expressed among replicate evolved populations. In total, we identified 3,062
157 genes which differed in at least one evolved population from the others (Figure 5 —
158 Supplementary Figure 3, Supplementary file 1). Functional enrichment analysis revealed that
159 a substantial portion of these divergently expressed genes is highly expressed in testis and is
160 involved in reproduction-related processes (e.g. multicellular organism reproduction
161 (GO:0032504); Supplementary file 2). Focusing on 255 reproduction-associated genes that

162 diverged in expression across evolved replicated populations (see Materials and Methods),
163 replicate-specific selection responses were detected (Figure 5a). The heterogeneity in gene
164 expression evolution of reproduction-related genes was significantly higher than for a
165 random set of genes (Figure 5b, Wilcoxon's test, $p < 0.001$). Most of these reproduction-
166 associated genes changed their expression in only one or two replicate population(s) towards
167 the same direction (mostly up-regulation) during the adaptation (Figure 5a and
168 Supplementary Figure 4, and Supplementary file 2). Hence, reproduction related genes are
169 probably under selection in the new environment, but different sets of genes respond in
170 different replicate populations. Given that mutation-order speciation can be mediated by
171 inter-locus antagonistic coevolution [2, 27], our results provide a very compelling functional
172 explanation to our observation: reproduction-associated genes are known to be expressed in
173 males and females and their interaction is important for reproductive success [28, 29]. This
174 implies that changes in gene expression in one replicate probably require adjustments of
175 other genes to maintain functionality. The fitness decrease of hybrids between two
176 independently evolved populations may be explained by a perturbation of the co-evolved
177 gene expression patterns of reproduction-associated genes.
178
179

180 **Discussion**

181 The analysis of replicate populations adapting to the same environment demonstrated that
182 reproductive isolation could rapidly evolve during adaptation to a novel temperature regime.
183 The observed patterns of premating reproductive isolation were consistent with the
184 predictions for ecological speciation - populations evolved to different habitats showed
185 reproductive isolation, but no difference was detected between populations adapted to the
186 same environment. Connecting the adaptive response in lipid metabolism of hot evolved flies
187 to shifts in CHC profiles provides a plausible link between the adaptive response and the
188 pleiotropic byproduct of reproductive isolation. Hence, our study presents one example of
189 ecological adaptation where the target of selection (lipid metabolism) has been identified and
190 the pleiotropic effects of involved genes explain the premating reproductive isolation.

191

192 Remarkably, we did not only find evidence for ecological speciation, but also for mutation-
193 order speciation in the same experiment. While it has been previously proposed that inter-
194 locus antagonistic coevolution could drive speciation [2, 27] and the decoupling of the
195 coevolution of males and females has fitness consequences [30], the underlying molecular
196 processes have remained unclear. Our study provides an important step forward. By showing
197 that reproductive isolation between replicates is accompanied by an elevated variance in
198 gene expression evolution of reproduction-associated genes, we provide a plausible
199 explanation for the previously indirectly inferred coevolution between both sexes.
200 Importantly, several lines of evidence suggest that the differences between replicates are not
201 caused by genetic drift: 1) the variance in gene expression evolution of reproduction-
202 associated genes among replicates is higher than that of randomly selected genes (under
203 random genetic drift no enrichment for a GO category would be expected). 2) most of the

204 reproduction-associated genes changed their expression in the same direction (mostly up-
205 regulation) during adaptation 3) the higher expression of reproduction-related genes is
206 consistent with the increased reproductive activity of the evolved populations (Figure 2a and
207 [19]). This upregulation pattern was specific for reproduction-related genes and could not
208 found for other divergently expressed genes (Figure 5 — Supplementary Figure 5). 4) the post-
209 mating isolation among replicate evolved population is stronger for pairs with more diverged
210 gene expression. With the limitation of only three data points, this further supports that
211 alternative selection responses in replicate populations are the underlying genetic
212 mechanism (Figure 5 — Supplementary Figure 6). It is not clear, however, if the higher
213 reproductive activity of the hot-evolved flies is a direct adaptive response to temperature or
214 reflects an adaptation to laboratory conditions.
215
216 Mutation-order speciation assumes that independent new mutations are fixed in replicate
217 populations adapting to the same environment [6]. In our experiment, however, adaptation
218 is probably driven by standing genetic variation [31]. Given that most traits are polygenic [32],
219 it is conceivable that standing genetic variation is more important for speciation processes
220 than previously thought. Polygenic control and heterogeneous genomic divergence have
221 been documented in some iconic cases of replicated ecological speciation in nature such as
222 the convergent morphological radiation of cichlid species in East African lakes [33, 34] as well
223 as the repeated colonization to fresh water systems of oceanic stickleback species [35, 36].
224 We propose to expand the concept of mutation-order speciation by relaxing the condition of
225 novel mutations. Potential incompatibilities may evolve among sister species with the same
226 adaptive phenotypes, but alternative genetic pathways.
227

228 We observed the signature of mutation-order speciation from standing genetic variation in
229 an experimental evolution study in the laboratory, but the important question is whether
230 similar processes are likely to occur in natural populations. More work, both theoretical and
231 empirical, is needed to determine the importance of mutation order speciation form standing
232 genetic variation in a natural system.
233
234
235 **Material and Methods**
236 ***Experimental evolution***
237 The setup and maintenance of the experimental populations are detailed in [37]. In brief, 10
238 replicated outbred populations were constituted from the same 202 isofemale lines derived
239 from natural *Drosophila simulans* populations collected in Florida, USA in 2010. Five females
240 of each line were pooled to create one replicate of a reconstituted ancestral population
241 (henceforth called ancestral population). Replicated populations independently adapted to a
242 laboratory environment at 18/28°C with 12hr dark/12hr light photoperiod with the census
243 population size of 1250 adult individuals per population per generation Based on genomic
244 allele frequency changes, an effective population size of ~300 was inferred [37]. During the
245 evolution experiment, all replicate populations were reared in the same incubator with
246 randomization orientation/position across generations.
247 Each founder isofemale lines was maintained at small population size (typically less than 50
248 individuals) at 18 °C for >50 generations on standard laboratory food before the
249 reconstitution of the ancestral populations for common garden experiments. Potential
250 adaptation to the lab environment with the residual heterogeneity or de novo mutation is
251 unlikely, as discussed previously [37]. The small effective population size during the

252 maintenance of each line prevents adaptation to the lab environment. This has also been
253 experimentally tested by the lack of significant difference in allele frequencies in populations
254 which were reconstituted shortly after the establishment of the lines and after 50 generations
255 [38]. Furthermore, mutations occurring in the laboratory, if any, are mostly recessive [39],
256 and the effects are likely to be masked because after two generations of random mating
257 during common garden maintenance, most individuals will be heterozygous for isofemale
258 line-specific variants.
259
260 **Common garden experiment**
261 We performed multiple common garden experiments with the evolved populations and
262 reconstituted ancestral populations for the transcriptomic, metabolomic and phenotypic
263 assays in this study. The common garden procedures have been detailed in [19, 20]. Briefly,
264 replicates of ancestral populations were reconstituted by pooling five females each from the
265 founder isofemale lines. These reconstituted ancestral populations and evolved replicates
266 were reared at 18/28°C cycling with 12hr dark/12hr light photoperiod for at least two
267 generations to minimize transgenerational and/or environmental effects before subjected to
268 different phenotypic assays.
269
270 ***Assay for male reproductive activity***
271 We measured the reproductive activity for evolved and reconstituted ancestral populations
272 at generation 140. After four generations at common garden condition (18/28°C cycling), 10
273 pairs of 5-day-old mated males and females were placed together in an agar-based arena (4%
274 agar, 4% sugar) and filmed for 15 minutes at 20 FPS (frame-per-second) at 28°C using
275 FlyCapture2 system (PointGrey, Version 2.13.3.31). In total, for matched-setup (flies from the

276 same population are assayed together), eight trials within the evolved populations and six
277 trials within the ancestral populations were assayed. In addition, five and six trials of
278 reciprocal mismatched-setup (ancestral males were assayed together with evolved females
279 and vice versa) were included in the experiment. Each fly was tracked using flytracker [40] for
280 movement and behavior analysis. Janelia Automatic Animal Behavior Annotator (JAABA) was
281 used to annotate and recognize the chasing behaviors [41]. The time a male fly spent on
282 chasing females was quantified as its reproductive activity and the difference in the different
283 setups was tested with a Kruskal-Wallis test in R.
284
285 ***Multiple-choice mating assay for assortative mating***
286 We performed a multiple-choice mating assay for assortative mating between 1) three
287 ancestral populations and three selected evolved populations (repl. 1, 3 and 6, generation
288 194) 2) across the three independently evolved populations. We modified the experimental
289 protocol from [42]. After two generations of common garden condition (18/28°C cycling with
290 12hr dark/12hr light photoperiod), 400 unmated males and females were collected from each
291 population in six hours after eclosure. The collected flies were reared on corn meal in vials
292 with moderate density (50 flies per vial, 8 replicates per sex per population) for four days and
293 transferred onto yeast paste saturated with food coloring dyes (red or blue) one day before
294 assaying. During the mating assay, 50 flies of each sex from each population in a combination
295 of interest (one ancestral and one evolved populations/two evolved populations) were placed
296 together in a cylindrical cage with a sealable entrance for aspirator. Eight cages (four for each
297 color) for each population combination were setup and run simultaneously and each cage
298 was checked sequentially. Mating pairs were captured with aspirator and identified under
299 microscope based on the color of abdomen. The examination was terminated when the first

300 25 mating pairs in a cage were identified or when two hours elapsed. Based on a two-by-two
301 contingency table from the assay for each population combination, Yule's index (Y) [43] was
302 calculated to quantify the strength of assortative mating and Fisher's exact test was applied
303 for hypothesis testing.
304
305 ***Quantification and identification of cuticle hydrocarbons***
306 We measured the cuticle hydrocarbons (CHCs) for five replicates of the ancestral population
307 and 10 independently evolved populations at generation 158. After two generations in a
308 common garden, three replicates of 20 virgin flies of each sex were collected two hours after
309 eclosure from each population. The collected flies were aged for three additional days before
310 CHC extraction. We used 100 μl of heptane containing an internal standard (IS, 10 μg n-C30)
311 to wash the CHCs from each sample. The extracts were stored at -20°C until gas
312 chromatography. GC/MS analysis was performed to isolate and quantify each compound. The
313 analysis of chromatograph was performed with an Agilent 7890C gas chromatography system
314 coupled to an Agilent 5975C MSD. Injection volume was 1 µl in an inlet at a temperature of
315 320°C, a split of 10:1. He was used as carrier gas at a velocity of 1 ml/min and the compounds
316 separated in a HP-5MS column (30 m x 250 µm ID x 0.25 µm film thickness) with a
317 temperature program 180°C to 240°C with 6°C/min. and then with 20°C/min. to 320°C.
318 Transfer line temperature to the MSD was 280 °C, mass range of the MSD was 40-400 m/z
319 and electromagnetic voltage was set at 70 eV.
320 Identification of the constituents were done based on retention index (RI), determined with
321 reference to homologues series of n-alkanes (C8-C30) and mass spectra with the databases
322 NIST 05. Each identified chemical compounds were validated with published *Drosophila* CHCs
323 chromatograph (Dembeck et al., 2015; Pardy et al., 2019). For integration, identification and

324 quantification of compounds, Automated Mass spectral Deconvolution and Identification
325 System (AMDIS) [44] and openChrom [45] were used. The relative concentration of each CHC
326 compound (excluding cis-Vaccenyl acetate) was calculated and subjected to centered-log
327 ratio transformation [44]. Two-way analysis of variance (ANOVA) and principal component
328 analysis (PCA) were performed to dissect the effect of evolution and sex on different CHCs.
329
330 ***Cross-replicate compatibility assay***
331 At generation 194 of the evolved populations, we performed a diallel cross among three
332 selected evolved replicates (repl. 1, 3 and 6) and measured the number of progenies that
333 survived until adulthood to approximate the fitness of each cross. Since we did not find
334 assortative mating (Figure 2a), this assay can identify fitness differences caused by viability or
335 fertilization success. The ancestral populations were not included because of the potential
336 impact of assortative mating. After three generations of common garden rearing (18/28°C
337 cycling), 75 unmated males and females were collected from each evolved replicate
338 population and aged to four days old. For each cross, five males and five females from the
339 same/different population were placed together in a vial (with food surface area of 3.14cm^2;
340 no crowding effect is expected) at the same common garden condition and two transfers
341 were made for all vials with a 48-hour interval. The measurements of viable progeny numbers
342 after 14 days from both transfers were averaged and treated as one replicate for the data
343 analysis. Five replicates were generated for each cross and there were in total nine types of
344 cross in a full diallel cross design among three populations.
345
346 ***RNA-Seq data analysis for divergently expressed genes***

347 RNA-Seq data for the evolution experiment at generation 103 were obtained from previous
348 studies where contrasts between ancestral and evolved populations were made [19, 20]. In
349 this study, we reanalyzed the data for the 5-day-old whole-body male samples of the evolved
350 populations to identify the genes that were divergent across replicates. Only two replicates
351 are available from each evolution replicates for females, preventing similar analysis in females
352 [19, 46]. Genes with count per million (CPM) > 0.1 across all samples were retained and we
353 modeled their expression as: $Y = repl + \varepsilon$, where Y is the normalized expression values;
354 $repl$ indicates the effect among evolution replicates and ε is the random error. Likelihood
355 ratio tests implemented in edgeR [47] were used to perform differential expression analysis
356 on the effect of $repl$. Benjamini and Hochberg's FDR correction [48] was applied with a FDR
357 cutoff of 0.05. Genes differentially expressed in at least one replicate were identified for
358 further analysis.
359
360 ***Gene set enrichment analysis***
361 In order to explore the functional implication of the parallel changes and random divergence
362 in transcriptome, we tested for an enrichment of gene ontology (GO), tissue-specific
363 expression and reproductive functions among the genes evolving for a parallel response [19]
364 or those evolved for significant divergence across the 10 evolution replicates (this study). GO
365 enrichment was performed using default "weight01" algorithm implemented in topGO
366 (version 2.32.0) [49]. Genes highly expressed in each tissue were identified based on flyatlas2
367 expression dataset [50] (required > 2 fold higher expression in a certain tissue than whole-
368 body). Genes involved in the cuticle hydrocarbon (CHC) metabolism were obtained from [25].
369 Fisher's exact tests were applied for the enrichment analysis. Except for the GO enrichment

370 analysis which already accounts for multiple testing [51], Benjamini and Hochberg's FDR
371 correction [48] was applied to account for multiple testing.
372
373 ***Replicate-specific evolution in divergently evolving reproductive genes***
374 In order to understand the driving force underlying the divergent evolution in the expression
375 of genes involved in "multicellular organism reproduction" (GO:0032504), we investigated
376 the replicate-specific gene expression evolution of 255 reproductive genes in this term that
377 were significantly diverged across independent replicates. Using the expression matrix of
378 these genes in all samples (Figure 5 — Figure Supplement 2), we summarized the variation in
379 expression of the 255 genes by the principal components (PCs). The PC scores of each sample
380 reflect the expression difference across the sample. Thus, comparing different evolved
381 replicates to the ancestral samples could inform us about replicate-specific evolution in a
382 certain subset of genes. Additionally, we investigated the heterogeneity in the expression
383 changes across replicates as:
384 $$Heterogeneity = \sum_{i,j} (1 - cor(x_i, x_j))\,/N$$
385 Where x is a vector consist of the expression change of each gene in each evolution replicate,
386 $i, j \in [1,10], i \neq j$ and N is the total number of pairwise combination of 10 replicates.
387
388 ***Data and materials availability***
389 Sequence reads from this study is available from the European Sequence Read Archive
390 (http://www.ebi.ac.uk/ena/) under the study accession number PRJEB35504 and
391 PRJEB35506. Additional scripts and raw data are available on Github upon publication.
392

393 Table 1. Abundance of cuticular hydrocarbons (CHCs) in relation to sex and evolution

Peak No.	Retention index	Name	% in anc. female	% in anc. male	% in evo. female	% in evo. male	ANOVA (p-value)		
							Sex	Evolution	Sex:Evolution
1	2187	x-Docosene (x-C22:1)	0.96±0.02	0.69±0.02	0.88±0.02	0.66±0.01	1.18E-18	1.88E-09	9.04E-01
2	2195	cis-Vaccenyl Acetate (cVA)	-	-	-	-	-	-	-
3	2200	n-Docosane (n-C22)	1.35±0.04	0.58±0.01	1.26±0.03	0.67±0.03	1.37E-31	1.38E-02	3.40E-01
4	2279	9-Tricosene (9-T)	1.67±0.07	2.77±0.09	2.16±0.11	3.78±0.09	1.99E-27	1.23E-04	6.95E-01
5	2285	7-Tricosene (7-T)	75.48±0.42	75.29±0.44	73.64±0.44	71.70±0.52	2.64E-01	1.01E-06	6.02E-01
6	2294	5-Tricosene (5-T)	0.68±0.03	0.83±0.04	0.64±0.04	0.79±0.02	1.04E-03	1.47E-03	5.59E-01
7	2300	n-Tricosane (n-C23)	9.64±0.20	8.39±0.14	9.08±0.15	8.14±0.17	1.39E-05	1.65E-07	9.04E-01
8	2383	6-Tetracosene (6-C24:1)	0.64±0.01	0.58±0.02	0.87±0.03	0.86±0.03	1.43E-01	1.97E-12	5.49E-01
9	2391	5-Tetracosene (5-C24:1)	0.25±0.01	0.27±0.01	0.40±0.02	0.42±0.02	1.46E-01	1.33E-11	5.59E-01
10	2400	n-Tetracosane (n-C24)	0.34±0.02	0.13±0.02	0.39±0.02	0.15±0.02	1.52E-21	7.32E-01	6.27E-01
11	2465	7,11-Pentacosadiene (7,11-PD)	0.35±0.03	0.82±0.03	0.28±0.02	0.86±0.05	1.12E-29	1.53E-04	3.40E-01
12	2478	9-Pentacosene (9-P)	0.95±0.05	0.75±0.04	1.28±0.12	1.08±0.05	1.79E-06	9.40E-06	6.08E-01
13	2486	7-Pentacosene (7-P)	1.59±0.06	1.98±0.05	2.93±0.18	2.90±0.12	4.26E-02	8.37E-12	3.71E-02
14	2500	n-Pentacosane (n-C25)	1.44±0.05	1.00±0.04	1.31±0.03	0.85±0.03	2.45E-14	8.43E-07	5.59E-01
15	2665	7,11-Heptacosadiene (7,11-HD)	1.85±0.07	4.04±0.10	2.18±0.08	5.28±0.20	2.24E-45	1.38E-02	5.59E-01
16	2700	n-Heptacosane (n-C27)	1.56±0.08	0.68±0.05	1.28±0.09	0.45±0.05	1.22E-26	3.25E-05	7.45E-01
17	2863	7,11-Nonacosadiene (7,11-ND)	1.11±0.06	1.20±0.05	1.31±0.06	1.41±0.08	8.30E-02	7.21E-01	6.08E-01
18	3000	n-Triacotane (n-C30)	-	-	-	-			

394

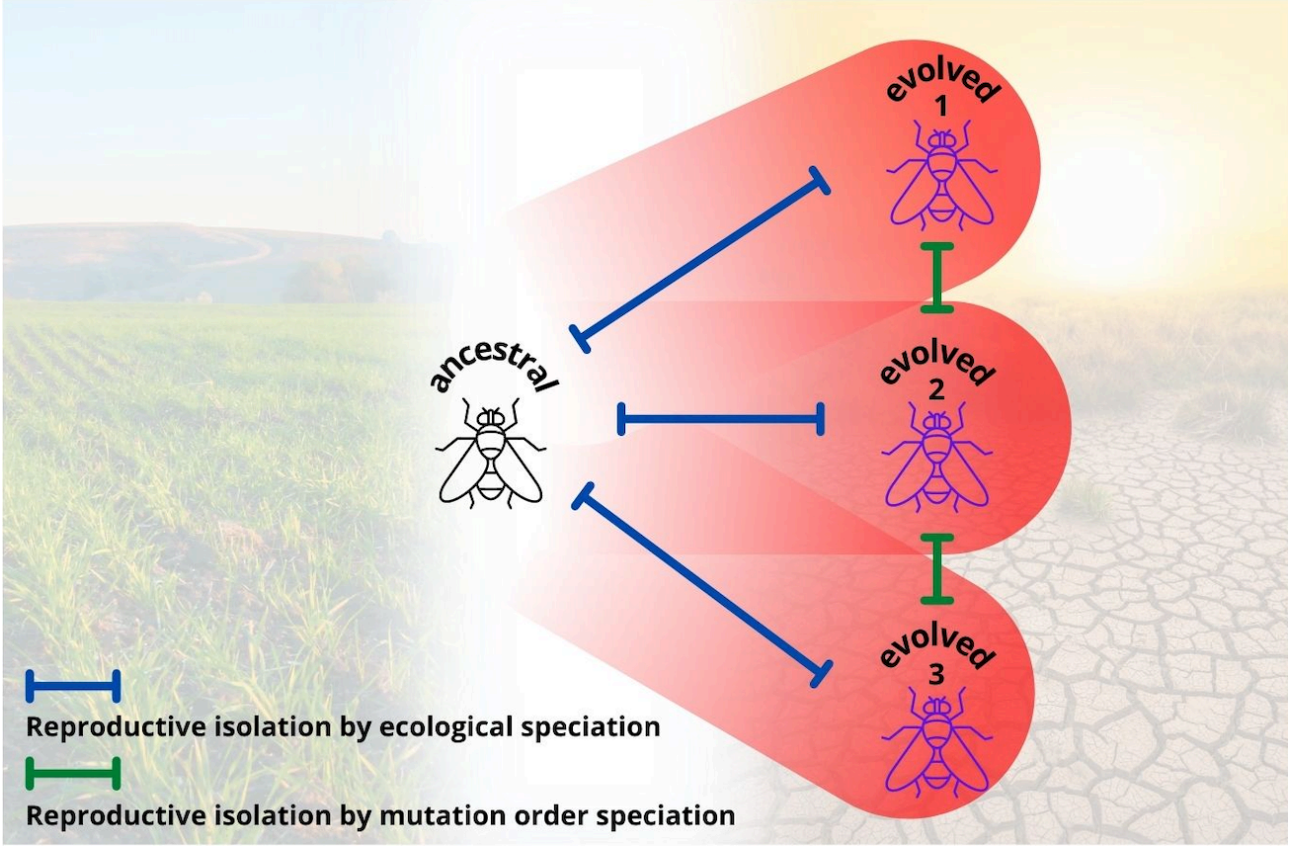

Figure 1. Schematic overview of the emergence of reproductive isolation during more than 100 generations of adaptation to a new hot environment. Replicate populations evolved at a hot temperature regime fluctuating between 18 and 28°C to mimic day and night conditions. Reproductive isolation evolved between ancestral and evolved population (ecological speciation) and among evolved replicates (mutation order speciation).

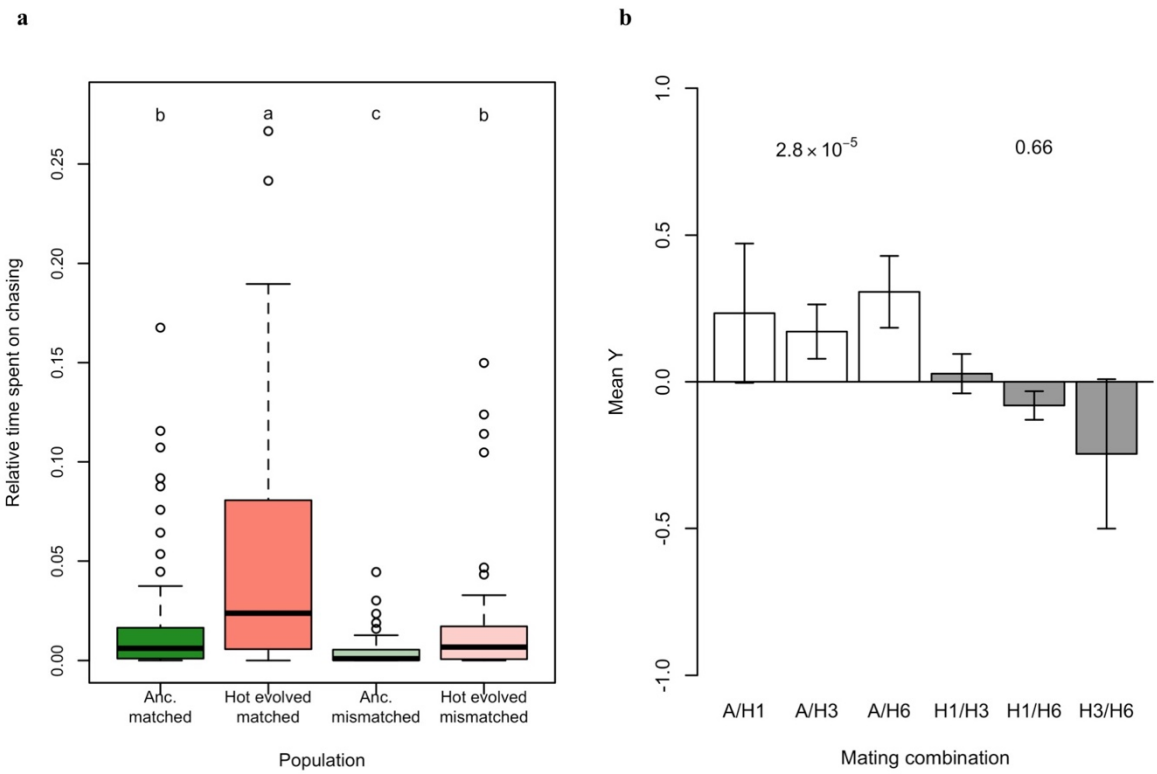

Figure 2. Evolution of mating preference and pre-mating reproductive isolation. a. Male reproductive activity in the presence of the females from the same evolved population or the ancestral one. The reproductive activity was measured as the time spent on chasing. Males spent significantly more time chasing females from the same populations (Kruskal-Wallis test, $p < 0.001$; post-hoc groups were indicated as the small-case letter above the boxes). **b.** Yule's index for assortative mating. There is significantly positive assortative mating between flies from the same populations (white bars) while there's no significant effect among the replicate evolved populations (p-values for Fisher's exact test).

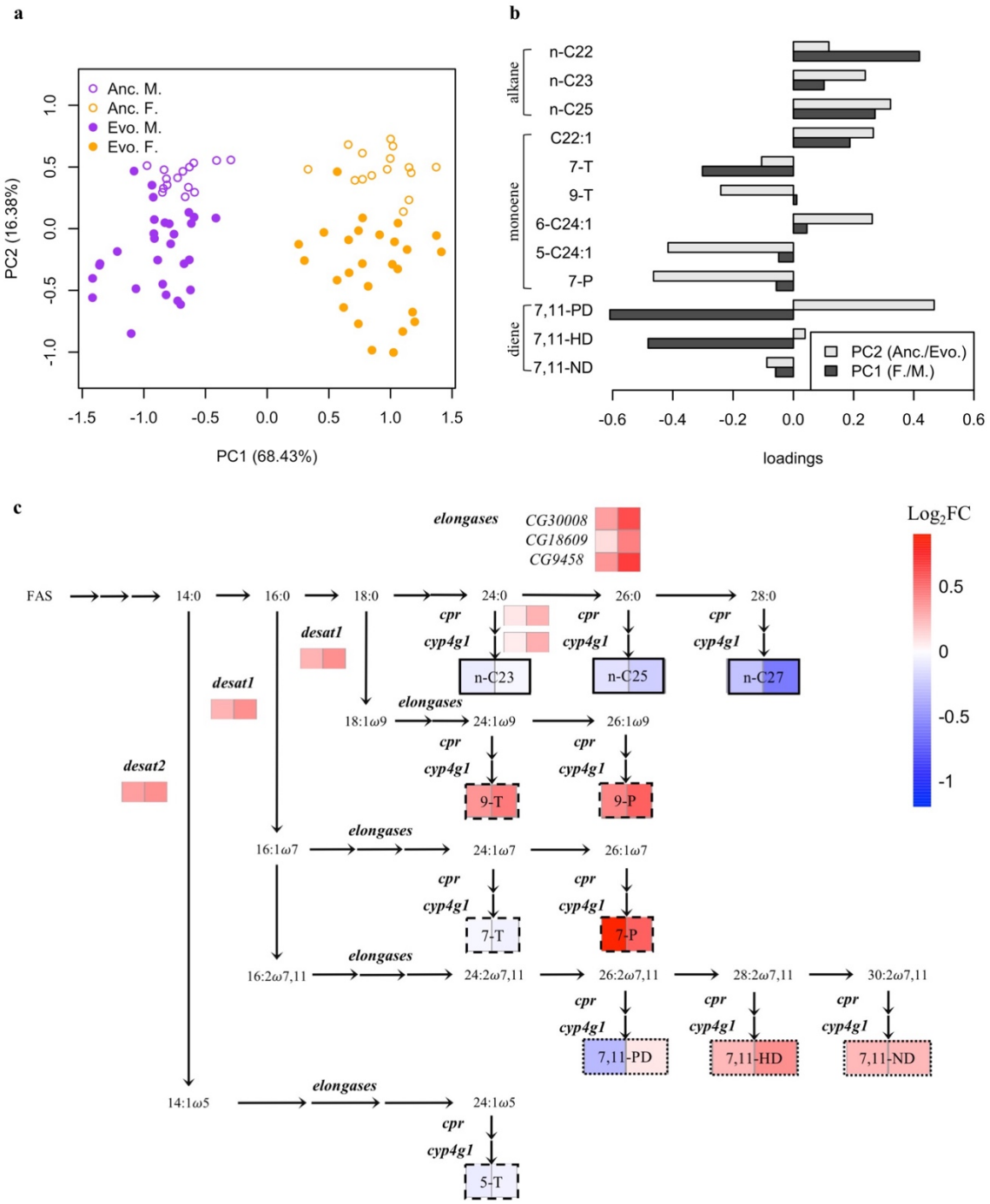

Figure 3. Evolution of cuticle hydrocarbons (CHCs) metabolism. a. Principal component analysis on the center-log transformed relative abundance of 12 CHCs detected in all samples. (orange: females; purple: males). Solid dots are evolved samples and empty ones are ancestral samples. **b.** Loading of each compound on the first and second principal components (PCs). Dark grey bars indicate the loadings for PC1. Positive loading suggests higher abundance in females. Light grey bars indicate the loadings for PC2 and positive values reflect higher abundance in the ancestral samples. **c.** Evolutionary changes in gene expression of the biosynthetic pathway for *Drosophila* cuticular hydrocarbons (CHCs). The heat maps underneath each compound (boxes with regular fonts) or genes (boxes with italic fonts) denote the changes in abundance during adaptation (log2-scale). The left cell denotes the changes in females and right cell for males. Red indicates increase and blue indicates decrease in abundance.

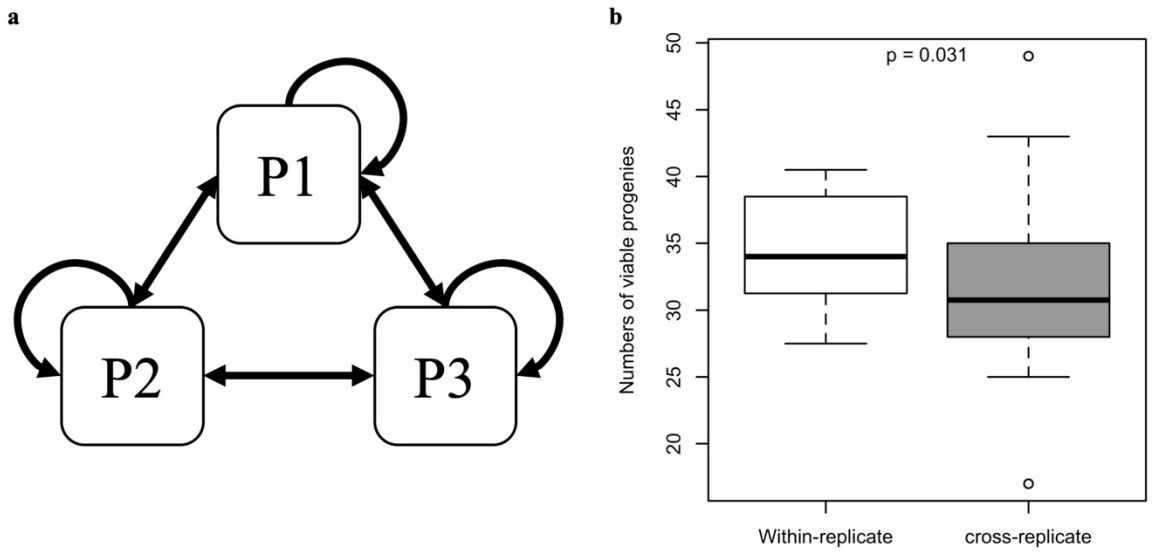

Figure 4. Evolution of post-mating incompatibility among independently evolved populations. a. Diallel cross design among three replicate evolved populations. Flies from each population were crossed with flies from the same population and the other populations reciprocally. **b.** Fitness difference between within-replicate crosses and cross-replicate ones. The fitness was measured as the numbers of viable adults in each cross. Fitness was significantly higher in within-replicate crosses ($p = 0.031$, Wilcoxon's test), indicating post-mating incompatibility between independently evolved replicates.

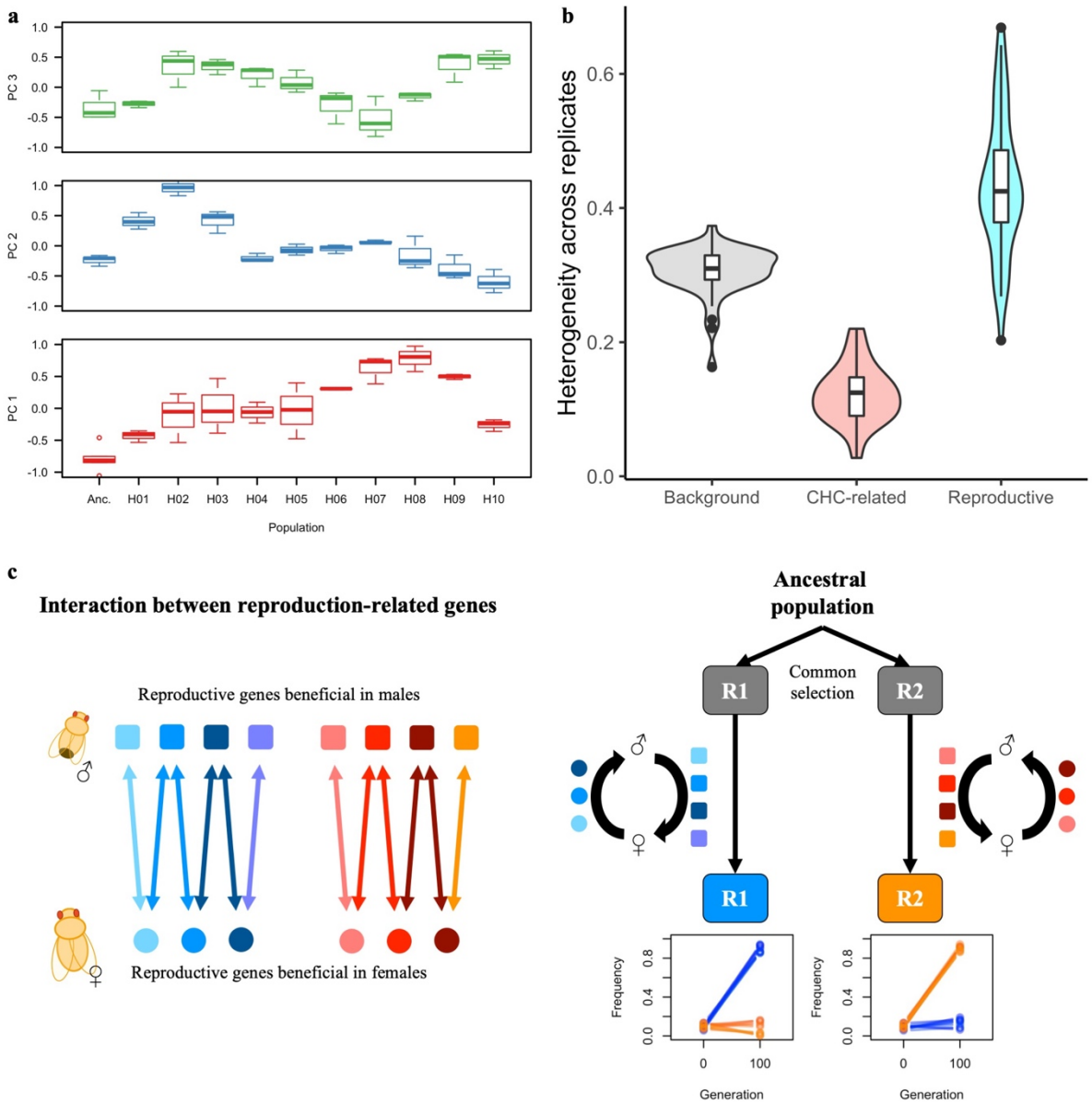

Figure 5. Heterogeneous adaptive changes in reproduction-related genes a. Replicate-specific evolution caused the expression divergence of reproduction-related genes. The expression variance among 255 reproduction-related genes (see materials and methods) is summarized by the leading principal components (PCs). The first three PCs suggest that different replicates evolved different sets of genes during adaptation. **b.** The heterogeneity in expression evolution differs among different sets of genes: background (all expressed genes), CHC-related (33 genes reported in (Dembeck et al., 2015)), reproduction-related (255 genes from GO term (GO:0032504)). The heterogeneity is measured as $\sum_{i,j}(1 - cor(x_i, x_j))/N$, where x is a vector consisting of the expression of each gene in each evolution replicate; $i, j \in [1,10], i \neq j$ and N is the total number of pairwise combination of 10 replicates. Reproduction-related genes evolved with significantly higher heterogeneity than background genes or CHC-related genes (Kruskal-Wallis test, $p < 0.001$). **c.** An illustration for how intersexual co-evolution could lead to heterogeneous and incompatible evolution in isolated populations: Genes involved in male reproduction also serve important functions in females via either synergistic or antagonistic interaction with female reproduction-related genes. Multiple epistatic modules with different genes involved (red- and blue-colors) may co-exist in the ancestral population. A cascade of intersexual co-evolution for a certain group of genes is expected if selection pressure is imposed on either sex. As such co-evolutionary cascades (modules) are not unique, but redundant, replicate populations may evolve for high divergence and

the alternative co-evolutionary paths may not be fully compatible with each other, resulting in fitness costs for between-replicate crosses.

References

1. Coyne JA, Orr HA: *Speciation.* Sunderland, MA: Sinauer; 2004.
2. Schluter D: **Evidence for Ecological Speciation and Its Alternative.** *Science* 2009, **323:**737-741.
3. Price T: *Speciation in Birds.* Greenwood Village, CO: Roberts & Company; 2008.
4. Darwin C: *On the Origin of Species by Means of Natural Selection.* London: J. Murray; 1859.
5. Schluter D: **Ecology and the origin of species.** *Trends Ecol Evol* 2001, **16:**372-380.
6. Mani GS, Clarke BC: **Mutational order: a major stochastic process in evolution.** *Proc R Soc Lond B Biol Sci* 1990, **240:**29-37.
7. Nosil P, Flaxman SM: **Conditions for mutation-order speciation.** *Proc Biol Sci* 2011, **278:**399-407.
8. Jiggins CD: **Ecological speciation in mimetic butterflies.** *Bioscience* 2008, **58:**541-548.
9. Villa SM, Altuna JC, Ruff JS, Beach AB, Mulvey LI, Poole EJ, Campbell HE, Johnson KP, Shapiro MD, Bush SE, Clayton DH: **Rapid experimental evolution of reproductive isolation from a single natural population.** *Proc Natl Acad Sci U S A* 2019, **116:**13440-13445.
10. Mallet J, Barton NH: **Strong Natural Selection in a Warning-Color Hybrid Zone.** *Evolution* 1989, **43:**421-431.
11. Nagel L, Schluter D: **Body Size, Natural Selection, and Speciation in Sticklebacks.** *Evolution* 1998, **52:**209-218.
12. Funk DJ, Nosil P, Etges WJ: **Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa.** *Proc Natl Acad Sci U S A* 2006, **103:**3209-3213.
13. Ono J, Gerstein AC, Otto SP: **Widespread Genetic Incompatibilities between First-Step Mutations during Parallel Adaptation of Saccharomyces cerevisiae to a Common Environment.** *Plos Biology* 2017, **15**.
14. Rice WR, Hostert EE: **Laboratory Experiments on Speciation: What Have We Learned in 40 Years?** *Evolution* 1993, **47:**1637-1653.
15. Dettman JR, Sirjusingh C, Kohn LM, Anderson JB: **Incipient speciation by divergent adaptation and antagonistic epistasis in yeast.** *Nature* 2007, **447:**585-588.
16. Kilias G, Alahiotis SN, Pelecanos M: **A Multifactorial Genetic Investigation of Speciation Theory Using *Drosophila melanogaster*.** *Evolution* 1980, **34:**730-737.
17. Dodd DMB: **Reproductive Isolation as a Consequence of Adaptive Divergence in *Drosophila pseudoobscura*.** *Evolution* 1989, **43:**1308-1311.
18. Rundle HD, Chenoweth SF, Doughty P, Blows MW: **Divergent selection and the evolution of signal traits and mating preferences.** *PLoS Biol* 2005, **3:**e368.
19. Hsu SK, Jaksic AM, Nolte V, Lirakis M, Kofler R, Barghi N, Versace E, Schlötterer C: **Rapid sex-specific adaptation to high temperature in *Drosophila*.** *Elife* 2020, **9**.
20. Jaksic AM, Karner J, Nolte V, Hsu SK, Barghi N, Mallard F, Otte KA, Svecnjak L, Senti KA, Schlötterer C: **Neuronal function and dopamine signaling evolve at high temperature in Drosophila.** *Mol Biol Evol* 2020.
21. Keller I, Seehausen O: **Thermal adaptation and ecological speciation.** *Mol Ecol* 2012, **21:**782-799.
22. Pardy JA, Rundle HD, Bernards MA, Moehring AJ: **The genetic basis of female pheromone differences between *Drosophila melanogaster* and *D. simulans*.** *Heredity* 2019, **122:**93-109.
23. Coyne JA: **Genetics of differences in pheromonal hydrocarbons between *Drosophila melanogaster* and *D. simulans*.** *Genetics* 1996, **143:**353-364.
24. Seeholzer LF, Seppo M, Stern DL, Ruta V: **Evolution of a central neural circuit underlies *Drosophila* mate preferences.** *Nature* 2018, **559:**564-569.

25. Dembeck LM, Boroczky K, Huang W, Schal C, Anholt RR, Mackay TF: **Genetic architecture of natural variation in cuticular hydrocarbon composition in Drosophila melanogaster.** *Elife* 2015, **4**.
26. Frentiu FD, Chenoweth SF: **Clines in cuticular hydrocarbons in two Drosophila species with independent population histories.** *Evolution* 2010, **64:**1784-1794.
27. Rice WR, Linder JE, Friberg U, Lew TA, Morrow EH, Stewart AD: **Inter-locus antagonistic coevolution as an engine of speciation: Assessment with hemiclonal analysis.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**6527-6534.
28. Ram KR, Wolfner MF: **Sustained post-mating response in *Drosophila melanogaster* requires multiple seminal fluid proteins.** *PLoS Genet* 2007, **3:**e238.
29. Chen DS, Delbare SYN, White SL, Sitnik J, Chatterjee M, DoBell E, Weiss O, Clark AG, Wolfner MF: **Female Genetic Contributions to Sperm Competition in *Drosophila melanogaster*.** *Genetics* 2019, **212:**789-800.
30. Rice WR: **Sexually antagonistic male adaptation triggered by experimental arrest of female evolution.** *Nature* 1996, **381:**232-234.
31. Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD: **Genome-wide analysis of a long-term evolution experiment with *Drosophila*.** *Nature* 2010, **467:**587-590.
32. Sella G, Barton NH: **Thinking About the Evolution of Complex Traits in the Era of Genome-Wide Association Studies.** *Annu Rev Genomics Hum Genet* 2019, **20:**461-493.
33. Kautt AF, Kratochwil CF, Nater A, Machado-Schiaffino G, Olave M, Henning F, Torres-Dowdall J, Harer A, Hulsey CD, Franchini P, et al: **Contrasting signatures of genomic divergence during sympatric speciation.** *Nature* 2020, **588:**106-111.
34. Kocher TD, Conroy JA, Mckaye KR, Stauffer JR: **Similar Morphologies of Cichlid Fish in Lakes Tanganyika and Malawi Are Due to Convergence.** *Molecular Phylogenetics and Evolution* 1993, **2:**158-165.
35. Lackey AC, Boughman JW: **Evolution of reproductive isolation in stickleback fish.** *Evolution* 2017, **71:**357-372.
36. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, et al: **The genomic basis of adaptive evolution in threespine sticklebacks.** *Nature* 2012, **484:**55-61.
37. Barghi N, Tobler R, Nolte V, Jaksic AM, Mallard F, Otte KA, Dolezal M, Taus T, Kofler R, Schlötterer C: **Genetic redundancy fuels polygenic adaptation in *Drosophila*.** *PLoS Biol* 2019, **17:**e3000128.
38. Nouhaud P, Tobler R, Nolte V, Schlötterer C: **Ancestral population reconstitution from isofemale lines as a tool for experimental evolution.** *Ecol Evol* 2016, **6:**7169-7175.
39. Charlesworth B, Charlesworth D: *Elements of Evolutionary Genetics.* Greenwood Village, Colorado: Roberts and Company Publishers; 2010.
40. Eyjolfsdottir E, Branson S, Burgos-Artizzu XP, Hoopfer ED, Schor J, Anderson DJ, Perona P: **Detecting Social Actions of Fruit Flies.** In *Computer Vision – ECCV 2014*; *2014//; Cham*. Edited by Fleet D, Pajdla T, Schiele B, Tuytelaars T. Springer International Publishing; 2014: 772-787.
41. Kabra M, Robie AA, Rivera-Alba M, Branson S, Branson K: **JAABA: interactive machine learning for automatic annotation of animal behavior.** *Nat Methods* 2013, **10:**64-67.
42. Kwan L, Rundle HD: **Adaptation to desiccation fails to generate pre- and postmating isolation in replicate Drosophila melanogaster laboratory populations.** *Evolution* 2010, **64:**710-723.
43. Bishop YM, Holland PW, Fienberg SE: *Discrete multivariate analysis, theory and practice.* New York: Springer; 2007.
44. Mallard WG, Reed J: **Automated Mass Spectral Deconvolution & Identification System AMDIS - User Guide. .** *Program* 1997.

45. Wenig P, Odermatt J: **OpenChrom: a cross-platform open source software for the mass spectrometric analysis of chromatographic data.** *BMC Bioinformatics* 2010, **11:**405.
46. Hsu SK, Jaksic AM, Nolte V, Barghi N, Mallard F, Otte KA, Schlötterer C: **A 24 h age difference causes twice as much gene expression divergence as 100 generations of adaptation to a novel environment.** *Genes (Basel)* 2019, **10**.
47. McCarthy DJ, Chen Y, Smyth GK: **Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation.** *Nucleic Acids Res* 2012, **40:**4288-4297.
48. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Ser B* 1995, **57:**289-300.
49. Alexa A, Rahnenfuhrer J, Lengauer T: **Improved scoring of functional groups from gene expression data by decorrelating GO graph structure.** *Bioinformatics* 2006, **22:**1600-1607.
50. Leader DP, Krause SA, Pandit A, Davies SA, Dow JAT: **FlyAtlas 2: a new version of the Drosophila melanogaster expression atlas with RNA-Seq, miRNA-Seq and sex-specific data.** *Nucleic Acids Res* 2018, **46:**D809-D815.
51. Kushwaha M, Salis HM: **A portable expression resource for engineering cross-species genetic circuits and pathways.** *Nat Commun* 2015, **6:**7832.