Host adaptation to viruses relies on few genes with different cross-resistance properties

Nelson E. Martins[1*], Vítor G. Faria[1*], Viola Nolte[2], Christian Schlötterer[2], Luis Teixeira[1], Élio Sucena[1,3**], Sara Magalhães[3,4**]

1 - Instituto Gulbenkian de Ciências, Oeiras, Portugal

2 - Institut für Populationsgenetik, Vetmeduni Vienna, Austria

3 - Universidade de Lisboa, Faculdade de Ciências, Departamento de Biologia Animal, Lisboa, Portugal

4 - Centro de Biologia Ambiental, Faculdade de Ciências da Universidade de Lisboa, Portugal

*Joint first authors; **Joint last authors

Corresponding authors:

Sara Magalhães (snmagalhaes@fc.ul.pt) and Élio Sucena (esucena@igc.gulbenkian.pt)

Universidade de Lisboa, Faculdade de Ciências, Departamento de Biologia Animal, Campo Grande, Edifício C2, 1749-016 Lisboa, Portugal

Abstract

Host adaptation to one parasite may affect its response to others. Yet, the genetics of these direct and correlated responses remain poorly studied. The overlap between these responses is instrumental for the understanding of host evolution in multi-parasite environments.

We determined the genetic and phenotypic changes underlying adaptation of *Drosophila melanogaster* to Drosophila C virus (DCV). Within 20 generations, flies selected with DCV showed increased survival after DCV infection, but also after Cricket Paralysis Virus (CrPV) and Flock House Virus (FHV) infection.

Whole-genome sequencing identified two regions of significant differentiation among treatments, from which candidate genes were functionally tested with RNAi. Three genes were validated, *pst*, a known DCV-response gene, and two novel loci: *Ubc-E2H* and *CG8492*. Knockdown of *Ubc-E2H* and *pst* also led to increased sensitivity to CrPV, while knockdown of *CG8492* increased susceptibility to FHV infection.

Therefore, *Drosophila* adaptation to DCV relies on few major genes, each with different cross-resistance properties, conferring host resistance to several parasites.

Significance statement

Despite ample knowledge of the genetics and physiology of host responses to parasites, little is known on the genetic basis of host adaptation to parasites. Moreover, adaptation to one parasite is likely to impact the outcome of different infections. Yet, these correlated responses, seminal to the understanding of host evolution in multi-parasite environments, remain poorly studied.

We determined the genetic and phenotypic changes underlying adaptation upon experimental evolution of a *Drosophila melanogaster* population under viral infection (DCV). After 20 generations, selected flies showed increased survival upon infection with DCV and two other viruses.

Using whole-genome sequencing and through RNAi, we identified and functionally validated three genes underlying the adaptive process and revealed their differential roles in the correlated responses observed.

\body

Introduction

Parasites impose a strong fitness cost on their hosts as they develop and reproduce at the expenses of host resources. Therefore, it is expected that host strategies will be selected to cope with parasite burden. There is ample variety of such strategies, from behavioral to intracellular responses (1). Because the range of possibilities is very broad, it is difficult to predict which strategy, if any, will evolve in host populations upon parasite attack. Moreover, in natural populations, hosts are exposed simultaneously to several parasite species and many other selection pressures. If these selection pressures do not vary independently of each other, a clear establishment of causality between changes in host traits and the selection pressure posed by a given parasite species may be hampered.

Experimental evolution enables the establishment of a direct link between the selection imposed by a given environment and the genetic and phenotypic changes observed in a population. The explanatory power of this methodology relies on three major characteristics: (i) knowledge of the ancestral state, (ii) control of the selection forces driving different sets of replicated populations and (iii) the ability to follow the dynamics of a process, instead of measuring only its end-product (2). In addition, this methodology allows addressing the consequences of the adaptation process for the performance in other environments (3–5).

Experimental evolution coupled with whole-genome approaches can provide a nearly unbiased view of the actual targets of selection, a long-standing aim of evolutionary biology (2). To this day few examples exist in which these combined

methodologies have been used in multicellular sexual organisms, in which most adaptation comes from standing genetic variation (SGV) instead of novel mutations (6–10). However, despite the centrality of host-parasite interactions in evolutionary biology, and several experimental evolution studies in host-parasite systems (11–16), no study of host-parasite interactions has combined experimental evolution with genomics.

Another important aspect of experimental evolution is that it allows measuring the consequences of evolving in one environment for the performance in other environments (3). Indeed, adaptation to one environment may entail a fitness decrease in other environments, possibly hampering future evolution in such settings (17, 18). Despite being common, these costs are not universal (4) even within experiments (17). Moreover, adapting to one environment may even lead to increased performance in other environments (e.g. 5, 19). In host-parasite interactions, the question is particularly important because of the epidemiological consequences of infecting or resisting multiple hosts or parasites, respectively.

Despite ample knowledge of the genes triggered by parasite attacks against *Drosophila*, only a few key studies have analyzed how an outbred fly population may adapt to a given parasite (11–13, 15). Yet, the genetic basis and the consequences of such adaptation for host susceptibility to other parasites have not been determined.

It has been shown that natural *D. melanogaster* populations contain standing genetic variation for resistance against natural viruses. Whereas some studies show that most of this variation can be attributed to a limited number of genes with major effect (20–23), others indicate that a significant fraction of the genetic variation for resistance is polygenic (24, 25). Interestingly, the alleles that contribute to the variation in resistance

to a given virus are of genes unrelated to the canonical insect anti-viral defense pathways (26). Moreover, this variation may be rather specific in mediating responses to distinct natural pathogens (21).

Here, we addressed the genetics of host adaptation to parasites and the effects in cross-resistance in a *D. melanogaster*-virus system. To this aim, we performed experimental evolution of an outbred *D. melanogaster* population exposed to a natural viral parasite (Drosophila C virus - DCV), analyzed the basis for the response using a genome-wide approach, and functionally tested candidate genes for their role in the response against DCV and other parasites.

Results

1. Adaptation to DCV infection

We have performed experimental evolution of an outbred *D. melanogaster* population exposed to recurrent systemic DCV infection (VirSys). DCV infection was imposed at every generation using the same (not co-evolved) ancestral virus strain. In parallel, two control conditions were established, where individuals were subjected to the same procedure as the virus-selected population but pricked with a buffer solution only (ContSys) or not pricked at all (Control). The experiment was performed with four replicates for each condition.

When exposed to DCV, VirSys populations showed higher survival than individuals from Control lines (Fig. 1A; general linear mixed model (GLMM), $\chi^2_1 = 154.98$; $p < 0.0001$). Changes in survival in the VirSys selection regime were consistent among replicate

populations (Fig. S1A). The difference in survival was absent in the early generations and increased with time, leading to a significant interaction between generation and selection regime (Fig. 1A, Dataset S1 and Fig. S1A, GLMM χ^2_{30}=163.54, p< 0.0001). When tested independently in the two sexes, both effects of selection regime (GLMM χ^2_1= 20.489 and 24.288, p< 0.0001 for males and females respectively) and interaction with generation (GLMM χ^2_{30}= 236.95 and χ^2_{26},=145.89, p< 0.0001, for males and females, respectively) were significant. Given that we were comparing Control (not pricked) with VirSys individuals, and that ContSys populations were used in all subsequent tests, survival of ContSys and Control populations were directly compared at generations 15 and 25. No significant differences were observed between the two sets of Control lines (Table S1).

VirSys lines showed a strong reduction of virus numbers when compared with ContSys lines (Fig. 1B; ANOVA, $F_{1,6}$=39.55, p= 0.0008) indicating that selection has relied (at least partially) on the evolution of resistance.

Next, we tested the contribution of *Wolbachia* to the evolution of resistance in our populations as this endosymbiont has been shown to protect *Drosophila* against viral infections (27). To this end, we removed *Wolbachia* from replicates of VirSys and ContSys populations, after 25 generations of selection and measured survival upon DCV infection (Fig. 1C). A significant interaction was found between sex and both *Wolbachia* and selection regime (Cox model, χ^2_1 = 56.705 and 17.150, respectively, p< 0.0001 in both comparisons). Therefore, we tested the effects of *Wolbachia* and selection regime independently for both sexes (Fig. S1B). In both cases, there was a significant *Wolbachia*

and selection regime effect (Cox model, $\chi^2_1 = 29.110$ and 34.94, for *Wolbachia* and selection regime effect in males; $\chi^2_1 = 24.865$ and 22.824 for *Wolbachia* and selection regime effects in females, respectively; $p < 0.0001$ in all comparisons). Therefore, the protective role of *Wolbachia* against viral infections (27) is confirmed in this study on both experimental and control lines. However, no significant effect of the interaction Wolbachia * selection regime was found for either sex (Cox model, $\chi^2_1 = 0.255$, $p = 0.613$ and $\chi^2_1 = 1.007$, $p = 0.316$ for males and females, respectively). This indicates a significant contribution of the host genome to the evolution of resistance, which is statistically independent of the effect of *Wolbachia* infection status.

2. Cross resistance to other parasites

As shown in figure 2, VirSys populations also had on average higher survival, relative to ContSys, after infection with the parasites Cricket Paralysis Virus (CrPV) or Flock House Virus (FHV) (Cox model, $|z| = 19.857$, 11.329 and 5.226, for infection with DCV, CrPV and FHV, respectively, $p < 0.0001$ for all comparisons). There was a significant interaction effect with the generation at which the test was conducted, for the different parasites (Cox model, $\chi^2_3 = 31.276$, $p < 0.001$ for DCV, $\chi^2_1 = 4.192$, $p < 0.05$ for CrPV and χ^2_2, $p < 0.05$ for FHV). However, the difference between the VirSys and ContSys regimes was significant in all separate tests performed at different generations and for the different viruses (Cox model, $|z| = 14.480$, 10.790, 13.454 and 7.337 for DCV infections performed at generations 15, 20, 25 and 30; $|z| = 1.122$ and 1.438 for CrPV infections at generations 15 and 30; and $|z| = 0.514$, 0.327 and 0.804 for FHV infections at generations 15, 20 and 30. $p < 0.001$ in all comparisons, except for the FHV infection at generation 20,

where $p < 0.05$). However, the hazard ratios between ContSys and VirSys exposed to FHV infection are significantly lower than those observed upon exposure to DCV (used for selection) or against CrPV, a very close DCV relative (Fig. 2).

No significant difference in survival among selection regimes was found when flies were infected with the bacteria *Pseudomonas entomophila* and *Enterococcus faecalis* (Cox model, $|z| < 0.446$, $p > 0.66$ for all comparisons after infection with *P. entomophila* at generations 15 and 25 or with *E. faecalis* at generations 34 and 35). We therefore conclude that evolution of resistance to DCV leads to partial protection against other positive strand RNA viruses, but not against bacterial pathogens.

3. Genetic basis of host adaptation

To identify the changes in allele frequencies underlying the observed increased resistance of *Drosophila* populations evolving in presence of DCV, we performed genome-wide sequencing of DNA pools ("Pool-Seq") of all populations (Fig. 3) (28). Patterns of overall genetic diversity are presented in the supplementary materials (Fig. S2).

Using a chromosome-wide cut-off, we observed consistent significant changes in allele frequencies of 853 SNPs over a region that spans approximately 4 Mb on chromosome arm 3L (most 5' SNP, 3L:5127093 and most 3' SNP, 3L:9149494) and five SNPs on the X chromosome across a 300 kb region (X:7638809-7984449). This result did not change qualitatively using a genome-wide cut-off, but the region of significance was reduced to positions 3L:5221901-8901948 (i.e., 384 SNPs), and to two SNPs on the X chromosome. The most significantly differentiated SNP in the 3L region corresponds to position 3L:7350895 and maps to the gene *pastrel* (*pst*). The two significantly

differentiated SNPs on the X chromosome (X:7984325 and X:7984449) are located in introns of the gene *Ubc-E2H*. Initial and final frequencies of the most significantly differentiated SNPs were 0.167 and 0.7 for 3L:7350895 (*pst*) and 0.267 and 0.6 for X:7984325 (*Ubc-E2H*) respectively. Considering these changes in frequency, and assuming additive effects only, the estimated selection coefficients are 0.24 and 0.14 for the SNP in *pst* and *Ubc-E2H*, respectively. Changes in other significantly differentiated SNPs are described in Dataset S2.

4. Functional validation of the candidate genes

We then used RNAi to functionally validate the two genes associated to the most significant SNPs identified in the genome-wide analysis. We further tested 12 genes in the 3L region, which contained non-synonymous mutations (Fig. 4).

Knockdown of *pastrel* and *Ubc-E2H* (with stock w^{1118}; P{GD9765}v33510, see Table S2 for details) led to reduced survival of flies when exposed to DCV or to CrPV infection (Fig. 4A: *Ubc-E2H*: $|z| = 3.98$ and 3.09, $p < 0.01$ and $p < 0.05$, after DCV and CrPV infection, respectively; Fig. 4B: *pst* $|z| = 5.94$ and 5.93, $p < 0.001$ after DCV and CrPV infection), but not when exposed to FHV infection (*Ubc-E2H*: $|z| = 1.35$, $p > 0.9$ and *pst*: $|z| = 0.08$ for knockdown of both genes). Using another RNAi line targeting *Ubc-E2H* (with stock P{KK108626}VIE-260B, see Table S2 for details) did not show differences in survival against any of the viruses ($|z| = 2.25$, 0.11 and 0.12, for DCV, CrPV and FHV respectively, $p > 0.3$) (Fig. 4A). We attribute this survival difference using two different RNAi lines to a lower knockdown efficiency of this construct, as revealed by semi-quantitative gene expression analysis (Fig. S3). No differences in

susceptibility to virus were observed when comparing the negative control with the respective genetic background (|z| = 0.71, 0.93 and 0.19 for DCV, CrPV and FHV respectively, $p > 0.97$).

RNAi knockdown of another 12 genes within the 3L region revealed only one other case, gene *CG8492* (stock P{KK100300}VIE-260B), with reduced survival upon exposure to DCV and to FHV (Fig. 4B, |z| = 4.23 and 3.23, $p < 0.001$ and $p < 0.05$ for DCV and FHV, respectively), but not to CrPV (|z| = 0.24, $p = 1$). All p values were Bonferroni corrected for the number of performed comparisons.

Discussion

In this study, we found that resistance to DCV evolved rapidly in experimental *Drosophila* populations. Cross-resistance was detected for infection with other viruses (CrPV and FHV) but not with bacteria. Using whole-genome-sequencing, we identified two regions in which genetic changes occurred in populations evolving under DCV challenge, one in the 3L chromosome arm, and a smaller region on the X chromosome. Through RNAi assays against candidate genes in these regions, we confirmed the role of *pastrel* (*pst*), a gene with variants previously associated with differential response to DCV infection in *Drosophila* (21), as well as two loci that had not been linked previously to anti-viral response: *Ubc-E2H* on the X chromosome and *CG8492* on the 3L chromosome arm. Knockdown of *pst* and *Ubc-E2H* led to increased sensitivity to CrPV, but not to FHV, whereas the opposite pattern was found for *CG8492*. Hence, flies that

177 have adapted to resist to DCV are also better at surviving infection with other viruses, but
178 these correlated responses rely on different sets of genes.

179 **Genetic basis of resistance**

180 Using a combination of genomics with experimental evolution, we identified the genetic
181 changes underlying the evolution of a host population (*Drosophila melanogaster*)
182 adapting to a natural parasite (DCV). We find two regions of differentiation between the
183 populations evolving in presence of a virus and control populations. These changes were
184 parallel across four replicates (Fig. S2 and dataset S2) and correlate with the observed
185 parallel changes in survival (Fig. S1A). This indicates that selection, rather than drift,
186 shaped this adaptive response. In one region, the peak of differentiation matched *pst*, a
187 gene previously shown to be involved in *Drosophila* response to DCV through an
188 association study (21). The high number of differentiated SNPs around this locus,
189 extending to a region of approximately 4Mb, and the observed pattern of local decrease
190 of heterozygosity suggests the occurrence of an incomplete soft sweep around pst (29).
191 However, the influence of other genes in the region cannot be excluded, as shown by the
192 increased susceptibility of flies expressing RNAi against *CG8492*, a gene located near the
193 centromeric end of the peak. The determination of the haplotype structure in this region,
194 as well as the effect in virus resistance of the variants of *CG8492* and their possible
195 interactions with *pst*, deserve further examination.
196 This result is particularly interesting in that it departs from the inconsistency observed
197 when comparing genome-wide-association-studies (GWAS) using inbred lines *versus*
198 outbred populations (30). Thus far, only a weak but significant correlation has been found

between SNPs associated with polygenic traits by GWAS and "Evolve and Resequence" (E&R) approaches (31). Here, we confirm *pst*, a gene found through a GWAS approach (21), as a central player in the adaptation of an outbred population of *Drosophila* to DCV infection.

Furthermore, using RNAi we confirmed the role of *pst*, and unraveled an effect of *Ubc-E2H* and *CG8492* in antiviral defense. These results confirm the power of the E&R approach in the identification of targets of selection (32). This methodology has been used to identify changes in allele frequencies following selection in complex traits such as developmental time (7), body size (8), hypoxia tolerance (6), increased life span (33), adaptation to high/low temperatures (9, 34) and courtship behavior (10, 31). These studies have identified a polygenic basis for the studied traits, hampering the identification of candidate genes and a subsequent functional analysis. One exception is the study of Zhou et al. (6), in which most of the differentiated genes belonged to the Notch signaling pathway, thus permitting a functional validation of this pathway in hypoxia tolerance evolution. However, the relatively high number of genes involved in these responses do not permit the assessment of the role played by each gene and how the phenotypic effect may be partitioned. In our case, the few genes underlying the evolution of resistance to DCV seem to work in an (partially) additive fashion, as each gene tested independently confers resistance. Yet, further studies are needed to establish the relative role of additivity and genetic interactions in this response.

Cross-resistance

We find a strong positively correlated response with CrPV, but only a moderate response to FHV, and no response to bacteria. Hence, the correlated response is positive and diminishes with decreasing similarity to DCV. Both these findings match recent theoretical predictions for one-sided host evolution (14). However, other studies on host evolution have found trade-offs (16, 35) or no significant correlated response (36, 37) among resistance to different parasites, hence the generality of our finding remains to be shown.

We analyzed the correlated responses of the genes involved in DCV resistance when flies were infected with other viruses. To our knowledge, this constitutes the first direct test of the genetic basis of correlated responses to selection driven by standing genetic variation. Analysis of the effects of *de novo* mutations that arise in *E. coli* populations adapting to a glucose-limited environment when placed in other environments, had also shown that the set of mutations conferring fitness increases varies between environments (38). Similarly to that study, we find that distinct genes for which allelic frequencies have changed in response to DCV infection, affect correlated responses differently. Indeed, knockdown of *pst* does not affect susceptibility to FHV, confirming earlier results (21); but knockdown of either *pst* or *Ubc-E2H* affects cross-resistance to CrPV. In contrast, knockdown of *CG8492* does not affect the response to CrPV but leads to higher susceptibility to FHV. Therefore, in our populations, the evolution of a generalized response to viral parasites is specifically partitioned into different loci.

241 Until now, the genetic analysis of correlated responses has relied on measuring
242 the genetic correlation among traits in different environments using quantitative genetics
243 designs (3). This methodology has also been used in the study of host-parasite
244 interactions (39, 40). However, it has been shown that genetic correlations are poor
245 predictors of the evolution of correlated responses to selection, mainly because the latter
246 hinges on the genetic architecture of traits under each environment (41). In our study, we
247 do not measure the whole genetic architecture of the traits under selection, namely
248 because we miss genes involved in resistance that are fixed and those with changes
249 occurring below our threshold value. Still, we detect those genes in which allele
250 frequencies change across generations, and hence contribute to the evolutionary response.
251 By describing that these genes have different cross-resistance properties against different
252 parasites, we show that the genetics of correlated responses may be complex, even in
253 cases where the genetic basis of adaptation is relatively simple.
254 Our findings raise an important issue: which forces maintain the standing genetic
255 variation (SGV) upon which is based host adaptation to viral infection? We have not
256 found costs in susceptibility to other parasites associated to the evolution of resistance to
257 DCV. Hence, our results do not support the maintenance of diversity via antagonistic
258 pleiotropy (3). This does not rule out that trade-offs with susceptibility to other parasites
259 exist, which we have not included in our tests. Still, for the parasites tested, we show
260 evolution of positively correlated responses, which depend on different genetic
261 architectures in a parasite-specific manner. This raises the possibility that, even in cases
262 where a generalized response evolves, specificities at the genetic level may lead to
263 different genetic responses in environments with qualitatively different parasite

challenges. This extends the possibility of maintaining genetic diversity across host populations (42), even when phenotypic responses suggest a generalized response to several parasites. A formal test of this hypothesis will require evolving and re-sequencing outbred populations in environments with different combinations of viruses.

It is generally believed that the occurrence of specific host genotype x parasite genotype interactions (Gh x Gp) relies on simple genetic bases (43–45). Here, we show that although the genetic basis of host adaptation to a parasite is simple, a generalist response has evolved. Therefore, a simple genetic basis is a necessary, but not sufficient condition for the evolution of specific interactions. However, it should be noted that our findings concern the outcome of an evolutionary process in which no coevolution has occurred. Therefore, more studies identifying the genetic basis of coevolution are required (44, 46). In particular, it will be highly informative to compare the genetic architecture of cross-correlations in coevolved systems with that of the present study.

Materials and Methods

Fly populations

We used an outbred population of *Drosophila melanogaster* founded and maintained as described in Martins et al. (15) and kept at high effective populations size (see Suppl. Information). Prior to the initiation of experimental evolution, this population was serially expanded for two generations to allow the establishment of 36 new populations, of which twelve were used in this work. Except otherwise noted, flies were maintained under constant temperature (25°C), humidity (60–70%) and light-darkness cycle (12:12), and

fed with standard cornmeal-agar medium. The populations were fully infected with *Wolbachia* at the onset of the experiment, and this infection status of the populations was monitored throughout the experiment.

Parasite stocks and cultures

Drosophila C Virus (DCV), Cricket Paralysis Virus (CrPV) (a kind gift from Peter Christian) and Flock House Virus (FHV), were grown and titrated as described before (27). Virus aliquots were kept at -80 ºC and thawed prior to infection. *Pseudomonas entomophila* and *Enterococcus faecalis* were generous gifts from B. Lemaitre and T. Rival, respectively. Bacteria stocks were kept in glycerol at -80 ºC. Prior to use, they were streaked in fresh Petri dishes, then a single colony was picked and let to grow in LB at 30 ºC (*P. entomophila*) or 37 ºC (*E. faecalis*). The culture was then centrifuged and adjusted to the desired O.D.

Experimental evolution

Starting from the base population, we derived 12 lines evolving under three different regimes (4 replicates per treatment). In the VirSys treatment, adult flies were pricked in the thoracic region with DCV ($2x10^7$ $TCID_{50}$) at each generation. A second treatment consisted of a control for pricking, in which the needle was dipped in sterile medium (ContSys). Finally, a second group of control lines consisted of flies kept in standard food without being pricked (Control). No differences between ContSys and Control lines were found for any test made with both sets of lines. The dose of DCV was used caused an average mortality of 66% in the initial population, 10 days after infection (Fig. S4).

These treatments were administrated to 310 males and 310 females (4-6 days after eclosion). Selection lines were kept in large population cages, surviving individuals mated randomly, and reproduction took place at days 5-7 after infection, by providing fresh oviposition substrate. The number of individuals in the control populations was always reduced to the initial number of infected individuals (i.e. 600). Since several selection lines were running in parallel, each with different selection dynamics (15), we opted to maintain a constant number of individuals in the controls, recognizing a possible upward bias in census sizes of the control lines.

Egg density was limited to 400 per cup, a density determined experimentally to enable optimal larval development. Each generation cycle lasted three weeks. Prior to the beginning of the experiment, absence of vertical transmission of the parasite to the progeny was verified (Fig. S5).

To monitor survival across generations, we infected at each generation an additional sample males and female flies from each of the VirSys lines and Control lines and monitored their survival in vials for at least 10 days (Dataset S1).

Parasite loads

Virus quantifications were performed as described in Teixeira et al. (27) with minor modifications. For each assay, 75 to 125 females from each population of ContSys and VirSys at generation 33 were infected as in the survival assays. Surviving flies were collected on day 5 after infection, pooled in 5 replicates of 10 individuals per population, and snap frozen in liquid N_2. RNA was extracted using TRIZOL®. To avoid possible

328 artefacts due to different maternal effects, flies used in these tests were the progeny of
329 flies that spent one generation in a common environment without the virus.

330 **Wolbachia**

331 *Wolbachia*-free replicates of the ContSys and VirSys populations were derived at
332 generation 25, by raising the progeny for two generations on food with tetracycline (0.05
333 mg/ml). Two generations after tetracycline treatment, 100 individuals (males and
334 females) from each replicate population of the VirSys and ContSys selection regimes and
335 their *Wolbachia*-free counterparts, were systemically infected with DCV and their
336 survival was followed for 16 days.

337 **Cross resistance with other parasites**

338 To test how adaptation to a specific parasite affected host responses to other parasites,
339 100 individuals (males and females) from each replicate population of the VirSys and
340 ContSys selection regimes, which had spent one generation in a common environment,
341 were systemically infected with the following parasites: CrPV (undetermined $TCID_{50}$),
342 FHV ($TCID_{50}=5x10^6$), *P. entomophila* ($OD_{600}=0.01$) and *E. faecalis* ($OD_{600}= 3$). These
343 tests were performed at generations 15, 20, 25 and 30 (DCV), 15, 20 and 30 (FHV), 15
344 and 25 (*P. entomophila*), 15 and 35 (CrPV) and at 34 and 35 (*E. faecalis*).

345 **Whole genome sequencing**

346 Genomic DNA preparation and sequencing were done as in Orozco-TerWengel et al. (9).
347 Briefly, a pool of 200 individuals of each selection line was homogenized with an
348 Ultraturrax T10 (IKA-Werke, Staufen, Germany), and DNA was extracted from the
349 homogenate using a high salt extraction protocol. Genomic DNA was sheared using a

Covaris S2 device (Covaris, Inc. Woburn, MA, USA) and paired-end libraries were prepared using the TruSeq v2 DNA Sample Prep Kit (Illumina, San Diego, CA, USA). Libraries were size-selected for a mean insert size of 300 bp on agarose gels, amplified with 10 PCR cycles, and 2x100 bp paired-end reads were sequenced on a HiSeq 2000. Three groups of populations were sequenced: four replicates of the base population ("Ancestral") and the four replicates of the ContSys and VirSys selection regimes at generation 20.

Read quality control and mapping

Reads were mapped following the previously described pipeline for pooled-sequencing analysis. Briefly, 100 bp paired-end reads were filtered for a minimum average base quality score of 18 and trimmed using PoPoolation (28). Reads with a minimum length ≥50 bp were then mapped against a reference containing the FlyBase *D. melanogaster* genome r5.38 (http://flybase.org). For details on filtering parameters and coverage, see supplementary information.

SNP calling

Only SNPs that met the following quality criteria were considered: (i) occurrence in at least two replicate populations; (ii) the minor allele was covered by at least 10 reads across all populations analysed; (iii) the maximum coverage did not exceed 500.

Genetic diversity

To characterize genome-wide patterns of genetic diversity, we estimated per site heterozygosity (π), following the PoPoolation analysis pipeline (28). We only

considered polymorphic sites for which the minor allele was supported by at least two reads after standardizing the coverage to 30 - and used unbiased estimators for pooled data that correct for pool size and coverage (28, 47). For graphical representation, we calculated average values in sliding 500-kb windows, with a step size of 100kb across the entire genome (Fig. S1A).

Identification of candidate SNPs

We used the Cochran–Mantel–Haenszel (CMH) test, as implemented in PoPoolation2 (48) to identify SNPs with changes in allele frequencies between the different regimes that were consistent among replicates as described in Orozco-terWengel et al (9) (see also Suppl. Information).

RNAi

We performed *in vivo* RNAi knockdown assays for the candidate genes in the 3L and X (pst and Ubc-E2H) and for a set of genes in the 3L peak of differentiation, selected according to whether (a) they had significantly differentiated non-synonymous SNPs or (b) gene ontology or previous functional assays suggested a role in antiviral immunity. We took advantage of the two large RNAi collections of the VDRC (49), and used the Gal80ts/Tub-Gal4 inducible system to rescue from developmental lethality. The tested constructs are shown in Table S2. More details are available as Suppl. Information.

Statistical analysis

All statistical analyses were done using R (v 2.15). Full details are provided as Suppl. Information.

Acknowledgements

The authors wish to thank Dieter Ebert and Britt Koskella for critical comments that significantly improved the manuscript, the Beldade, Mirth, and Sucena Labs and Margarida Matos for inspiring discussions on the project. NM (#SFRH/BPD/62964/2009) and VF (#SFRH/BD/82299/2011) are funded by Fundação para a Ciência e a Tecnologia (FCT, Portugal). CS is funded by the Austrian Science Funds (FWF, P22725, P19467) and the ERC (ArchAdapt). This work was supported by Fundação para a Ciência e Tecnologia (PTDC/SAU-IMU/120673/2010), Instituto Gulbenkian de Ciência/ Fundação Calouste Gulbenkian and VetMedUni Funding.

References

1. Schmid-Hempel P (2011) *Evolutionary Parasitology: The Integrated Study of Infections, Immunology, Ecology, and Genetics* (Oxford Biology, Oxford).
2. Kawecki TJ et al. (2012) Experimental evolution. *Trends Ecol Evol* 27:547–60.
3. Kassen R (2002) The experimental evolution of specialists, generalists, and the maintenance of diversity. *J Evol Biol* 15:173–190.
4. Lopes PC, Sucena E, Santos ME, Magalhães S (2008) Rapid experimental evolution of pesticide resistance in *C. elegans* entails no costs and affects the mating system. *PLoS One* 3:e3741.
5. Magalhães S, Blanchet E, Egas M, Olivieri I (2009) Are adaptation costs necessary to build up a local adaptation pattern? *BMC Evol Biol* 9:182.
6. Zhou D et al. (2011) Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 108:2349–54.
7. Burke MK et al. (2010) Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* 467:587–90.
8. Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM (2011) Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet* 7:e1001336.
9. Orozco-terWengel P et al. (2012) Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol Ecol* 21:4931–41.

10. Turner TL, Miller PM (2012) Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics* 191:633–42.
11. Ye YH, Chenoweth SF, McGraw E a (2009) Effective but costly, evolved mechanisms of defense against a virulent opportunistic pathogen in *Drosophila melanogaster*. *PLoS Pathog* 5:e1000385.
12. Kraaijeveld AR, Godfray HC (1997) Trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster*. *Nature* 389:278–80.
13. Vijendravarma RK, Kraaijeveld AR, Godfray HCJ (2009) Experimental evolution shows *Drosophila melanogaster* resistance to a microsporidian pathogen has fitness costs. *Evolution* 63:104–14.
14. Antonovics J et al. (2013) The origin of specificity by means of natural selection: evolved and nonhost resistance in host-pathogen interactions. *Evolution* 67:1–9.
15. Martins NE, Faria VG, Teixeira L, Magalhães S, Sucena E (2013) Host Adaptation Is Contingent upon the Infection Route Taken by Pathogens. *PLoS Pathog* 9:e1003601.
16. Zbinden M, Haag CR, Ebert D (2008) Experimental evolution of field populations of *Daphnia magna* in response to parasite treatment. *J Evol Biol* 21:1068–78.
17. Bennett AF, Lenski RE (2007) An experimental test of evolutionary trade-offs during temperature adaptation. *Proc Natl Acad Sci U S A* 104 Suppl :8649–54.
18. Kubinak JL, Ruff JS, Hyzer CW, Slev PR, Potts WK (2012) Experimental viral evolution to specific host MHC genotypes reveals fitness and virulence trade-offs in alternative MHC types. *Proc Natl Acad Sci U S A* 109:3422–7.

19. Vogwill T, Lagator M, Colegrave N, Neve P (2012) The experimental evolution of herbicide resistance in *Chlamydomonas reinhardtii* results in a positive correlation between fitness in the presence and absence of herbicides. *J Evol Biol* 25:1955–64.

20. Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM (2011) Successive Increases in the Resistance of *Drosophila* to Viral Infection through a Transposon Insertion Followed by a Duplication. *PLoS Genet* 7:e1002337.

21. Magwire MM et al. (2012) Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS Genet* 8:e1003057.

22. Bangham J, Knott SA, Kim K-W, Young RS, Jiggins FM (2008) Genetic variation affecting host-parasite interactions: major-effect quantitative trait loci affect the transmission of sigma virus in *Drosophila melanogaster*. *Mol Ecol* 17:3800–7.

23. Contamine D, Petitjean AM, Ashburner M (1989) Genetic resistance to viral infection: the molecular cloning of a *Drosophila* gene that restricts infection by the rhabdovirus sigma. *Genetics* 123:525–33.

24. Bangham J, Kim K-W, Webster CL, Jiggins FM (2008) Genetic variation affecting host-parasite interactions: different genes affect different aspects of sigma virus replication and transmission in *Drosophila melanogaster*. *Genetics* 178:2191–9.

25. Carpenter J a, Hadfield JD, Bangham J, Jiggins FM (2012) Specific interactions between host and parasite genotypes do not act as a constraint on the evolution of antiviral resistance in *Drosophila*. *Evolution* 66:1114–25.

26. Xu J, Cherry S (2014) Viruses and antiviral immunity in *Drosophila*. *Dev Comp Immunol* 42:67–84.

27. Teixeira L, Ferreira Á, Ashburner M (2008) The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol*. 6(12): e1000002

28. Kofler R et al. (2011) PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PLoS One* 6:e15925.

29. Burke MK (2012) How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc Biol Sci* 279:5029–38.

30. Huang W et al. (2012) Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *Proc Natl Acad Sci U S A* 109:15553–9.

31. Turner TL, Miller PM, Cochrane VA (2013) Combining genome-wide methods to investigate the genetic complexity of courtship song variation in *Drosophila melanogaster*. *Mol Biol Evol* 30:2113–20.

32. Kofler R, Schlötterer C (2014) A Guide for the Design of Evolve and Resequencing Studies. *Mol Biol Evol*. 31:474-83

33. Remolina SC, Chang PL, Leips J, Nuzhdin S V, Hughes KA (2012) Genomic basis of aging and life-history evolution in *Drosophila melanogaster*. *Evolution* 66:3390–403.

34. Tobler R et al. (2014) Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Mol Biol Evol*. 31:364-75

35. Gwynn DM, Callaghan A, Gorham J, Walters KF a, Fellowes MDE (2005) Resistance is costly: trade-offs between immunity, fecundity and survival in the pea aphid. *Proc Biol Sci* 272:1803–8.

36. Lazzaro BP, Sackton TB, Clark AG (2006) Genetic variation in *Drosophila melanogaster* resistance to infection: a comparison across bacteria. *Genetics* 174:1539–54.

37. Kraaijeveld AR, Layen SJ, Futerman PH, Godfray HCJ (2012) Lack of phenotypic and evolutionary cross-resistance against parasitoids and pathogens in *Drosophila melanogaster*. *PLoS One* 7:e53002.

38. Ostrowski EA, Rozen DE, Lenski RE (2005) Pleiotropic effects of beneficial mutations in *Escherichia coli*. *Evolution* 59:2343–52.

39. Luong LT, Polak M (2007) Costs of resistance in the *Drosophila-Macrocheles* system: a negative genetic correlation between ectoparasite resistance and reproduction. *Evolution* 61:1391–402.

40. Magalhães S, Fayard J, Janssen A, Carbonell D, Olivieri I (2007) Adaptation in a spider mite population after long-term evolution on a single host plant. *J Evol Biol* 20:2016–27.

41. Czesak ME, Fox CW, Wolf JB (2006) Experimental evolution of phenotypic plasticity: how predictive are cross-environment genetic correlations? *Am Nat* 168:323–35.

42. Thompson JN (2005) *The Geographic Mosaic of Coevolution (Interspecific Interactions)* (University Of Chicago Press).

43. Luijckx P, Ben-Ami F, Mouton L, Du Pasquier L, Ebert D (2011) Cloning of the unculturable parasite *Pasteuria ramosa* and its *Daphnia* host reveals extreme genotype-genotype interactions. *Ecol Lett* 14:125–31.

44. Luijckx P, Fienberg H, Duneau D, Ebert D (2013) A matching-allele model explains host resistance to parasites. *Curr Biol* 23:1085–8.

45. Clarke B (1976) in *Genetic Aspects of Host-Parasite Relationships*, eds Taylor AER, Muller R (Blackwell, Oxford), pp 87–103.

46. Thrall PH et al. (2012) Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecol Lett* 15:425–35.

47. Futschik A, Schlötterer C (2010) The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* 186:207–18.

48. Kofler R, Pandey RV, Schlötterer C (2011) PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27:3435–6.

49. Dietzl G et al. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448:151–6.

Figure legends

Figure 1. Evolution of increased resistance to DCV

(A) Experimental evolution trajectories of control (Control) and virus-exposed (VirSys) populations over 34 generations of experimental evolution. Circles: populations exposed to the virus; Squares: control lines. Vertical bars correspond to the standard error of the mean survival among the four selected populations (VirSys) and of the pool of Control individuals; the straight dotted line corresponds to the original mortality rate imposed on the populations (66%). **(B)** Relative DCV loads (DCV/rpl32 copies) in females, 5 days post infection, of ContSys andVirSys populations. Points represent individual measurements; horizontal lines the mean and 95% confidence intervals). **(C)** Survival after DCV infection of control and virus selected lines, with or without *Wolbachia* (solid lines/closed symbols, Wol+ or dotted lines/open symbols, Wol-, respectively).

Figure 2. Specificity of the evolved response

Hazard ratios between ContSys and VirSys populations, when exposed to DCV, Cricket Paralysis Virus (CrPV), Flock House Virus (FHV), *Pseudomonas entomophila* (P.ent) and *Enterococcus faecalis* (E.fae). Shown are the average hazard ratios of at least 2 independent experiments, done at different generations. Vertical bars correspond to the 95% confidence intervals of the estimated hazard ratios. (* - $p< 0.05$; ** - $p< 0.01$; *** $p< 0.001$)

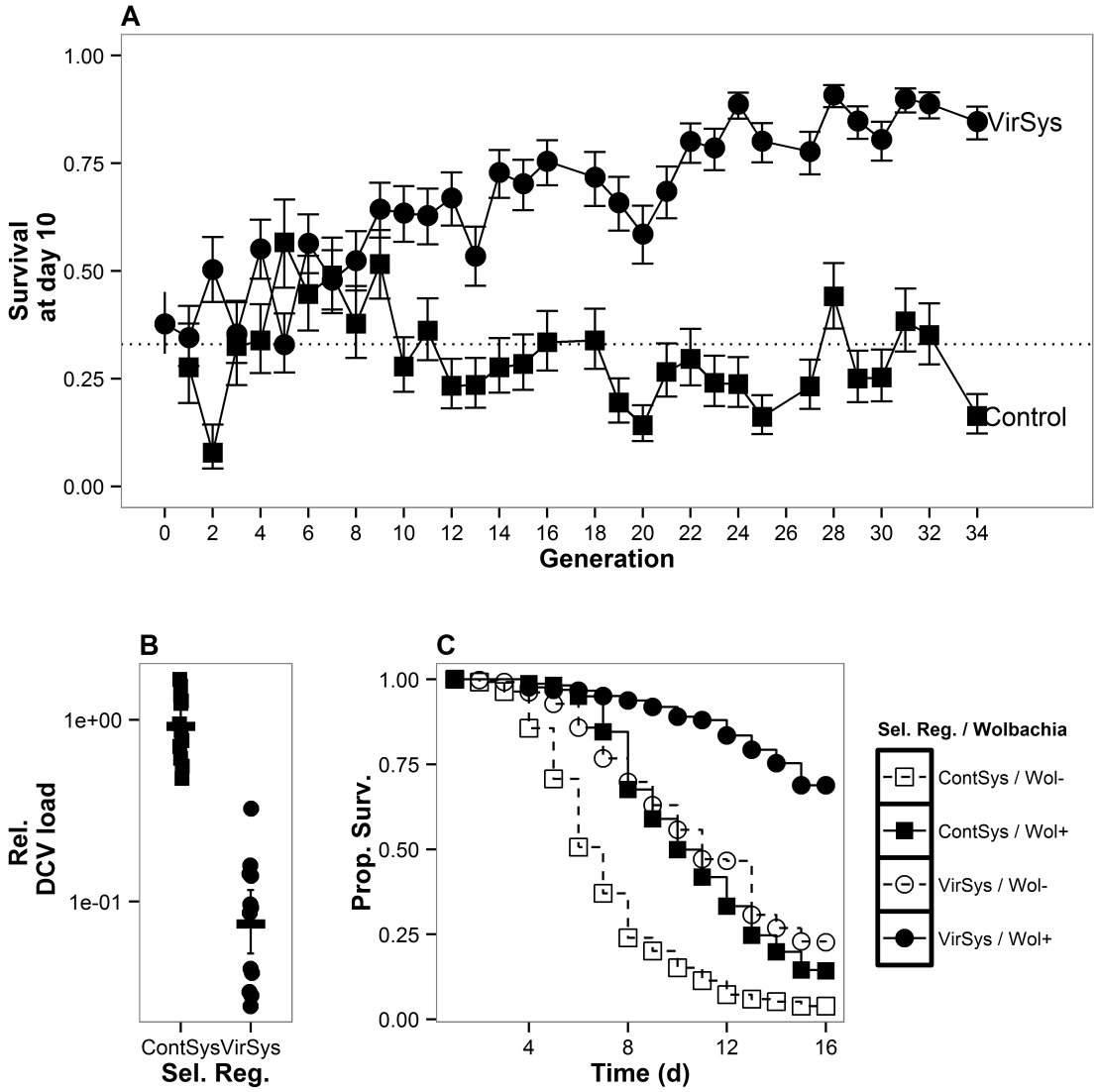
Figure 3. Differentiation between selection regimes

-log10 values of the CMH (Cochran-Mantel-Haenszel) statistic for every polymorphic SNP, across the 5 major chromosomal arms through pairwise comparison of allele frequencies between Ancestral and ContSys populations at generation 20 (top panel), Ancestral and VirSys populations at generation 20 (middle panel) and, between ContSys and VirSys at generation 20 (bottom panel). The black and red lines represent the 99.99% quantile of the p-values in the ancestral *vs* ContSys comparison at a genome wide and chromosome wide levels, respectively.

Figure 4. RNAi knockdown of candidate genes

Natural logarithm of hazard ratios between survival of flies with knocked-down candidate genes and their controls upon infection with DCV (first row), CrPV (second row) and FHV (third row), using as genetic background KK (grey bars), GD (black bars) or both, whenever a construct was available in both backgrounds. **(A)** RNAi interference against the candidate genes identified by the peaks in Figure 3, *pst* and *Ubc-E2H*. **(B)** Tests to other genes in the large 3L peak. Vertical bars correspond to the 95% confidence intervals of the estimated hazard ratios. (* - $p < 0.05$; ** - $p < 0.01$; *** $p < 0.001$)

569 Figure 1



570

571

572

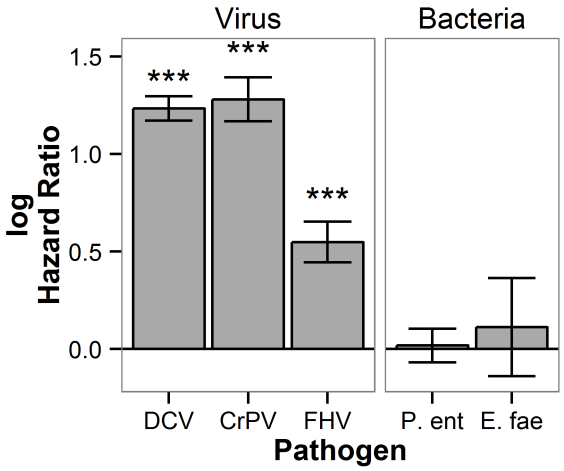
573

574

575

576

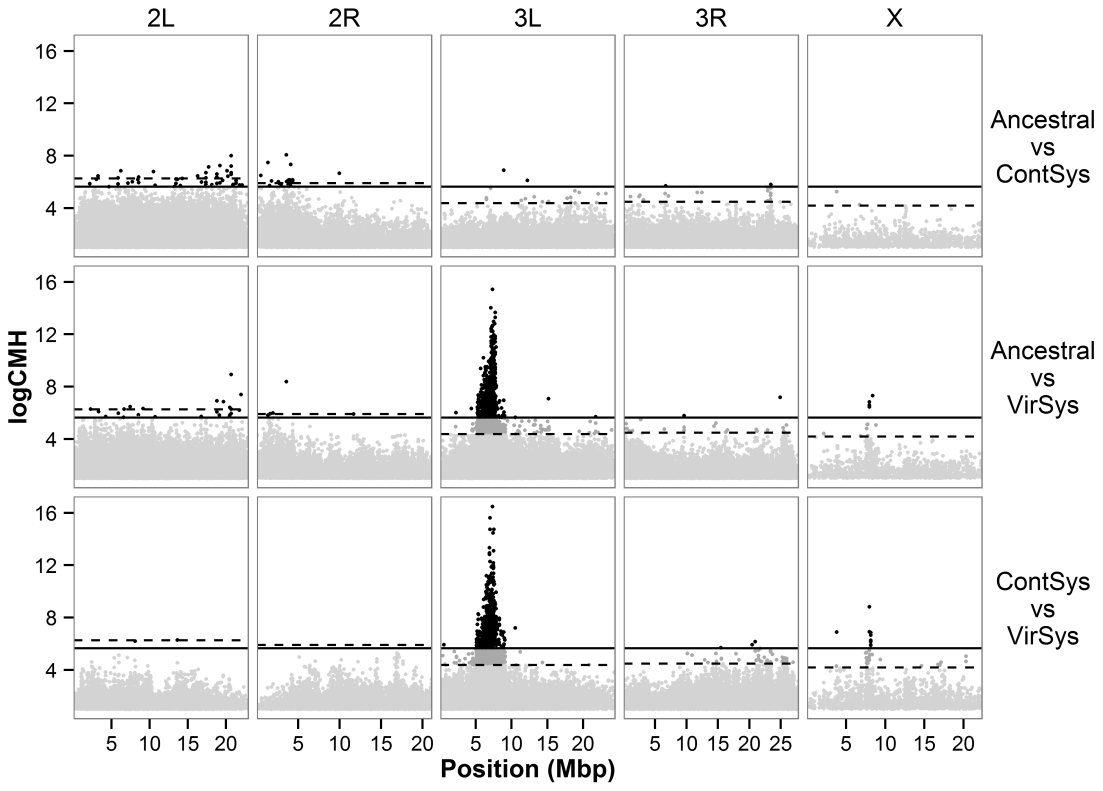
577 Figure 2



578

579

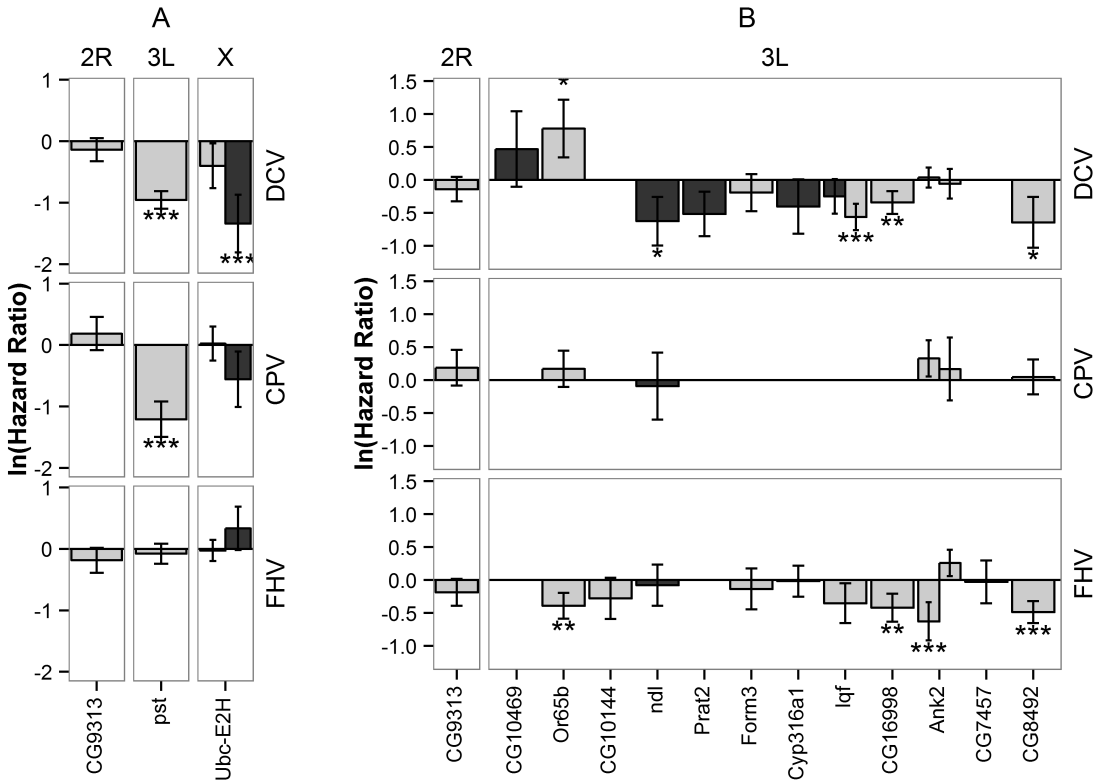
580 figure 3



581

582 Figure 4

583



584

585

586

587

588