Innate immune pathways act synergistically to constrain RNA virus evolution in Drosophila melanogaster

Vanessa Mongelli1, Sebastian Lequime2, Athanasios Kousathanas3, Valerie Gausson1, Hervé Blanc1, Jared Nigg1, Lluís Quintana-Murci3,4, Santiago F. Elena5,6 and Maria-Carla Saleh6

Host-pathogen interactions impose recurrent selective pressures that lead to constant adaptation and counter-adaptation in both competing species. Here, we sought to study this evolutionary arms-race and assessed the impact of the innate immune system on viral population diversity and evolution, using Drosophila melanogaster as model host and its natural pathogen Drosophila C virus (DCV). We isogenized eight fly genotypes generating animals defective for RNAi, Imd and Toll innate immune pathways as well as pathogen-sensing and gut renewal pathways. Wild-type or mutant flies were then orally infected with DCV and the virus was serially passaged ten times via reinfection in naive flies. Viral population diversity was studied after each viral passage by high-throughput sequencing and infection phenotypes were assessed at the beginning and at the end of the evolution experiment. We found that the absence of any of the various immune pathways studied increased viral genetic diversity while attenuating virulence. Strikingly, these effects were observed in a range of host factors described as having mainly antiviral or antibacterial functions. Together, our results indicate that the innate immune system as a whole and not specific antiviral defence pathways in isolation, generally constrains viral diversity and evolution.

Interaction between hosts and pathogens trigger defence and counter-defence mechanisms that often result in reciprocal adaptation and co-evolution of both organisms1. Empirical evidence of such arms-races involving both species can be drawn from genome-wide analysis of hosts and pathogens and in experimental evolution settings. For example, evolutionary analysis of mammalian genomes has revealed evidence of host–virus co-evolution between different retroviruses and antiviral factors12 and, in plants, host resistance genes and virulence genes encoded by pathogens have been found to co-evolve14. Likewise, between bacteria and their infecting bacteriophages, experimental co-evolution studies resulted in the occurrence of genetic variants in both a bacterial lipopolysaccharide synthesis gene and the phage tail fibre gene which binds to lipopolysaccharide during adsorption15. In nematodes and their pathogenic bacteria, the number of toxin-expressing plasmids varies during adaptation to the host16.

In insects, analyses of sequences within and between Drosophila species have shown evidence of adaptive evolution in immunity-related genes17–19. In a study that deep-sequenced small interfering RNAs (siRNAs) from mosquitoes infected with West Nile virus, it was found that the regions of the viral genome more intensively targeted by RNA interference (RNAi) contained a higher number of mutations than genomic regions less affected by this antiviral defence mechanism imposes a selective pressure on the viral population11. Similar observations on the selective pressure imposed by the RNAi pathway on viral evolution have been made in plant- and human-infecting viruses12–16. Drosophila melanogaster is a well-studied insect model to decipher virus–host interactions and therefore the impact of host antiviral immunity on viral diversity and evolution. Different Drosophila immune pathways and mechanisms are involved in antiviral defence17–19. As is the case for all invertebrates, defence against pathogens in Drosophila relies on innate immunity, which constitutes the first and only defence against microbes. Innate immunity is characterized by the recognition of pathogen-derived molecules, called pathogen-associated molecular patterns (PAMPs), by host-encoded receptors (pathogen-recognition receptors, PRRs), which leads to a rapid defence response.

The RNAi mechanism is known to play a central role in Drosophila antiviral defence, mainly through the action of the siRNA pathway19–22. Antiviral RNAi is triggered by virtually all insect-infecting viruses, resulting in targeting of the viral genome in a sequence-specific manner to control infection. Several other pathways have antiviral properties in flies but their roles against viruses seem to be virus specific. The Toll and Imd (immunedeficiency) pathways, originally described to be involved in antibacterial and antifungal responses, have been shown to play a role in antiviral defence against Drosophila C virus (DCV), Cricket paralysis virus (CrPV), Drosophila X virus, Nora virus and Flock house virus23–26. The Janus kinase signal transducers and activators of transcription (JAK–STAT) pathway can be activated on DCV or CrPV infection in flies, triggering the expression of antiviral factors27,28. DCV, a positive-sense single-stranded RNA virus from the genus Cripavirus within the Dicistrioviridae family and Picornavirales order29, is a well-characterized natural pathogen of the fruit fly that can be found in laboratory and wild populations30. As for many other Drosophila-infected viruses, defence against DCV depends on the joint action of different innate immune pathways and mechanisms.
RNAi, Toll and Imd pathways, but also the protein encoded by the gene Vago, play a role in the defence against this virus. DCV is thought to be naturally acquired by ingestion. For orally acquired pathogens, the digestive tract, and the gut in particular, represents the first host defence barrier. Despite many studies using oral bacterial infections, the role of gut-specific antiviral responses in Drosophila is not fully understood. Responses triggered against bacterial pathogens in the gut include the production of reactive oxygen species and antimicrobial peptides, as well as tissue repair and regeneration mechanisms. Furthermore, the maintenance of gut homeostasis after tissue damage caused by pathogenic bacteria relies on the activity of JAK-STAT and epidermal growth factor receptor (EGFR) pathways, amongst others. In the hallmark of viral infections, a role of the Imd and extracellular-signal-regulated kinase (ERK) pathways in the antiviral response in the gut has been suggested. It is important to note that, like many other RNA viruses with error-prone polymerases and fast replication kinetics, DCV exists as large populations composed of a cloud of genetically related mutant variants known as viral quasi-species or mutant swarms. Viral mutant swarms constitute a dynamic repertoire of genetic and phenotypic variability that renders great adaptability.

In this work, we leveraged the vast knowledge on antiviral mechanisms, the extensive genetic tool-box available for D. melanogaster, the intrinsic variability of the DCV mutant swarm and the great depth of power of next-generation sequencing (NGS) to study the impact of innate immune pathways on viral diversity and evolution. We aimed to determine not only if each pathway has a specific impact on the selective pressure imposed on DCV mutant swarm but also their relative impact. In addition, we investigated possible links between selected viral variants (viral function) and specific defence mechanisms. Our results with infections in flies defective for several immune pathways show that the host genotype has an impact on viral genetic diversity regardless of the immune pathway being affected and this is accompanied by an attenuation of the virulence along evolutionary passages. We also describe complex mutation dynamics, with several examples of clonal interference in which increases in frequency of adaptive mutations have been displaced by other mutations of stronger effect that arose in different genetic backgrounds. Overall, our results highlight that innate immune pathways constrain RNA virus evolution and further demonstrate that antiviral responses in Drosophila are probably polygenic.

### Results

**Production of fly mutant lines for innate immune pathways.** To determine the impact of the innate immune system on virus population diversity and evolution, we selected fly lines with impaired function in genes belonging to most of the Drosophila innate immune pathways: RNAi, Toll and Imd. We selected genes encoding for proteins involved both upstream and downstream of the immune pathways, such as receptors or ligands that trigger the immune response and effectors of the response (Fig. 1a): for the RNAi pathway, Dicer 2 (Dcr-2) and Argonaute 2 (AGO-2); for the Toll pathway, the ligand of Toll receptor Spätzle (spz) and the NF-kB transcription factor dorsal-related immunity factor (Dif); for the Imd pathway, the NF-kB transcription factor Relish (Rel). We also added to the study the host factor Vago, that is upregulated during viral infections in a Dicer 2-dependent manner. Because DCV is orally acquired, and to explore the impact of gut homeostasis on the antiviral response, a mutant line for epidermal growth factor receptor (Egfr), a gene involved in gut epithelium renewal, was also included in our panel. With the exception of Egfr and Dif, all of the selected genes were previously described to play an antiviral role against DCV infection. It is important to mention that, in contrast to the RNAi antiviral mechanism that relies on the direct interaction between the components of the RNAi pathway and the viral genome, the molecular mechanisms underlying the antiviral responses mediated by Toll, Imd and Vago in Drosophila remain largely unknown.

To reduce genetic variation due to differences in genetic background, mutant flies were isogenized before beginning viral evolution experiments. Homozygous loss-of-function lines for Dcr-2 (Dcr-2;w1118), AGO-2 (AGO-2;w1118), spz(spz);Dif(Dif), Rel (RelE20) and Vago (Vago;w1118) and a hypomorph mutant line for Egfr (Egfr+) were produced in the same genetic background by crossing parental lines at least ten times to w1118 flies. Infection phenotypes of the newly produced fly lines were characterized by following their survival after inoculation with DCV by intrathoracic injection (Supplementary Fig. 1a). As previously described, Dcr-2;w1118 (Dcr-2R416X/R416X), AGO-2;w1118 (AGO-2L811fsX/L811fsX) and AGO-2;w1118 mutants infected with DCV died faster than w1118 flies, as well as Vago;w1118 flies (Supplementary Fig. 1a); however, these mutants maintained the previously observed increased susceptibility to infection by Gram-positive and Gram-negative bacteria, respectively (Supplementary Fig. 1b,c). No difference in virus-induced mortality was found between w1118 and Egfr;w1118 mutant flies (Supplementary Fig. 1a). This set of isogenic mutant flies with contrasting phenotypes to DCV infection provided us with the host model system to perform the viral evolution experiment.

**Experimental DCV evolution.** To study the impact of innate immune pathways on virus population diversity and evolution, DCV from a viral stock was serially passaged (P1 to P10) in w1118 flies and in the isogenic innate immune-deficient fly lines (Fig. 1a,b). DCV population diversity was studied after each passage by NGS and DCV virulence was analysed at the beginning and at the end of the evolution experiment.

To follow viral infection during the course of the experiment, viral load (TCID50) was determined by end-point dilution and prevalence (percentage of flies positive for TCID50) was calculated for all passages in individual flies from DCV-contaminated cages. We found that for most fly genotypes and for both biological replicates, DCV infection prevailed along the ten viral passages (Extended Data Fig. 1a,b). When considering viral loads along passages, only w1118, AGO-2;L811fsX and RelE20 flies displayed significant temporal dispersion. Whether remnants of non-replicating virus remained in the fly surface was not assessed. Of note, the DCV stock was experimentally introduced to the system only once, to start the P1.

To assess the impact that fly genotype, biological replicate and viral passage has on viral loads, the log-transformed TCID50 values from each fly genotype (Extended Data Fig. 1d) were fitted to the generalized linear model (GLM) described in Methods. In short, the model incorporates fly genotype and experimental block as orthogonal factors and passage as covariable. Highly significant differences were observed in viral load among fly genotypes (test of the intercept: $\chi^2 = 146.734$, 8 d.f., $P < 0.001$) that were of very large magnitude ($\eta^2 = 0.8485$), thus confirming that DCV load strongly varied among host genotypes. A significant effect was also observed for the viral passages (test of the covariable: $\chi^2 = 5.075$, 1 d.f., $P = 0.024$), indicating overall differences in viral accumulation among passages, although the magnitude of this effect was rather small ($\eta^2 = 0.028$). Regarding second-order interactions among factors and the
Fig. 1 | Experimental design. a, Simplified scheme of *D. melanogaster* immune pathways. The siRNA pathway is triggered by virus-derived double-stranded RNA (dsRNA), recognized by Dcr-2 and cleaved into viral siRNAs, which guide the recognition and cleavage of viral RNA by Ago-2 controlling virus infection. The Toll pathway is activated when spz binds to the Toll receptor, leading to the activation of NF-κB transcription factors (for example, Dif). The Imd pathway is triggered after the recognition of microbial peptidoglycans (PGN) by PGRP-LC, ultimately leading to the activation of Rel. Toll and Imd pathways induce the expression of antimicrobial peptides to control infection. The expression of Vago is induced after infection with DCV. The EGFR pathway is triggered in the gut after bacterial damage and leads to delamination of enterocytes and renewal. Created with BioRender.com. b, Scheme of the DCV evolution experiment. To produce the DCV stock, *w^{1118}* female flies were injected with DCV from a stock produced in S2 *Drosophila* cells (S2 DCV stock), placed in cages containing fresh *Drosophila* medium, left for 3 days and then removed to place in these DCV-contaminated cages (n = 500 *w^{1118}* or immune-deficient males and females. Flies were fed ad libitum for 3 days, moved to a clean cage for 1 day and further placed into a new clean cage for 4 days, when they were harvested (DCV passage 1, P1). A new group of 500 flies was placed in contaminated cages. This procedure was repeated ten times (ten DCV passages, P1 to P10) and replicated twice (biological replicates BR1 and BR2). For each passage and fly genotype, high-throughput sequencing and viral stocks for phenotypic characterization were obtained. c, Scheme of DCV genome and the location of primers used to amplify the genome. The viral genome is composed of single-stranded positive-sense RNA and contains two open reading frames (ORFs). ORF1 encodes for the non-structural viral proteins: 1A, viral silencing suppressor; 2C, RNA helicase; VPg, viral genome-linked protein; 3C, protease; RdRp, RNA-dependent RNA polymerase; 2B and 3A, assembly of the viral replication complex. ORF2 encodes for DCV structural proteins VP1 to VP4, which constitute the viral capsid.
covariable, a significant interaction exists between fly genotype and experimental block ($\chi^2 = 27.082, 8$ d.f., $P < 0.001$) indicating that some of the differences observed in virus accumulation among host genotypes differed among biological replicates and between fly genotype and evolutionary passage ($\chi^2 = 52.511, 8$ d.f., $P < 0.001$). However, despite being statistically significant, these two effects were of very small magnitude ($\eta^2 = 0.288\%$ and $\eta^2 = 1.49\%$, respectively), casting doubts about their biological irrelevant. Likewise, the third-order interaction was statistically significant ($\chi^2 = 86.023, 8$ d.f., $P < 0.001$), suggesting that the differences in viral load among experimental blocks observed for a particular host genotype also depended on the evolutionary passages, although once again the effect could be considered as minor ($\eta^2 = 1.49\%$). Next, we evaluated whether differences exist in viral load between immune-competent ($w^{1118}$) and the different mutant fly genotypes. In all eight cases, DCV accumulated to significantly higher levels in the immune-deficient flies than in the wild-type flies ($P < 0.001$), with the smallest significance difference corresponding to viral populations replicating in $Reb^2R20$ and $Dif^2$ and the largest to those replicating in Egfr$^{1/11}$ and Dcr-2$^{R416X/R416X}$ (Extended Data Fig. 1d).

Overall, these results show that in both immune-competent ($w^{1118}$) and immune-deficient flies, DCV oral infection was maintained along passages and confirm that mutant flies are more permissive to DCV infection.

**Viral nucleotide diversity increases in the absence of a fully functional immune response.** To look into the selective pressure imposed by the *Drosophila* innate immune pathways on DCV population variation and dynamics, we analysed virus genome diversity after each passage. Half of the population of infected flies was used to sequence the full-length DCV genome by NGS (Fig. 1b,c). The viral stocks used to start the experiment, S2 DCV stock and DCV stock, were also sequenced (Methods). Analysis of the NGS data was performed using the computational pipeline Viral Variance Analysis (ViVan)\(^42\). Sequence coverage was at least 8,000 reads per position on the genome. To determine the error rate of the sequencing procedure, including library preparation, four sequencing technical replicates of the S2 DCV stock were used (Supplementary Fig. 2). An allele frequency threshold of 0.0028 was used for all subsequent analyses based on variant detection and frequency correlation between technical replicates (Methods). We next calculated the site-averaged nucleotide diversity ($\pi$) on all polymorphic sites ($n = 1,869$) across the full-length viral genome and present in the full dataset (Fig. 2), with the aim of determining if the lack of activity of a given innate immune pathway had an impact on viral population genetic diversity, in terms of size of the viral mutant swarm.

First, we asked if there was any difference in DCV population diversity and dynamics between the different fly genotypes along the complete evolution experiment. To answer this question, we analysed if the host genotype, viral passages, biological replicate or the interactions between these factors had an impact on the evolution of viral population diversity, considering the full-length DCV genome, across all passages. We found that only the fly genotype had a statistically significant impact on $\pi$ ($\chi^2 = 25.545, 8$ d.f., $P = 0.001$) (Table 1). We then compared the DCV population diversity present in each fly genotype to each other. We found that, except for viral diversity found in Dcr-2$^{R416X/R416X}$ and $Dif^2$ lines, for which no difference was found compared to $\pi$ in $w^{1118}$ flies ($P \geq 0.303$), DCV population diversity significantly differed from $w^{1118}$ line in the rest of the innate immune mutants analysed ($P \leq 0.013$) (Supplementary Table 1).

A post hoc Bonferroni test further sorted overlapping groups according to their increasing viral nucleotide diversity; group 1 (less diversity)—$w^{1118}$, Dcr-2$^{A141/141A141}$ and $Dif^2$ flies; group 2—$Dif^2$, Dcr-2$^{A141/141A141}$, $Reb^2R20$, sp$^{z2/z}$ and Dcr-2$^{A1414X/R414X}$ flies; group 3—Dcr-2$^{A141/141A141}$, $Reb^2R20$, sp$^{z2/z}$ and Dcr-2$^{A1414X/R414X}$ and Ago-2$^{1414/1414}$ flies; group 4 (more diversity)—containing sp$^{z2/z}$, Dcr-2$^{A1414X/R414X}$, Ago-2$^{1414/1414}$, Egfr$^{1/11}$ and Vago$^{1/11}$ fly lines (Extended Data Fig. 2 and Supplementary Table 1).

Next, we wondered if the general differences observed in viral nucleotide diversity, between fly genotypes were associated with a particular viral genomic region (that is, if a determined viral function was affected during the evolution experiment) (Fig. 1c). Of note, the intergenic region internal ribosome entry site (IGR IRES) was not included in the analysis because its lack of genetic variation prevented us from determining its nucleotide diversity value. We found that the fly genotype had a statistically significant effect on the nucleotide diversity found in each DCV genomic region ($\chi^2 = 27.178, 8$ d.f., $P < 0.001$), which further differed between each specific viral genomic region ($\chi^2 = 11.698, 8$ d.f., $P = 0.008$). As a second-order interaction, an effect of the fly genotype and the biological replicate was found ($\chi^2 = 16.314, 8$ d.f., $P = 0.038$) (Table 1). Comparison of viral genetic diversity within the genomic regions allowed us to distinguish three main groups: group 1 (less diversity), 3'UTR; group 2, 5'UTR IRES; and group 3 (more diversity), ORF1 and ORF2 (Extended Data Fig. 2 and Supplementary Table 1).

Finally, we wondered if viral diversity evolved from the starting viral stock (DCV stock) in each fly genotype. The $\pi$ present in P1, P5 and P10 was compared between fly genotypes and with the diversity present in the DCV stock. We found that pairwise comparisons of viral nucleotide diversity present in each fly genotype in P1, between each other and versus DCV stock, yield no statistically significant difference ($P = 1.000$) (Supplementary Table 1). In P5 viral diversity was reduced only in $w^{1118}$ (group 1/2; $P = 0.026$ and $P = 0.032$) compared to the starting viral stock (Extended Data Fig. 2 and Supplementary Table 1). In P10, viral nucleotide diversity present in $w^{1118}$ (group 1, $P = 0.032$ and $P = 0.041$), sp$^{z2/z}$ (group 1, $P = 0.020$ and $P = 0.025$), $Dif^2$ (group 1, $P = 0.005$ and $P = 0.006$) and $Reb^2R20$ (group 1/2, $P = 0.046$) mutant flies was reduced when compared to DCV diversity from the DCV stock (Extended Data Fig. 2 and Supplementary Table 1).

Altogether, the results show that the absence of a fully functional immune system results in an increase of viral population diversity that remains constant along passages. They also show that the coding regions of the virus are more prone to accumulate variation than the non-coding regions where regulatory elements are present.

**Viral population diversity derives from pre-existing standing genetic variation.** Next, we examined if the levels of viral diversity observed in DCV populations from innate immune mutants compared to the $w^{1118}$ line were accompanied with the fixation of particular genetic changes in the mutant swarms and whether (1) these changes can be associated with fitness effects, (2) potentially adaptive mutations arose in response to particular immune responses. To do so, we estimated the selection coefficients for each single nucleotide polymorphism (SNP) using their variation in frequency across evolutionary time (Fig. 3 and Extended Data Fig. 3), using a classic population genetics approach\(^43\) (Table 2). Thirty-six SNPs yielded significant estimates of selection coefficients (this number reduces to ten if a stricter false discovery rate (FDR) correction is applied; Table 2). Twenty-one of them were already detected in the ancestral S2 DCV stock, hence a maximum of 15 new SNPs might have arisen during the evolution experiment. Estimated selection coefficients for all these SNPs ranged between $-0.304$ per passage (synonymous mutation RdRp/C5713U) and $1.204$ per passage (VP2/G6311C non-synonymous change R16P), with a median value of 0.286 per passage (interquartile rank = 0.265). Nineteen mutations were observed in more than one lineage (range 2–7 lineages), with synonymous mutations VP3/U7824C appearing in seven lineages of six different host genotypes and mutation 5'UTR/A280U in five lineages of five host genotypes (Table 2). These nine SNPs were all present in the S2 DCV stock. Indeed, the frequency of SNPs among evolving lineages is significantly correlated with their frequency in the ancestral
S2 DCV stock (Pearson’s $r=0.401$, 36 d.f., $P=0.013$) but not with their measured fitness effect ($r=-0.091$, 36 d.f., $P=0.588$).

An interesting question is whether the fitness effects associated with each of these nine SNPs were the same across all genotypes or, conversely, whether fitness effects were host genotype-dependent. To test this hypothesis, we performed one-way analysis of variance (ANOVA) tests comparing fitness effects (Table 2) across the corresponding host genotypes. In all cases, significant differences were observed ($F>15.637$ and $P<0.001$ and $\geq 93.99\%$ of total observed variance in fitness effects explained by true genetic differences among host genotypes), supporting the notion that fitness effects are indeed host genotype-dependent. A pertinent example is the case of the synonymous mutation VP3/U7824C, which was the most prevalent mutation ($F_{1,15}=158.862$, $P<0.001$, 99.37% of genetic variance). In this case, a post hoc Bonferroni test shows that host genotypes can be classified into three groups according to the estimated fitness effect of this SNP. In genotypes $Dcr-2^{R416X/R416X}$, $Ago-2^{414/414}$, the mutation has a deleterious effect (on average, $-0.2260$ per passage); in genotypes $Egf^{R20/R20}$ and $Vago^{A310D/A310D}$, the mutation is moderately beneficial (on average, 0.1257 per passage); and in genotypes $w^{118}$ and $Ago-2^{414/414}$, the mutation had a strong beneficial effect (on average, 0.502 per passage).

As shown in Fig. 3 and Extended Data Fig. 3a, some SNPs show a strong parallelism in their temporal dynamics, suggesting that they might be linked into haplotypes. This is particularly relevant for mutations shown in Table 2. To test this possibility, we computed all pairwise Pearson correlation coefficients between mutation frequencies along evolutionary time. The results of these analyses are shown in Extended Data Fig. 3b–k as heatmaps. Again, as an illustrative example, we discuss here the case of the viral population BR2 evolved in $Ago-2^{414/414}$ (Extended Data Fig. 3d). Synonymous mutations VP3/U7824C and VP1/C8424U and non-synonymous mutation VP1/C8277U (H655Y) are all linked into the same haplotype ($r=0.998, P<0.001$). Since these three mutations already existed in the S2 DCV stock, it is conceivable that the haplotype already existed and has been selected as a unit. Indeed, the fitness effects estimated for these three mutations are indistinguishable (one-way ANOVA: $F_{1,15}=1.781$, $P=0.192$; average fitness effect 0.590 ± 0.032 per passage), thus suggesting that the estimated value corresponds to the haplotype as a unit. The absence of this haplotype in $Ago-2^{414/414}$ BR1 suggests that it was lost during the transmission bottleneck from S2 cells to flies. Interestingly, mutations VP1/C8424U and VP1/C8277U appear also linked into the same haplotype in population BR2 evolved in $Dcr-2^{R416X/R416X}$ (Extended Data Fig. 3b). These two cases, as well as populations BR1 evolved in $Rel^{E20/E20}$, BR2 evolved in $spz^{2/2}$ and BR1 and BR2 evolved in $Vago^{A310D/A310D}$, illustrate some examples of haplotypes (Extended Data Fig. 3e,f,h,i). Other viral populations, especially those evolved in $Egf^{R20/R20}$ flies, show much more complex patterns (Extended Data Fig. 3j,k) in which haplotypes change over time by acquiring de novo mutations.

When mapping the 36 SNPs found to have significant estimates of selection coefficients in the viral genome (Table 2 and Extended Data Fig. 4), we found that two mapped to the 5′UTR IRES, 12 to ORF1, one to the IGR IRES, 20 to ORF2 and one to the 3′UTR. Of the 12 mutations observed in ORF1, which encodes the non-structural proteins, four mapped to the 3C viral protease and
five to the RdRp. Only one of these mutations in the 3C viral protease was non-synonymous. Of the 20 mutations in ORF2, which encodes the viral structural proteins, eight mapped to VP2, five to VP3 and seven to VP1. These correspond to the three major predicted DCV capsid proteins.

Taken together, these results show that viral population diversity over these ten in vivo passages mainly derived from pre-existing standing genetic variation in the ancestral DCV population. Furthermore, temporal dynamics of population diversity were linked to the fly genotype in which the virus evolved.

DCV virulence decreases along passages in the absence of immune pathways. Finally, we wondered if DCV virulence varied among each lineage in the different fly genotypes. Infectious DCV stocks were produced from viral passages P1 and P10 and from all fly genotypes. Because the viral evolution experiment was performed by DCV orofecal transmission, we first evaluated DCV viral population diversity by feeding flies with DCV stocks derived from P10 since their median survival time was longer than those inoculated from stocks from P1 for most DCV stocks (Fig. 4a and Supplementary Table 2). Notable exceptions were DCV stocks from BR2 of Vago_pspz2/2 mutant flies, for which w1118 flies were more sensitive to P10 than to P1 and stocks from BR1 of spz2/2 and BR2 of EgfpM10 mutant flies, for which no difference in median survival time after infection with DCV between P1 and P10 was detected.

A fundamental question in evolutionary biology is the role that past evolutionary events may have in the outcome of evolution41. If ongoing evolution is strongly contingent with past evolutionary events, ancestral phenotypic differences should be retained to some extent, while if other evolutionary forces such as selection and stochastic events (mutation and genetic drift) dominate, then ancestral differences can be eroded and, in the extreme case, even fully removed. Here, we observed significant differences in the performance of the ancestral DCV across the eight host genotypes. To test whether these differences are still observable in the evolved population, we compared the median survival time (Fig. 4a and Supplementary Table 2) for DCV populations isolated at the beginning of the evolution experiment P1 and at the end P10 (Fig. 4b). Under the null hypothesis of strong historical contingency, it is expected that data will fit to a regression line of slope 1 and intercepting the ordinate axis at 0. However, if ancestral differences have been removed, data would fit significantly better to a regression line with a slope <1 and with an intercept >0 (ref. 16). Figure 4b shows the data and their fit to the null hypothesis (solid black line) and the alternative hypothesis (dashed red line). A partial F-test shows that adding an intercept to the regression equation significantly improves the fit (F1,16 = 28.437, P = 0.001), thus supporting the notion that ancestral differences among host genotypes have been removed by the action of subsequent adaptation, that is, the fixation of beneficial mutations.

Discussion

In this work we aimed at determining the overall impact of innate immunity on viral evolution. On the basis of the arms-race hypothesis, we speculated that if a given host defence mechanism imposes a specific selective pressure on a particular pathogen function, the absence of this defence mechanism would result in the relaxation of the selective constraint, which would in turn be detectable in the pathogen at the genomic and phenotypic levels. We found that viral population diversity evolved differently according to each fly genotype; however, viral population diversity mostly derives from ancestral standing genetic variation (that is, few 'new' mutations were selected). Our results further confirm the polygenic nature of antiviral responses; there is not a specific, main immune defence mechanism against a particular virus but instead a repertoire of defence mechanisms that are triggered after infection and that might interact with each other.

Our results are compatible with a pervasive presence of clonal interference. In the absence of sexual reproduction, clonal interference is the process by which beneficial alleles originated in different clades within a population compete with each other, resulting in one of them reaching fixation. Subsequently, the outcompeted beneficial allele may appear in the new dominant genetic background and, assuming no negative epistasis among both loci, become fixed. As a consequence, beneficial mutations may fix sequentially, thus slowing down the rate of adaptation44. Given their large effective population size and high mutation rates, viral populations are expected to contain considerable amounts of potentially beneficial standing genetic variation, making them prone to clonal interference. Indeed, it has

| Experimental variable | χ² | d.f. | P      |
|-----------------------|----|-----|--------|
| Full-length genome     |    |     |        |
| BR                    | 2.2528 | 1 | 0.1334 |
| VP                    | 1.6460 | 1 | 0.1995 |
| FG                    | 25.5447 | 8 | 0.0013 ** |
| (BR) × VP             | 0.0024 | 1 | 0.9606 |
| (BR) × FG             | 14.2963 | 8 | 0.0744 |
| VP × FG               | 12.1679 | 8 | 0.1439 |
| (BR) × VP × FG        | 10.4253 | 8 | 0.2364 |
| Each DCV genomic region |    |     |        |
| BR                    | 1.2107 | 1 | 0.2712 |

The site-averaged nucleotide diversity (π) on the evolution of DCV nucleotide diversity (mean log₁₀(π) per site).

Table 1 | Analysis of the impact of each experimental variable on the evolution of DCV nucleotide diversity (mean log₁₀(π) per site)
been previously shown to operate in experimental populations of vesicular stomatitis virus adapting to cell cultures in bacteriophage ϕX174 populations adapting to harsh saline environments, in tobacco etch virus adapting to new plant host species, among HIV-1 escape variants within individual patients and also at the epidemiological level among influenza A virus lineages diversifying antigenically. In our own results, clonal interference can be observed in populations BR1 evolved in Dcr-2 L811fsX/L811fsX, BR1 evolved in Ago-2 414/414, BR1 evolved in spz2/2, BR2 evolved in Rel E20/E20 and BR2 evolved in VagoΔM10/ΔM10. All of these viral populations share similar patterns in which some beneficial allele (or haplotype) rose in frequency, reached a peak at some intermediate passage, then declined in frequency and were finally outcompeted by a different beneficial mutation (or haplotype) that had lower initial frequency. For example, the non-synonymous mutation VP2/G6931A (A223T) appeared de novo in population BR1 evolved in spz2/2 and outcompeted several mutations probably linked in a haplotype (Fig. 3). Tightly linked to clonal interference is the concept of leap-frogging, in which the beneficial mutation that ends up dominating the population is less genetically related to the previously dominant haplotype than to the common ancestor of both (Fig. 3). The VP2/G6931A mutation illustrates this case well, as it

Fig. 3 | Trajectories of DCV variants across passages. Muller plots illustrating the dynamics of frequencies of SNPs along evolutionary time. Each colour represents the dynamics of a different SNP.
## Table 2 | Mutations for which significant estimates of fitness effects have been obtained

| Fly genotype | Biological replicates | Mutation | Standing variation (frequency) | Selection coefficient per passage (± s.e.m.) | P  |
|--------------|-----------------------|----------|--------------------------------|-----------------------------------------------|----|
| **w**1118    | 1                     | VP2/G6311C R16P | Yes (0.0104)                  | 1.2039 ± 0.2543                              | 0.0418 |
| **w**1118    | 2                     | VP3/U7824C | Yes (0.1457)                  | 0.4780 ± 0.0617                              | <0.0001* |
| **Dcr-2**1885/1885 | 1                   | RPd/U5302C | No                             | 0.3877 ± 0.0973                              | 0.0073 |
|              | 2                     | VP1/C8227U H655Y | Yes (0.0147)                  | 0.3735 ± 0.1368                              | 0.0258 |
| **Dcr-2**1885/1885 | 1                   | VP1/C8424U | Yes (0.0139)                  | 0.3880 ± 0.1407                              | 0.0248 |
| **Dcr-2**1885/1885 | 2                   | VP2/C6932U A223V | Yes (0.0084)                  | 0.2135 ± 0.0169                              | <0.0001* |
| **Dcr-2**1885/1885 | 1                   | VP2/G6379A A39T | Yes (0.0098)                  | 0.2074 ± 0.0555                              | 0.0057 |
| **Dcr-2**1885/1885 | 2                   | VP3/A7465G I401V | Yes (0.0088)                  | 0.1185 ± 0.0338                              | 0.0100 |
| **Ago-2**414/414 | 1                   | VP3/U7824C | Yes (0.1457)                  | −0.2887 ± 0.0884                             | 0.0309 |
| **Ago-2**414/414 | 2                     | VP3/U7824C | Yes (0.1457)                  | 0.6238 ± 0.1077                              | 0.0007* |
| **Sproze**2/2  | 1                     | VP3/U7824C | Yes (0.1457)                  | −0.2804 ± 0.0206                             | 0.0467 |
| **Dif**1/1    | 1                     | VP3/U7824C | Yes (0.1457)                  | −0.0917 ± 0.0277                             | 0.0130 |
| **Dif**1/1    | 2                     | VP3/U7824C | Yes (0.1457)                  | −0.2804 ± 0.0206                             | 0.0467 |
| **Rel**20/E20 | 1                     | VP1/C8227U H655Y | Yes (0.0147)                  | 0.1368 ± 0.0553                              | 0.0426 |
|              | 2                     | VP1/C8424U | Yes (0.0139)                  | 0.1915 ± 0.0283                              | 0.0001* |
| **Rel**20/E20 | 2                     | VP2/C6932U A223V | Yes (0.0084)                  | 0.1936 ± 0.0291                              | 0.0002* |
| **Rel**20/E20 | 1                     | VP3/U7824C | Yes (0.1457)                  | 0.1368 ± 0.0553                              | 0.0426 |
| **Rel**20/E20 | 2                     | VP3/U7824C | Yes (0.1457)                  | −0.2804 ± 0.0206                             | 0.0467 |
| **Vago**ΔM10/ΔM10 | 1                   | VP2/G6311C R16P | Yes (0.0104)                  | 1.2039 ± 0.2543                              | 0.0418 |
|              | 2                     | VP3/U7824C | Yes (0.1457)                  | 0.4780 ± 0.0617                              | <0.0001* |
| **Vago**ΔM10/ΔM10 | 1                   | RPd/U5302C | No                             | 0.3877 ± 0.0973                              | 0.0073 |
|              | 2                     | VP1/C8227U H655Y | Yes (0.0147)                  | 0.3735 ± 0.1368                              | 0.0258 |
| **Vago**ΔM10/ΔM10 | 1                   | VP1/C8424U | Yes (0.0139)                  | 0.3880 ± 0.1407                              | 0.0248 |
| **Vago**ΔM10/ΔM10 | 2                   | VP2/C6932U A223V | Yes (0.0084)                  | 0.2135 ± 0.0169                              | <0.0001* |
| **Vago**ΔM10/ΔM10 | 1                   | VP2/G6379A A39T | Yes (0.0098)                  | 0.2074 ± 0.0555                              | 0.0057 |
| **Vago**ΔM10/ΔM10 | 2                   | VP3/A7465G I401V | Yes (0.0088)                  | 0.1185 ± 0.0338                              | 0.0100 |
| **Vago**ΔM10/ΔM10 | 1                   | VP3/U7824C | Yes (0.1457)                  | −0.2887 ± 0.0884                             | 0.0309 |
| **Vago**ΔM10/ΔM10 | 2                   | VP3/U7824C | Yes (0.1457)                  | 0.6238 ± 0.1077                              | 0.0007* |
| **Vago**ΔM10/ΔM10 | 1                   | VP3/U7824C | Yes (0.1457)                  | −0.2804 ± 0.0206                             | 0.0467 |
| **Vago**ΔM10/ΔM10 | 2                   | VP3/U7824C | Yes (0.1457)                  | 0.6238 ± 0.1077                              | 0.0007* |
| **Vago**ΔM10/ΔM10 | 1                   | VP3/U7824C | Yes (0.1457)                  | −0.2887 ± 0.0884                             | 0.0309 |

Continued.
Figure 4 | DCV virulence decreases in the absence of immune pathways. DCV infectious stocks were prepared from viral passages P1 and P10 and from each fly genotype. The w^{1118} flies were intrathoracically inoculated with ten TCID_{50} units of each DCV stock and survival of the flies was measured daily. a. Survival curves shown in the figure are the combination of the two independent replicates, with three technical replicates each, of a total of at least n = 98 flies per treatment. Error bars indicate ± s.e.m.; NS, not significant. Survival curves were compared via log-rank (Mantel–Cox) tests. b. Test of the contribution of historical contingency evolved (P10) versus ancestral (P1) DCV virulence. The dashed red line represents the linear regression and the black line represents the expected relationship under the null hypothesis of ancestral differences in DCV virulence which are maintained after evolution despite noise introduced by random events (mutation and drift).
of immune response (relaxation factor). Because DCV replication is significantly increased in immune-deficient mutants, the potential for population diversification is higher. This effect is clearly observed in w1118 flies where the virus is only adapting to the new environment and DCV populations evolved in w1118 flies show less variation than all other lineages. Future experimental evolution studies using viral stocks derived from flies, instead of cell cultures, are warranted to address this topic.

In a study published recently9, Navarro et al. used Arabidopsis thaliana and turnip mosaic virus to carry out experimental virus evolution assays with a similar design to ours. In their work, the authors used plant mutants compromised in their antiviral response (more permissive to viral infection) or with an enhanced antiviral response (less permissive to viral infection) and allowed the virus to evolve for 12 passages. Similarly to what we found in the D. melanogaster–DCV system, the authors showed that viral population evolution dynamics, as well as viral loads, depend on host genotype. Interestingly, a reduction of ancestral genetic variation regardless of the immune pathway affected was also clearly observed, in agreement with our observations.

Taken together, this result points to the concerted action of the different immune pathways to limit viral evolution. Response to infection does not simply consist of activating immune pathways, it also encompasses a broad range of physiological consequences including metabolic adaptations, stress responses and tissue repair. Critically, on infection, the homeostatic regulation of these pathways is altered. However, such alterations do not always result in increased disease severity and in fact can even lead to improved survival (or health) despite active virus replication.

Methods

Fly strains and husbandry. Flies were maintained on a standard cornmeal diet (Bloomington) at a constant temperature of 25 °C. All fly lines were cleaned of possible chronic infections (viruses and Wolbachia) as described previously25. The presence or absence of these chronic infections was determined by PCR with reverse transcription with specific primers for Noria virus, Drosophila A virus, DCV (NoraV for ATGCCAGGTAGTCGAACGT; NoraV for CTGTGGTTCCAGTTGGGTTCGA; Dav0V for AGAAGGCTTGTTGAGGACAT; DavV for GCCATATGCAACACAGCTTG; DcV for GTTGCCTATCTGTGCTCTG; DcV for CGCAATACGTGCTTCTCTG) and by PCR with specific primers for Wolbachia sp. (wspF for TGTTGCAAAATAGTGAAGAAAC, wsp4R for AAAAATGCACAG, Dif1R for ATCGTGGTCTCCTGTGTGACG, Rel_Ex4R for AGCTCTCTCACGCTA and wsp45R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG, Vagod10R for TTGGCCAACGGAAAGGATGTG,
instructions. A forward primer containing the non-target tag sequence (Tag, qPCR-F: AATTCAGCCTGTTCTCTCCTG) and a dCv- specific primer reverse (dCv_qPCR-R: AATGCGAAGGCACACACTATA) were used for qPCR.

RNA extraction, cDNA synthesis and NGS library production. To produce the NGS libraries from the evolution experiment, half of the total population of flies infected with dCv from each fly genotype, viral passage and biological replicates (=250 flies) was used. To produce the NGS libraries from the viral stock from S2 cells (S2 dCv stock), two different aliquots of the stocks were used. To produce the NGS libraries from the dCv stock (virus infecting w1118 female flies used to contaminate the cages to start the evolution experiment), half of the population of the infected flies (=800 flies: female flies used to contaminate the cages to start the evolution experiment), at 98 °C for 30 s; 30 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C with specific primers (dCvFor ATATGTACACACGGCTTTTAGGT and dCvRev primers reverse transcription with the Maxima H Minus Reverse Transcriptase. Then 300 ng of total RNA were used to produce the cDNA using oligo(dT) as primer against the poly-A tail. The resulting first-strand cDNA was purified using the PureLink RNA Kit (Thermo Fisher Scientific) and its concentration was determined using a NanoDrop ND-1000 Spectrophotometer. A total 200 ng of the purified RNA were used to produce the cDNA library using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs), according to manufacturer’s instructions. Then 100 ng of total RNA were used to produce the cDNA library using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs), according to manufacturer’s instructions. The quality of the libraries was verified using a High Sensitivity DNA Chip (Agilent) and the quality of the libraries was verified using a High Sensitivity DNA Chip (Agilent) and quantified using the Quanti-it RNA assay kit (Thermo Fisher Scientific). A 1% NDM dilution of the libraries was used for the sequencing that was performed on a NextSeq sequencer (Illumina) with a NextSeq 500 Mid Output kit v.2 (Illumina) (151 cycles). Two of the four technical replicates for S2 dCv stock and dCv stock were included in each run.

Sequencing of dCv populations from Dif1′1′ mutant flies from P4 to P6 from BR1 and P8 from BR2 did not work.

Genetic diversity analyses. Variant frequency threshold. To determine the error rate of the sequencing procedure, including library preparation, four sequencing technical replicates from S2 dCv stock were used (Supplementary Fig. 3a). First, pairwise comparison was done to identify the variant frequency threshold above which at least 95% of the variants were detected in both considered replicates (high-confidence detection threshold: 0.0028). All variants above detection threshold were then correlated between each technical replicate to ensure good correlation between reported frequency values: the Pearson correlation coefficient between the detected frequency for variants was ≥0.982 for all pairwise correlation (P < 0.001). The R packages used for these analysis were described elsewhere.

Nucleotide diversity () . Nucleotide diversity of the viral population was computed using the following formula:

\[ \pi = \frac{1}{2N} \sum_{i=1}^{n} \left( 1 - \left( p_i^2 + (1 - p_i)^2 \right) \right) \]

with \( D \) the sequencing depth and \( p \) the frequency of the minority variant at each nucleotide site. For diallelic SNV, the selection coefficient of the minority allele is given by:

\[ s = \frac{1}{2} \sqrt{1 - 2 \lambda} \]

with \( \lambda \) the fitness cost of the minority allele. In the subsequent analyses, \( \lambda \) was averaged over all polymorphic nucleotide sites of the DCV genome of each sample. A site was considered polymorphic if at least one sample showed the presence of a nucleotide variant at said position of the DCV genome. The log10-transformed site-averaged \( \pi \) values were then compared between fly genotypes (orthogonal factor), biological replicates (orthogonal factor), passages (continuous variable) and genomic regions (orthogonal factor) and their interactions using a GLM. The significance of each term in the model was evaluated using a likelihood ratio test that follows a \( \chi^2 \) distribution probability.

Estimation of relative mutational fitness effects. We have followed the classic population genetics method described in Hartl and Clark. In short, let \( x(i) \) be the frequency of a mutant allele (SNP) at genomic position \( i \) and passage \( t \) and, therefore, \( 1 - x(i) \) the frequency of the wild-type allele. It holds that \( \log_{10} = \log_{10} + \log \left( 1 - x(i) \right) \), where \( s \) is the selection coefficient of the mutant relative to the wild-type allele at locus \( i \). Selection coefficients calculated this way have units of inverse time (per passage in our case). This equation was fitted to the time-series data of each locus \( i \) shown in Fig. 3 by least squares regression, obtaining an estimate of \( s \) and its s.e.m.

Haplotype inference was done using two different statistical approaches. First, by assessing the similarity between temporal dynamics of all possible pairs of loci.

To this end, Pearson partial correlation coefficients (controlling for passages) were computed and their significance level corrected for multiple tests of the same null hypothesis using Benjamini and Hochberg FDR method. Correlation coefficient matrices were visualized as heatmaps in which more similar alleles were clustered together. Second, we confirmed the results from the first method using the longitudinal variant allele frequency factorization problem (LVAFFP) method as implemented in CALDER). LVAFFP generates spanning trees of a directed graph constructed from the variant allele frequencies. The output of CALDER was used as input of ‘TimeScape’ to generate the Muller plots that illustrate the aneuploidy of mutations and haplotypes along the evolution experiment (Fig. 3).

Statistical analyses described in this section have been done with R v.4.0.2 in RStudio v.1.3.1073.

Data availability
All raw data from high-throughput sequencing were deposited to NCBI BioProjects under accession number PRJNA782868. Source data are provided with this paper.

Code availability
Scripts are provided in Supplementary Data 1.

Received: 19 May 2021; Accepted: 14 December 2021; Published online: 10 March 2022

References
1. Morgan, A. D. & Koskella, B. in Genetics and Evolution of Infectious Diseases (ed. Taberyrenc, M.) 115–140 (Elsevier, 2017).
2. Daugherty, M. D. & Malik, H. S. Rules of engagement: molecular insights from host–virus arms races. Annu. Rev. Genet. 46, 677–700 (2012).
3. Barreiro, L. B. & Quintana-Murci, L. From evolutionary genetics to human immunology: how selection shapes host defence genes. Nat. Rev. Genet. 11, 17–30 (2010).
4. Thompson, J. N. & Burdon, J. J. Gene-for-gene coevolution between plants and parasites. Nature 360, 121–125 (1992).
5. Buckling, A. & Rainey, P. B. Antagonistic coevolution between a bacterium and a bacteriophage. Proc. Biol. Sci. 269, 931–936 (2002).
6. Masi, L. et al. Host–pathogen coevolution: the selective advantage of Bacillus thuringiensis virulence and its Cry toxins genes. PLoS Biol. 13, e1002169 (2015).
7. Obbard, D. J., Gordon, K. H. J., Buck, A. H. & Jiggins, F. M. The evolution of RNAi as a defence against viruses and transposable elements. Philos. Trans. R. Soc. B 364, 99–115 (2009).
8. Lazzaro, B. P. Natural selection on the Drosophila antiviral immune system. Cur. Opin. Microbiol. 11, 284–289 (2008).
9. Lazzaro, B. P. Molecular population genetics of inducible antibacterial peptide genes in Drosophila melanogaster. Mol. Biol. Evol. 20, 914–923 (2003).
10. Sackton, T. B. et al. Dynamic evolution of the innate immune system in Drosophila. Nat. Genet. 39, 1168–1172 (2007).
11. Brackney, D. E., Beane, J. E. & Ebel, G. D. RNAi targeting of West Nile virus in mosquito midguts promotes virus diversification. PLoS Pathog. 5, e1000502 (2009).
12. Lin, S.-S. et al. Molecular evolution of a viral non-coding sequence under the selective pressure of amiRNA-mediated silencing. PLoS Pathog. 5, e1003012 (2009).
13. Lafourge, G. et al. Tempo and mode of plant RNA virus escape from RNA interference-mediated resistance. J. Virol. 85, 9686–9695 (2011).
14. Martinez, F. E. et al. Ultra-deep sequencing analysis of population dynamics of virus escape mutants in RNAi-mediated resistant plants. Mol. Biol. Evol. 29, 1327–1337 (2012).
15. Das, A. T. et al. Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. J. Virol. 78, 2601–2605 (2004).
16. Gitlin, L., Stone, J. K. & Andino, R. Poliovirus escape from RNA interference: short interfering RNA-target recognition and implications for therapeutic approaches. J. Virol. 89, 1027–1035 (2005).
17. Mondote, I. A. & Saleh, M.-C. Antiviral immune response and the route of infection in Drosophila melanogaster. Adv. Virus Res. 100, 247–278 (2017).
18. Swevers, L., Liu, J. & Smaghe, G. Defense mechanisms against viral infection in Drosophila: RNAi and non-RNAi. Viruses 10, 230 (2018).
19. Gorlana-Armoux, D., Dostert, C., Schneemann, A., Hoffmann, J. A. & Imler, J.-L. Essential function in vivo for Dicer-2 in host defense against RNA viruses in Drosophila. Nat. Immunol. 7, 590–597 (2006).
20. van Ruij, R. P. et al. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. Genes Dev. 20, 2985–2995 (2006).
21. Wang, X.-H. et al. RNA interference directs innate immunity against viruses in adult Drosophila. Science 312, 452–454 (2006).
22. Zambon, R. A., Vakharia, V. N. & Wu, L. P. RNAi is an antiviral immune response against a DNA virus in Drosophila melanogaster. Cell. Microbiol. 8, 880–889 (2006).
23. Costa, A., Jan, E., Sarnow, P. & Schneider, D. The IMD pathway is involved in antiviral immune responses in Drosophila. PLoS ONE 4, e7436 (2009).
24. Sansone, C. L. et al. Microbiota-dependent priming of antiviral intestinal immunity in Drosophila. Cell Host Microbe 18, 571–581 (2015).
25. Ferreira, A. G. et al. The Toll-Dorsal pathway is required for resistance to viral oral infection in Drosophila. PLoS Pathog. 10, e1004570 (2014).
26. Zambon, R. A., Nandakumar, M., Vakharia, V. N. & Wu, L. P. The Toll pathway is important for an antiviral response in Drosophila. Proc. Natl Acad. Sci. USA 102, 7257–7262 (2005).
27. Dostert, C. et al. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of Drosophila. Nat. Immunol. 6, 946–953 (2005).
28. Merkling, S. H. et al. The epigenetic regulator G9a mediates tolerance to RNA virus infection in Drosophila. PLoS Pathog. 11, e1004692 (2015).
29. Charlton, P. D. & Johnson, K. N. The novel genome organization of the previously undiscovered virus Drosophila C virus suggests this virus belongs to a (1993).
30. Steffens, N. R. & Neher, R. A. Quantifying selection against synonymous mutations in HIV-1 env evolution. J. Virol. 87, 11843–11850 (2013).
31. Martínez, M. A., Jordan-Paiz, A., Franco, S. & Nevot, M. Synonymous virus genome recoding as a tool to impact viral fitness. Trends Microbiol. 24, 134–147 (2016).
32. Kieft, J. S., Zhou, K., Rubin, J. & Doudna, J. A. Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA 7, 194–206 (2001).
33. Zhang, J. et al. IncRNA sensing of a viral repressor of RNA activates non-canonical innate immune signaling in Drosophila. Cell Host Microbe 27, 115–128 (2020).
34. Navarro, R. et al. Defects in plant immunity modulate the rates and patterns of RNA virus evolution. Preprint at bioRxiv https://doi.org/10.1101/2020.10.33.74702. (2021).
35. Merkling, S. H. & van Ruij, R. P. Analysis of resistance and tolerance to virus infection in Drosophila. Nat. Protoc. 10, 1084–1097 (2015).
36. Lee, Y. S. et al. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117, 69–81 (2004).
37. Buchon, N., Broderick, N. A., Poidevin, M., Pradervand, S. & Lemaitre, B. Invasive and Drosophila melanogaster 19, 233–239 (1975).
38. Buchon, N., Broderick, N. A., Chakrabarti, S. & Lemaitre, B. Viral infection and stress affect protein levels of Dicer 2 and Argonaute 2 in Drosophila melanogaster. Front. Immunol. 11, 362 (2020).
39. Deddouche, S. et al. The DEAH-box helicase Dicer-2 mediates the induction of antiviral activity in Drosophila. Nat. Immunol. 9, 1424–1432 (2008).
40. Gomariz-Zilber, E., Jeune, B. & Thomas-Orillard, M. Limiting conditions of the horizontal transmission of the Drosophila melanogaster virus in its host (D. melanogaster). Proc. Natl Acad. Sci. USA 102, 236–241 (2005).
41. Steffens, N. R. & Neher, R. A. Quantifying selection against synonymous mutations in vesicular stomatitis virus. J. Mol. Biol. 342, 1415–1421 (2004).
42. Zanini, F. & Neher, R. A Quantifying selection against synonymous mutations in vesicular stomatitis virus. J. Mol. Biol. 342, 1415–1421 (2004).
43. Desai, M. M. & Fisher, D. S. Beneficial mutation–selection balance and the effect of linkage on positive selection. Proc. R. Soc. B 267, 1759–1798 (2000).
44. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. Biom. J. 50, 346–363 (2008).
45. Lenth, R. V. et al. Estimated Marginal Means, aka Least-Squares Means. R package version 1.5-5 (2021).
46. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. Am. J. Epidemiol. 27, 493–497 (1938).
47. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. Biom. J. 50, 346–363 (2008).
48. Fox, J. & Weisberg S. An R Companion to Applied Regression 3rd edn (Sage, 2019).
49. Lenth, R. V. et al. Estimated Marginal Means, aka Least-Squares Means. R package version 1.5-5 (2021).
50. Wickham, H. et al. Welcome to the Tidyverse. J. Open Source Softw. 4, 668 (2019).
51. Cormann, R. S. et al. Population-genomic variation within RNA viruses of the western honey bee, Apis mellifera, inferred from deep sequencing. BMC Genomics 14, 154 (2013).
52. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300 (1995).
53. Myers, M. A., Satas, G. & Raphael, B. J. CALDER: inferring phylogenetic trees from longitudinal tumor samples. Cell. 87, 514–522 (1999).
54. Smith, M. A. et al. E-scape: escape from single-cell phylogenetics and cancer evolution. Nat. Methods 14, 549–550 (2017).

**Acknowledgements**

We thank members of the Saleh Lab, M. Vignuzzi and J. Pfeiffer for fruitful discussions. We thank C. Meignin for Vamos and Vago files. This work was supported by the European Research Council (FP7/2013–2019 ERC CoG 615220) and the French Government’s Investissement d’Avenir programme, Labolatoire d’Excellence Integrative Biology of Emerging Infectious Diseases (grant no. ANR-10-LABX-62-IBED) to M. V. Work in S.F.E.’s laboratory was supported by grant nos BFU2015-65037-P and PID2019-103998GB-I00 (Spain Agencia Estatal de Investigación—FEDER) and PROMETEO2019/012 (Generalitat Valenciana).

**Author contributions**

V.M. and M.-C.S. conceived the study. V.M., M.-C.S., A.K. and L.Q.-M. established the experimental design. V.M., V.G., H.B. and I.N. performed the investigations. S.L. and S.F.E. performed the formal analyses. V.M., S.F.E. and M.-C.S. wrote the paper and acquired funding.
Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41559-022-01697-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-022-01697-z.
Extended Data Fig. 1 | Viral load and prevalence across the DCV evolution experiment. Viral load of 10 individual flies coming from DCV inoculated cages and four individual flies coming from mock inoculated cages was determined by TCID$_{50}$. a) Prevalence, calculated as the percentage of flies positive by TCID$_{50}$. b) Viral load determined by TCID$_{50}$ in each genotype across the 10 DCV passages. c) DCV replication assessed by negative-strand RT–qPCR. Left panel: standard curve produced from a tenfold dilution series over a range from $10^8$ to $10^3$ copies per reaction of in vitro transcribed RNA corresponding to a portion of the full-length negative-strand DCV RNA (slope = $-3.644$, $R^2 = 0.990$, efficiency = 88.25%). Right panel: amount of negative-strand DCV RNA present in the viral stocks produced from each fly genotype in P10, S2 DCV stock and DCV stock. Mock-infected flies were added as controls. LOD: Limit of detection of DCV negative stranded RNA. d) Average viral loads per individual fly of each genotype estimated from the GLM fitted to the data shown in panel b. Error bars represent ±1SD.
Extended Data Fig. 2 | Grouping of DCV population swarms by similarity and increasing nucleotide diversity (\(\pi\)). Viral nucleotide diversity (\(\pi\)) was determined in each condition and grouped using a post hoc Bonferroni test based on the pairwise comparisons from Supplementary Table 1. SE: standard error. asymp.LCL: asymptomatic lower confidence level; asymp.UCL: asymptomatic upper confidence level.

| Condition                        | mean \(\log_{10}\) \((n)\) per site | SE     | d.f. | asymp.LCL | asymp.UCL | Group |
|----------------------------------|--------------------------------------|--------|------|-----------|-----------|-------|
| **Full length genome, all viral passages** |                                      |        |      |           |           |       |
| \(w^{118}\)                      | 0.0005                               | 0.0001 | Inf  | 0.0003    | 0.0009    | 1     |
| \(Dif^{1/1}\)                    | 0.0007                               | 0.0002 | Inf  | 0.0004    | 0.0013    | 12    |
| \(Dcr-2^{B11fsX/B11fsX}\)       | 0.0011                               | 0.0002 | Inf  | 0.0006    | 0.0019    | 123   |
| \(Reo^{E20/20}\)                | 0.0014                               | 0.0003 | Inf  | 0.0008    | 0.0025    | 23    |
| \(spz^{2/2}\)                   | 0.0015                               | 0.0003 | Inf  | 0.0009    | 0.0027    | 234   |
| \(Dcr-2^{R416X/R416X}\)         | 0.0017                               | 0.0003 | Inf  | 0.0009    | 0.0029    | 234   |
| \(Ago-2^{114/414}\)             | 0.0023                               | 0.0005 | Inf  | 0.0013    | 0.0041    | 34    |
| \(Egfr^{1/11}\)                 | 0.0036                               | 0.0007 | Inf  | 0.0020    | 0.0063    | 4     |
| \(Vago^{D10/DM10}\)             | 0.0036                               | 0.0007 | Inf  | 0.0021    | 0.0063    | 4     |
| **Full fly genotype, all viral passages** |                                      |        |      |           |           |       |
| \(3'UTR\)                       | 1.98-10^{-65}                       | 2.44-10^{-66} | Inf | 1.46-10^{-65} | 2.70-10^{-66} | 1     |
| \(5'UTR\)                       | 0.0001                               | 1.25-10^{-6}  | Inf | 0.0001    | 0.0002    | 2     |
| ORF1                             | 0.0004                               | 3.68-10^{-6}  | Inf | 0.0004    | 0.0005    | 3     |
| ORF2                             | 0.0006                               | 4.91-10^{-6}  | Inf | 0.0005    | 0.0007    | 3     |
| **Full length DCV genome, \(p = 5\)** |                                      |        |      |           |           |       |
| \(w^{118}\)                      | 0.0004                               | 0.0002 | 10   | 0.0001    | 0.0023    | 12    |
| \(Reo^{E20/20}\)                | 0.0007                               | 0.0003 | 10   | 0.0001    | 0.0040    | 12    |
| \(Dif^{1/1}\)                    | 0.0011                               | 0.0005 | 10   | 0.0002    | 0.0064    | 12    |
| \(spz^{2/2}\)                   | 0.0014                               | 0.0007 | 10   | 0.0003    | 0.0080    | 12    |
| \(Ago-2^{114/414}\)             | 0.0015                               | 0.0007 | 10   | 0.0003    | 0.0082    | 12    |
| \(Dcr-2^{R416X/R416X}\)         | 0.0016                               | 0.0007 | 10   | 0.0003    | 0.0088    | 12    |
| \(Dcr-2^{B11fsX/B11fsX}\)       | 0.0018                               | 0.0008 | 10   | 0.0003    | 0.0099    | 12    |
| \(Egfr^{1/11}\)                 | 0.0021                               | 0.0010 | 10   | 0.0004    | 0.0117    | 12    |
| \(Vago^{D10/DM10}\)             | 0.0041                               | 0.0019 | 10   | 0.0007    | 0.0228    | 12    |
| DCV stock R1                     | 0.0120                               | 0.0057 | 10   | 0.0022    | 0.0667    | 2     |
| DCV stock R2                     | 0.0110                               | 0.0052 | 10   | 0.0020    | 0.0612    | 2     |
| **Full length DCV genome, \(p = 10\)** |                                      |        |      |           |           |       |
| \(Dif^{1/1}\)                    | 0.0003                               | 0.0001 | 10   | 0.0001    | 0.0015    | 1     |
| \(spz^{2/2}\)                   | 0.0006                               | 0.0002 | 10   | 0.0001    | 0.0026    | 1     |
| \(w^{118}\)                      | 0.0007                               | 0.0003 | 10   | 0.0002    | 0.0032    | 1     |
| \(Reo^{E20/20}\)                | 0.0008                               | 0.0003 | 10   | 0.0002    | 0.0036    | 12    |
| \(Dcr-2^{B11fsX/B11fsX}\)       | 0.0011                               | 0.0004 | 10   | 0.0003    | 0.0046    | 123   |
| \(Vago^{D10/DM10}\)             | 0.0019                               | 0.0007 | 10   | 0.0004    | 0.0080    | 123   |
| \(Egfr^{1/11}\)                 | 0.0020                               | 0.0008 | 10   | 0.0005    | 0.0085    | 123   |
| \(Dcr-2^{R416X/R416X}\)         | 0.0023                               | 0.0009 | 10   | 0.0005    | 0.0099    | 123   |
| \(Ago-2^{114/414}\)             | 0.0023                               | 0.0009 | 10   | 0.0005    | 0.0099    | 123   |
| DCV stock R2                     | 0.0110                               | 0.0044 | 10   | 0.0026    | 0.0472    | 23    |
| DCV stock R1                     | 0.0120                               | 0.0048 | 10   | 0.0028    | 0.0514    | 3     |
Extended Data Fig. 3 | Evolution of DCV variants. a) Trajectories of DCV variants across passages. N: total number of SPNs found above the estimated frequency threshold (≥ 0.0028). Trajectories of viral variants found significant after FDR correction are show in green (p ≤ 0.006) and yellow (0.047 ≤ p ≤ 0.006) (based on data from Table 2). b) to k) Heatmaps showing the Pearson correlation coefficients between mutations’ frequencies along evolutionary time, ranging from blue, where no linkage between the SNPs was found, to red, where the SNPs were linked in a same viral haplotype.
Extended Data Fig. 4 | SNPs on the DCV genome with significant estimates of fitness effects. Green triangles represent synonymous mutations, pink triangles non-synonymous mutations and grey triangles mutations in non-coding sequences. Cases significant after FDR correction ($p \leq 0.006$) are marked with an asterisk (based on data from Table 2).
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

☐ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐ The statistical test(s) used AND whether they are one- or two-sided
☐ Only common tests should be described solely by name; describe more complex techniques in the Methods section.
☐ A description of all covariates tested
☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

n/a

Data analysis

a) Comparison of survival curves was performed using a log-rank (Mantel–Cox) test and Prism v.8.4.3 (www.graphpad.com).
b) Bioinformatics Analysis of NGS Libraries. VIVAN (http://www.vivanbioinfo.org) was used for SNPs detection. The quality of fastq files was assessed using graphs generated by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/).
Using cutadapt (https://cutadapt.readthedocs.io/en/stable/), low-quality bases and adaptors were trimmed from each read. Codes for algorithms used in this study will be deposited in GitHub and are referenced in the manuscript.
c) Variant frequency threshold, viral nucleotide diversity and estimation of relative mutational fitness effects were performed in R version 4.0.2 in RStudio version 1.3.1073.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.
**Data**

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data that support the findings of this study are available from the corresponding author upon request. In-house codes are also available at any time upon request to the authors.

**Field-specific reporting**

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- [ ] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

**Life sciences study design**

All studies must disclose on these points even when the disclosure is negative.

| Sample size | a) DCV experimental evolution assay: |
|-------------|-------------------------------------|
| N = 500 | 5 to 6 days old wild type or mutant flies (males and females) were used in each viral passage. This procedure was repeated in 2 biological replicates for viral infected flies and in 1 biological replicate for mock inoculated flies. |
| i) To produce the NGS library total RNA was extracted from half of the population of infected flies from each fly genotype (approx. 250 flies) and each viral passages. |
| One NGS library was produced and sequenced per fly genotype/viral passage/biological replicate. |
| Four NGS libraries (2 biological and 2 technical replicates) were produced and sequenced from S2 DCV stock and DCV stock. |
| ii) To asses DCV virulence, infectious DCV stocks were produced from viral passages P = 1 and P = 10 using half of the population of infected flies from each fly genotype (approx. 250 flies). Survival curves were performed in w1118 flies. Two independent experiments with three biological replicates of 20 flies each were done per condition. Total number of flies per viral stock is indicated between brackets. |

**Biological Replicate 1**

- Mock infected flies (235)
- S2 DCV stock (235)
- DCV stock (231)

**Biological Replicate 1 - P = 1**

- w1118 stock (119)
- Dcr-2 L811fsX/L811fsX stock (120)
- Dcr-2 R416X/R416X stock (118)
- Ago-2 414/414 stock (120)
- Spz 2/2 stock (119)
- Dif 1/1 stock (117)
- Rel E20/E20 stock (118)
- VagoΔ10/Δ10 stock (120)
- Egfr t1/t1 stock (119)

**Biological Replicate 1 - P = 10**

- w1118 stock (116)
- Dcr-2 L811fsX/L811fsX stock (114)
- Dcr-2 R416X/R416X stock (106)
- Ago-2 414/414 stock (115)
- Spz 2/2 stock (118)
- Dif 1/1 stock (117)
- Rel E20/E20 stock (110)
- VagoΔ10/Δ10 stock (108)
- Egfr t1/t1 stock (115)

**Biological Replicate 2**

- Mock infected flies (235)
- S2 DCV stock (225)
- DCV stock (233)

**Biological Replicate 2 - P = 1**
iii) To determine viral load and persistence during the course of the viral passages, viral load of 10 individual flies (5 males and 5 females) from DCV inoculated cages and four individual flies (2 males and 2 females) from mock inoculated cages was determined.

iv) To determine DCV replication, production of negative strand RNA was quantified using ssRT-qPCR according to the protocol published in DOI: 10.1016/j.jmb.2021.167308. The amount of negative strand DCV RNA present in the viral stocks produced from each fly genotype in P = 10, S2 DCV stock, and DCV stock was determined.

b) To characterize the newly produced back-crossed fly lines, survival curves were performed, the number of flies per condition is indicated between brackets. Two to three independent experiments with three biological replicates of 15 to 25 flies were done per condition.

DCV
- w1118 (185); Dcr-2 L811fsX/L811fsX (182)
- w1118 (191); Dcr-2 R416X/R416X (183)
- w1118 (207); Ago-2 414/414 (161)
- w1118 (131); Spz 2/2 (131)
- w1118 (126); Dif 1/1 (132)
- w1118 (194); Rel E20/E20 (180)
- w1118 (132); VagoΔ10/Δ10 (131)
- w1118 (131); Egfr t1/t1 (132)

Enterococcus faecalis
- w1118 (83); Spz 2/2 (82)
- w1118 (83); Dif 1/1 (84)
- w1118 (83); Rel E20/E20 (86)

Erwinia carotovora
- w1118 (90); Spz 2/2 (84)
- w1118 (90); Dif 1/1 (87)
- w1118 (90); Rel E20/E20 (71)

Data exclusions
- No data were excluded from the analysis.

Replication
- All attempts at replication were successful.

Randomization
- For DCV evolution experiment and survival curves, synchronized flies reared in standard medium were randomly collected from different tubes and pooled, and the treatment was assigned.

Blinding
- Blinding was not performed during the experiment, data acquisition, or analysis.

Reporting for specific materials, systems and methods
We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.
### Materials & experimental systems

- Antibodies
- Eukaryotic cell lines
- Palaeontology and archaeology
- Animals and other organisms
- Human research participants
- Clinical data
- Dual use research of concern

### Methods

- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

### Eukaryotic cell lines

**Policy information about cell lines**

- **Cell line source(s):** Drosophila S2 cells, Life Technologies.
- **Authentication:** None of the cell lines were authenticated.
- **Mycoplasma contamination:** All cell lines used in this study tested negative for mycoplasma contamination.
- **Commonly misidentified lines:**
  - Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

### Animals and other organisms

**Policy information about studies involving animals:** ARRIVE guidelines recommended for reporting animal research

- **Laboratory animals:** Drosophila melanogaster (non ethical permission required).
- **Wild animals:** n/a
- **Field-collected samples:** n/a
- **Ethics oversight:** n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.