

On the Fly: Host Determinants of RNA Virus Evolution

Óscar Morales Lezcano



**RADBOD
UNIVERSITY
PRESS**

Radboud
Dissertation
Series

On the Fly:

Host Determinants of RNA Virus Evolution

Óscar Morales Lezcano

On the Fly: Host Determinants of RNA Virus Evolution

Óscar Morales Lezcano

Radboud Dissertation Series

ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS

Postbus 9100, 6500 HA Nijmegen, The Netherlands

www.radbouduniversitypress.nl

Design: Proefschrift AIO | Annelies Lips

Cover: Iballa Vera Almeida and Camila Valentina Alonso Yépez

Printing: DPN Rikken/Pumbo

ISBN: 9789465152837

DOI: 10.54195/9789465152837

Free download at: <https://doi.org/10.54195/9789465152837>

© 2026 Óscar Morales Lezcano

**RADBOUD
UNIVERSITY
PRESS**

This is an Open Access book published under the terms of Creative Commons Attribution-Noncommercial-NoDerivatives International license (CC BY-NC-ND 4.0). This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator, see <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

On the Fly:

Host Determinants of RNA Virus Evolution

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 18 september 2026
om 12.30 uur precies

door

Óscar Morales Lezcano
geboren op 6 april 1995
te Las Palmas de Gran Canaria, Spanje

Promotoren:

Prof. dr. ir. R.P. van Rij

Prof. dr. M.A. Huynen

Manuscriptcommissie:

Prof. dr. G.J.C. Veenstra (voorzitter)

Prof. dr. A.E. Gorbalenya (Universiteit Leiden)

Dr. B. Longdon (University of Exeter, Verenigd Koninkrijk)

On the Fly:

Host Determinants of RNA Virus Evolution

Dissertation to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.M. Sanders,
according to the decision of the Doctorate Board
to be defended in public on
Friday, September 18, 2026
at 12.30 pm

by

Óscar Morales Lezcano

born on April 6, 1995
in Las Palmas de Gran Canaria, Spain

Supervisors:

Prof. dr. ir. R.P. van Rij

Prof. dr. M.A. Huynen

Manuscript Committee:

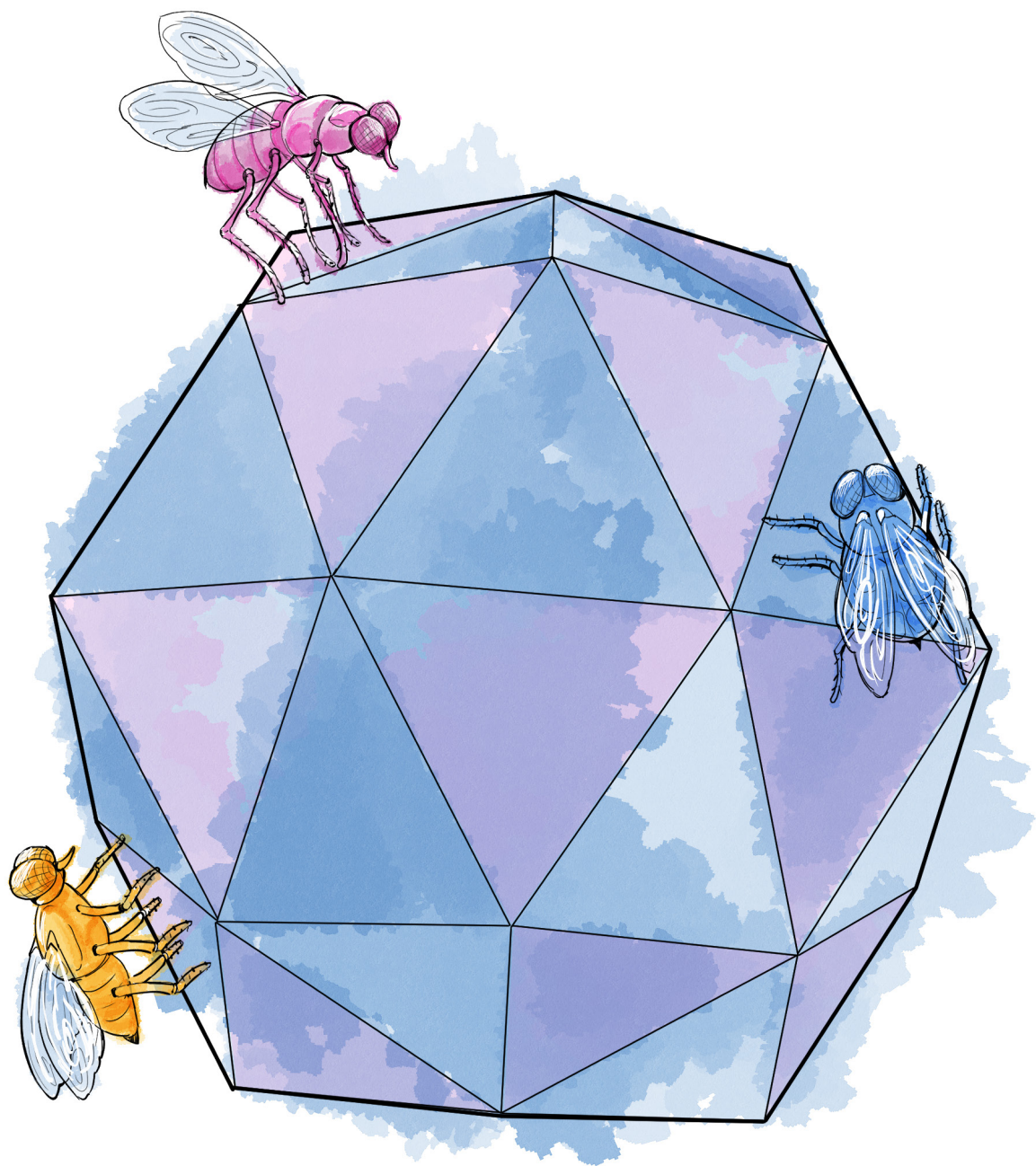
Prof. dr. G.J.C. Veenstra (chair)

Prof. dr. A.E. Gorbalenya (Leiden University)

Dr. B. Longdon (University of Exeter, United Kingdom)

Table of contents

Chapter 1	General introduction	9
Chapter 2	Parallel evolution and enhanced virulence upon <i>in vivo</i> passage of an RNA virus in <i>Drosophila melanogaster</i>	39
Chapter 3	A recurrent adaptive mutation in the transmembrane 2B protein of an insect picorna-like virus in a nonnative host	79
Chapter 4	Exploring the role of RNA-modifying enzymes in insect RNA virus infection	121
Chapter 5	General discussion	147
Appendix	Summary	176
	Samenvatting	178
	Research data management	180
	Acknowledgements	181
	Curriculum vitae	187
	List of publications	188
	PhD portfolio	189



Chapter 1

General introduction

From devastating pandemics to persistent infections, viruses have long shaped the biosphere. Their capacity to cause disease has imposed profound health and economic burdens throughout history. Beyond their pathogenic role, viruses exhibit such genetic diversity and rapid adaptation that they serve as useful models for studying fundamental principles of molecular evolution. Deciphering these processes is not only of theoretical interest but is also critical for anticipating emerging viruses and designing effective control strategies. My thesis explores how host innate defenses influence virus evolution, using *Drosophila melanogaster* as a model system to provide insights into the complex process of virus-host adaptation.

In this general introduction, I will outline early discoveries that marked the birth of virology and revolutionized the understanding of viruses, as well as review the known sources of their genetic diversity. I will then discuss the fruit fly *D. melanogaster* as a model for studying immunity and summarize its principal antiviral immune responses. Finally, I will introduce other regulators of immunity and virulence, focusing on RNA modifications.

A brief history of virology

Viruses are obligate intracellular parasites that infect all cellular life forms and rely on the host's intracellular machinery to reproduce. This feature led to a long-lasting debate about whether viruses are alive, which is largely semantic, as it depends on the employed definition of life (1). They are striking for their simplicity: selfish replicators of their genetic material (either DNA or RNA) packed inside a protein capsid, which forms the viral particle or virion in non-enveloped viruses. In enveloped viruses, the capsids are additionally covered in a host-derived lipid membrane. Like any other life form, viruses maximize their fitness by adapting to their environment and are possibly the most successful entities of the biosphere, considering that they are the most abundant and evolvable Earth inhabitants (2).

The word "virus" is derived from the Latin word for "poisonous substance" and was historically used in a broad disease-related context for centuries before being linked specifically to infectious agents after the discovery of viruses. During the late nineteenth century, three researchers independently complemented each other's work on a disease that produced mosaic-like mottling on the leaves of tobacco plants and was ruining tobacco plantations.

They are today considered the fathers of virology (3). The German scientist Adolf Mayer found an infective agent in the sap of diseased plants that produced the infection of healthy plants upon inoculation in year 1886. In 1892, the Russian botanist Dimitri Ivanovsky uncovered that the sap remained infective after filtration with a Chamberland filter, indicating that the disease agent was smaller than any previously described pathogen, such as bacteria. Lastly, the Dutchman Martinus Beijerinck, unaware of Ivanovsky's results, validated his observations in 1898 and noticed the ability of the pathogen to replicate within the growing parts of the plants, claiming that the pathogen only replicated in tissues undergoing active cell division. Although they were not aware of it, they had discovered the tobacco mosaic virus (TMV). Soon after, other authors documented "filterable" agents able to infect animals, including foot-and-mouth disease virus (1898), yellow fever virus (1901) and poliovirus (1909).

TMV kept a central role in the conceptual advances of virology, as it was the first virus ever to be purified (4) and studied by x-ray crystallography (5, 6), allowing the visualization of viral proteins and particles. In the late 1950s, it was shown that the infectious genetic information of this virus is stored in RNA molecules (7) and, only in the early 1970s, that viral particles self-assemble (8).

Origin and diversification of viruses

The origin of viruses remains one of the most intriguing evolutionary puzzles. Three classical hypotheses have been proposed to explain their emergence (9). (i) The regression hypothesis suggests that viruses originated from the degeneration of cells that gradually lost functions necessary for independent survival and became parasitic entities. (ii) The "escaped genes" hypothesis proposes that viruses arose when cellular genes acquired the ability for selfish replication and spread. (iii) The "virus-first" hypothesis postulates that viruses preceded cellular life, emerging from the primordial pool of self-replicating molecules in the RNA world. The RNA world is a theory that considers RNA-based molecules to have appeared before DNA and proteins in the primordial soup (10), and it is supported by the discovery of autocatalytic and self-replicating RNA molecules, known as ribozymes (11, 12).

A chimeric scenario integrating the "virus-first" and the "escaped genes" hypotheses was proposed by Krupovic, Dolja, and Koonin in 2019 (13). In this framework, most viral genes are thought to have been acquired from cellular hosts through 20 independent recruitment events over evolutionary history. In contrast, a small set of hallmark viral genes involved in genome replication

lack detectable homologues in extant cellular life, supporting the idea that these genes originated from primitive genetic replicators (14, 15). The diversification of the virus ancestors that generated the current virosphere is a product of millions of years of co-evolution of viruses and hosts, host switches, and the ongoing emergence of new viruses.

What are the forces driving virus evolution?

Understanding viral molecular evolution—the process by which viral genomes change over time—is critical, as it provides information for developing effective strategies for virus control.

High mutation rates and huge population sizes

Despite RNA virus evolution being dominated by the same forces as any other organism—namely mutation, selection and drift—RNA viruses are arguably unique entities that differ profoundly from organisms with a double-stranded DNA (dsDNA) genome. In dsDNA-based organisms, reproduction is generally characterized by highly accurate replication of genetic material, e.g. the human genome has an error rate of approximately 1×10^{-8} per base pair per generation (16). Moreover, dsDNA-based organisms typically possess large genomes, with eukaryotic genome sizes ranging from ~10–100,000 Mb (17), bacterial and archaeal genomes from ~0.6–14.3 Mb (18), and dsDNA virus genomes spanning from ~2 kb to nearly 2.5 Mb (19).

In contrast, RNA virus replication is characterized by the highest mutation rate in nature—with perhaps the exception of single stranded DNA viruses (20)—combined with small genomes and immense population sizes. RNA viruses have an average genome size of around 10 kb, although it ranges from 1.7 kb up to 64 kb in the recently discovered giant Nidoviruses (21, 22). RNA virus replication is conducted by either the RNA-dependent RNA polymerase (RdRp) or, in the case of retroviruses, by a reverse transcriptase (RT). Both enzymes lack proofreading activity, thus cannot repair nucleotide misincorporations and are notoriously error-prone as a consequence. The RdRp is predicted to introduce approximately one mutation per genome per replication cycle (23, 24), although coronaviruses are an exception generating roughly 0.1 mutations per genome per cycle, as further discussed below (25). In comparison, RT introduces between 0.1 and 0.3 mutations per genome per replication cycle (26, 27). Estimates for within-host viral population sizes can range from

10^9 to 10^{12} (28). For example, it has been approximated that 10^{12} new virions are produced daily in a host infected with the RNA virus hepatitis C (HCV) (29) and about 10^{10} new virions in a human immunodeficiency virus-type 1 (HIV-1)-infected host (30). This combination of tremendous population size and unbridled mutations provides sufficient genetic diversity for the RNA viruses to adapt to rapidly changing environments (31, 32).

The notion that high mutation rates benefit RNA viruses by enhancing adaptability is widely accepted in virology. However, most mutations are deleterious and are rapidly eliminated from viral populations (33–36). As a consequence, some authors have challenged this view, arguing that the accumulation of deleterious mutations ultimately has a net negative effect on viral fitness (37). Indeed, RNA viruses have such mutation load that they are thought to replicate close to the so-called error threshold (38), defined as the maximum number of mutations that a virus population can sustain before experiencing error catastrophe—a decline in viability due to an excess of deleterious mutations—and extinction.

The high mutation rate of RNA viruses has several important implications worth discussing. First, their small genome size is likely a consequence of the constraints imposed by mutational load as protection from the error catastrophe, and forces RNA viruses to adopt strategies that compact genetic information. Second, a trade-off with replication rate has been proposed, whereby high error rates arise as an inevitable byproduct of selection for rapid replication (37, 39, 40). Lastly, the error-prone replication of RNA viruses may reflect Eigen's paradox, in which major improvements in replication fidelity would require a larger genome encoding more complex enzymes, yet such complexity cannot evolve without first reducing the error rate (31). In other words, fidelity is required to evolve complexity, but complexity is needed to achieve fidelity. Indeed, the largest known RNA virus genomes, ranging from 26 to 64 Kb, are present in the *Coronaviridae* and *Roniviridae* families (in the order *Nidovirales*) (21, 41). These viruses have a distinctive replication strategy, with an expanded replicase machinery containing several functional domains in addition to the RdRp. One of the best-characterized of these is the non-structural protein 14 of coronaviruses, which increases replication fidelity due to its RNase activity and a 3'-to-5' exoribonuclease (ExoN) that confers proofreading activity to the replicase complex (42–44). Loss of ExoN results in a 21-fold increase in the mutation rate in severe acute respiratory syndrome coronavirus and a 15-fold increase in murine hepatitis virus, in addition to an increased sensitivity to a mutagen and reduced viral fitness (45–47).

Interestingly, an insect nidovirus was discovered with an intermediate genome size (~20 kb) that had acquired ExoN, and was suggested to be a missing evolutionary link between small and large nidoviruses (48). A model of unidirectional nidovirus genome expansion was then proposed, in which increased replication fidelity allowed the subsequent expansion of genome size and complexity, thus escaping Eigen's paradox (49). Overall, it is clear that high mutation rates can help RNA viruses adapt to rapidly changing environments, although it also imposes a deleterious mutation burden that can be detrimental. RNA viruses tolerate this cost largely because their vast population sizes ensure sufficient reversal and compensatory mutations to produce enough viable progeny (31).

Natural selection and genetic drift

The concepts of evolution and natural selection are so intertwined that they are often regarded as synonymous. They constitute the basis of Darwinian theory:

"This preservation of favourable individual differences and variations, and the destruction of those which are injurious, I have called Natural Selection, or the Survival of the Fittest". (50, p.41)

In the most simplistic view, the rare viral genomes carrying a beneficial allele are positively selected, and the strength of their fitness advantage can be expressed as the selection coefficient, s . Conversely, those genomes with deleterious alleles are under purifying selection. RNA virus populations are expected to reach a mutation-selection balance, i.e., an equilibrium between the influx of genetic diversity created by mutation and the outflow cleansed by purifying selection and by genetic drift. RNA viruses seem to have a lower ratio of nonsynonymous to synonymous substitutions in protein-coding genes than DNA viruses and mammals, suggesting that they are under a stronger purifying selection that limits their protein evolution (51, 52). This is likely related to their compact genome, in some cases with overlapping reading frames, frequent pleiotropy—one gene influencing several phenotypic traits—and epistatic interactions, where the phenotype of an allele depends on the genetic background (38).

However, evolution in natural ecosystems is a multifactorial process, and natural selection should not be considered in isolation. The frequency of a given allele can also be affected by random genetic drift, a stochastic process fed by sampling error in a replicating population. Classical population genetics predicts that the effective population size (N_e : defined as the number of individuals contributing offspring to the next generation) determines the relative influence

of selection and drift (53). Specifically, when N_e is small and the product $N_e s \ll 1$, for example during transmission bottlenecks, genetic drift dominates evolutionary dynamics. This leads to reduced genetic diversity, promoting the accumulation of deleterious mutations and even the loss of wild-type alleles, thereby reducing the population fitness in a process known as Muller's ratchet (54, 55). Conversely, when $N_e s \gg 1$, natural selection dominates the fixation of mutant alleles. This principle has recently received empirical support from an experimental evolution study of several potato virus Y variants propagated under different N_e conditions (56). Low N_e limited viral fitness gains, whereas high N_e increased fitness—irrespective of the replicative fitness of the initial variant—through the fixation of adaptive mutations. In principle, estimating N_e would be sufficient to determine how deterministic evolution is for a given selection coefficient. However, because it is hard to directly measure how many individuals contribute genetically to the next generation, the estimates are often inferred from genetic diversity metrics or from allele frequency changes over time (57, 58). N_e has been documented to be orders of magnitude smaller than the total population size for several viruses—approximately 13,000-fold lower for influenza A virus and at least 300-fold lower for HIV1 (59–61). Nonetheless, the vast intra-host population size of RNA viruses still allow natural selection to play a dominant role in driving their evolution.

Recombination and reassortment

Recombination occurs in many RNA viruses and can play a relevant role in evolution. There are two types of recombination in RNA viruses: RNA recombination, which occurs in many viruses (62), and reassortment, which is unique to segmented viruses, like influenza A virus, whose genome is split into eight RNA segments (63). Both recombination and reassortment require at least two viruses to co-infect a single cell. During RNA recombination, the viral polymerase stops replicating the donor template and switches to the acceptor template while remaining bound to the nascent RNA (64). This produces a chimeric molecule of mixed ancestry in a process termed copy-choice replication (65, 66). Template switching is affected by the sequence identity between template molecules and by RNA secondary structure (67, 68).

The recombination frequency is highly variable among different viruses. It is abundant in retroviruses (27), with estimated rates for HIV-1 of around 1.4×10^{-5} per site per generation (69), and in some positive-sense single-stranded RNA ((+)ssRNA) viruses, like poliovirus from the *Picornaviridae* family, with a recombination rate in the order of 10^{-6} per site per generation (70).

Nonetheless, recombination is rarer in other (+) RNA virus families like the *Flaviviridae* (71), where the HCV recombination rate is estimated at 4×10^{-8} per site per generation in cell culture (72). Lastly, negative-sense ssRNA viruses have the lowest recombination rate (73).

Recombination could be conceived as a form of sexual reproduction in asexual populations that accelerates the appearance of beneficial allele combinations and purges deleterious combinations (74). However, recombination cannot offer a generalized advantage for virus evolution, considering the high variation in recombination rates across the virosphere. The high mutation rate, together with the extremely high population size, may offer enough genetic diversity to prevent the need for recombination in RNA viruses. Thus, some authors consider recombination as a mechanistic by-product of RNA polymerase processivity, which varies according to the genomic architecture of each virus (75).

Quasispecies

Since its formulation in the 1970s, quasispecies theory has reshaped the understanding of viral population evolution (76). Viral quasispecies are defined by a distribution of closely related variants (also named mutant swarms or mutant clouds) that originate from error-prone RNA replication. Although the term quasispecies was not coined until later (77), the concept originates from Eigen's seminal work on natural self-replicating macromolecules, like RNA replicons, and aims to explain the early evolution of life in the RNA world (78). The quasispecies theory was applied to viral populations for the first time by Esteban Domingo (79).

Quasispecies is the consequence of mutation-selection balance in which the mutant clouds are clustered around the main variant (referred to as master or consensus sequence) (80). The critical assumption of this theory is that the mutation rate is so high that the frequency of each variant is not only given by its replication rate but also by the probability of originating from erroneous replication of neighboring variants in sequence space—one or a few mutations away. This mutational coupling between neighboring variants favors group selection in the quasispecies, shifting selection from individual genotypes to the mutant swarm as a whole (81). As a consequence, the quasispecies evolves to maximize its average fitness, allowing slow-replicating variants to be selected if they are close to beneficial mutations in the sequence space. At high mutation rates, populations composed of highly fit variants can be outcompeted by populations with lower individual variant fitness but that tolerate mutagenesis better by being far from deleterious mutations. Thus, the average fitness of

the latter mutant swarm is greater due to its higher genome robustness (82), a phenomenon that has been termed survival of the flattest (83, 84). Of note, synergistic effects in quasispecies have only been demonstrated by mathematical models and *in vitro* assays where the presence of mutagens artificially increases the mutation rate (78, 82). Whether a survival of the flattest effect can be obtained in a population depends on the product of the population size and the mutation rate (85), however, whether mutation rates in natural viral populations are high enough to reach the mutational coupling of close variants required for this quasispecies behavior still needs to be shown.

***Drosophila melanogaster* as a model to study innate immunity and virus-host interactions**

During the last century, the fruit fly *Drosophila melanogaster* has been a pillar of genetic research, contributing to fundamental discoveries in development and immunity (86, 87). The fly immune system is well studied and, although it lacks adaptive immunity, it shares conserved innate immune pathways with mammals, making it a powerful model to study fundamental principles of immunology. One notable example is the discovery of the Toll pathway, initially described as required for establishing dorsal-ventral polarity during embryonic development (88) and later shown to be essential for antifungal immunity (89). This work, for which Jules Hoffmann was awarded the Nobel Prize, paved the way for the identification of Toll-like receptors in mammals (90), reshaping understanding of innate immune sensing.

A key strength of *Drosophila* as a model organism is the availability of sophisticated genetic tools that allow precise spatial and temporal manipulation of gene expression. One of the most widely used tools is the UAS-Gal4 system (Fig. 1A), originally adapted from a yeast transcriptional control circuit (91). This binary system consists of a driver line expressing the yeast transcription factor Gal4 under the control of a chosen promoter that confers tissue or developmental specificity, and a responder line carrying the sequence of interest downstream of the upstream activating sequence (UAS), which is recognized by Gal4. Crossing driver and responder lines produces F1 progeny where the sequence of interest is expressed according to the promoter driving Gal4. Further temporal control can be achieved by introducing a temperature-sensitive version of the Gal4 repressor Gal80 (Gal80^{ts}) (92) (Fig. 1B). At low temperature (18°C), Gal80^{ts} is active and represses Gal4

activity, whereas shifting flies to 29°C inactivates Gal80^{ts} and induces expression of the transgene. By leveraging the UAS-Gal4 system, genome-wide gene silencing screens have enabled the functional characterization of numerous previously unannotated genes, including the identification of host factors required for influenza A virus replication (93, 94).

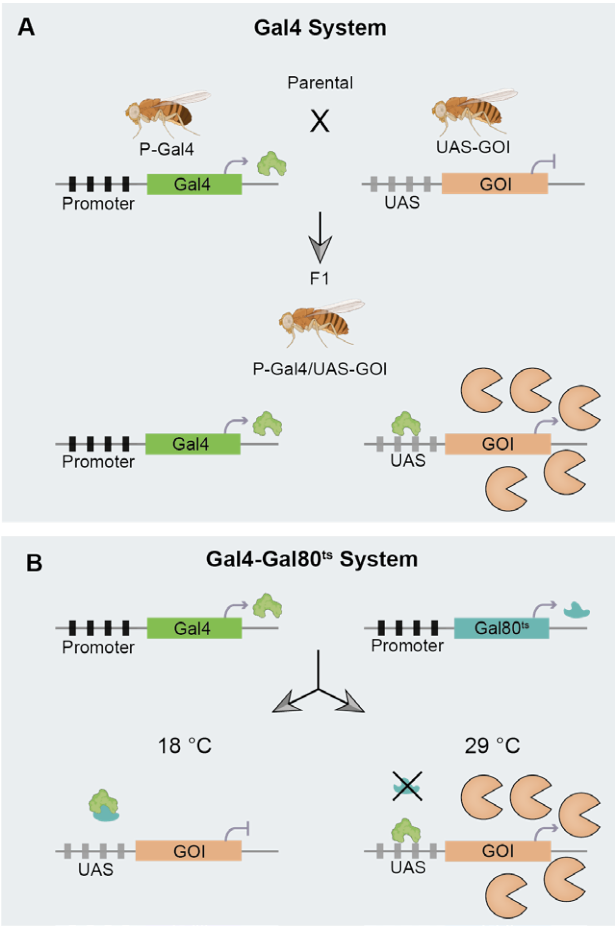


Figure 1. (A) The *Drosophila* UAS-Gal4 system is a binary genetic tool in which a driver fly line expressing the yeast transcription factor Gal4 under the control of a specific promoter (P-Gal4) is crossed with a responder line carrying a transgene encoding the gene of interest (GOI) downstream of the upstream activating sequence (UAS), referred to as UAS-GOI. The F1 progeny carrying both transgenes (P-Gal4/UAS-GOI) are selected based on phenotypic markers introduced during transgenesis and express the GOI specifically in the Gal4 expression domain. **(B)** The Gal4-Gal80^{ts} system adds the temperature-sensitive Gal80 repressor (Gal80^{ts}), which binds to Gal4 and inhibits its transcriptional activity at permissive temperatures (18 °C). Upon shifting flies to the restrictive temperature (29 °C), Gal80^{ts} becomes inactive, releasing Gal4 and thereby inducing temporal expression of the GOI. Fly icons were retrieved from BioRender.

Drosophila melanogaster has been successfully used as a model to study antiviral immunity (95, 96). We exploited this organism to deepen our understanding of virus-host interactions, focusing on one of the natural fly pathogens, *Drosophila C virus* (DCV) (97, 98). DCV belongs to the family *Dicistroviridae* within the *Picornavirales* order (99, 100). First identified in the 1970s, DCV was subsequently found to be present in numerous laboratory stocks (101). DCV has a (+)ssRNA genome of approximately 9,200 nucleotides that encodes two open reading frames (ORFs) (102). The genome is packaged into non-enveloped icosahedral particles and carries a VPg protein covalently bound to its 5' end (Fig. 2). The 5' ORF encodes the non-structural proteins, while the 3' ORF encodes the capsid proteins, and both ORFs are preceded by internal ribosomal entry sites (IRESs). The 5' IRES eliminates the need for capping the viral genome, whereas the intergenic region (IGR) IRES initiates viral protein synthesis independently of translation initiation factors (103). The IGR IRES binds to the ribosome, by mimicking the tRNA anticodon loop base-paired with mRNA, and functions as a translation elongation factor, strikingly mediating translation without requiring the initiating methionine-tRNA (104).

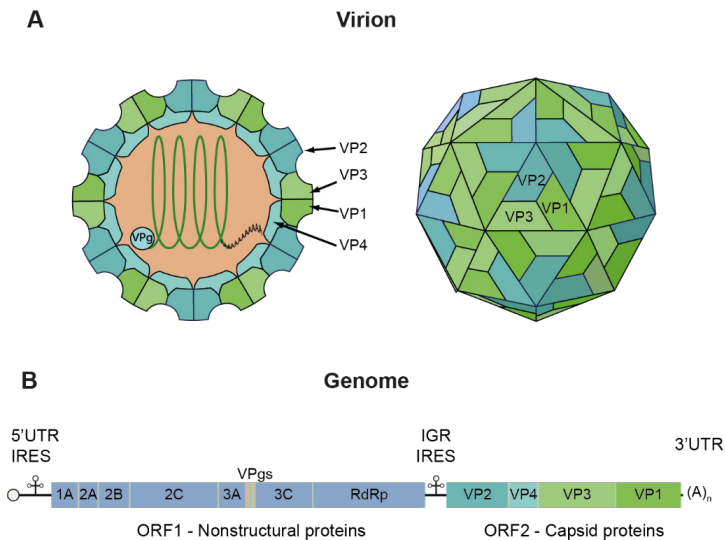


Figure 2. Virion (A) and genome organization (B) of viruses from the *Dicistroviridae* family. Virions are non-enveloped, icosahedral particles with pseudo-T=3 symmetry, composed of 60 protomers of VP1, VP2, and VP3. The small VP4 protein is located inside.

These genomic and translational features are shared with the closely related cricket paralysis virus (CrPV) (105-107), which was originally isolated from wild Australian crickets (108), and was used in Chapter 3 to study evolution in cross-species virus transmission. In the laboratory, CrPV has a broad host range, infecting a wide panel of fly species in the *Drosophilidae* family with relatively consistent replication rates across hosts, when compared to DCV (109, 110). However, CrPV has never been identified in metagenomic surveys of wild *Drosophila*, suggesting it is not a natural pathogen of fruit flies.

Determinants of cross-species virus transmission

Novel infectious diseases are frequently due to cross-species transmission (111), in which a virus shifts from its native host to a new species. While most viruses are restricted to a limited host range, transitions between host species occur repeatedly over evolutionary history. Host jumping requires that the new host is exposed to a virus, becomes infected, and produces enough viral progeny to ensure onward transmission. At the mechanistic level, this depends on the virus overcoming multiple host-specific barriers and generally requires the acquisition of one or several adaptive mutations that optimize cell entry, viral replication and host immune evasion. Perhaps the high mutation rate of RNA viruses explains their pronounced propensity to shift hosts compared to other pathogens (112). Interestingly, the same mutations often arise independently in parallel during adaptation to a new host, a pattern observed repeatedly in both natural populations and laboratory experiments (113-118). Parallel mutations are likely a sign of the limited number of viable evolutionary trajectories allowing adaptation to the new host environment (119), presumably due to the compact genomes of RNA viruses. Overall, some key determinants of effective host switching are the number of required adaptive changes and molecular solutions conferring adaptation, the order in which mutations get fixed, and their epistatic interactions.

Host phylogenetic relatedness is another important determinant of a virus' ability to infect novel species, but it seems to influence host shifting in two distinct ways. While viruses often replicate less efficiently in hosts that are increasingly distant from their natural host, clades of highly susceptible (or resistant) hosts have been observed independently of the phylogenetic distance from the original host in *Drosophilidae* species infected with different *Drosophila* sigma viruses (119, 120). Irrespective, variation in the ability to infect a new host

could reflect the compatibility of a virus with the new host's cellular machinery, or the virus' ability to evade the new host's immune response.

Beyond genetic compatibility, ecological factors strongly influence the likelihood of cross-species transmission. Contact rates between species and host population density modulate opportunities for viral spillover (121). Importantly, ecological opportunity alone is insufficient: even frequent exposure does not guarantee successful emergence unless the virus can replicate efficiently and transmit in the new host. The integration of ecological exposure with within-host evolutionary dynamics is therefore essential for understanding virus emergence.

Fly immune response and its role in virus-host co-evolution

Innate immunity is an evolutionarily ancient host defense system that relies on functionally analogous components shared across plants and animals, including pattern recognition-based pathogen detection. While the specific receptors generally differ between kingdoms, the underlying principles of innate immune sensing and signaling predate the divergence of plants and animals (122-124). An estimated five to ten million species in the kingdom Animalia depend entirely on innate immunity to combat pathogens, while only 45,000 extant vertebrates have developed the more elaborate adaptive immunity (125). Immune mechanisms must detect invading pathogens as non-self and mount defensive responses that efficiently target them while minimizing self-damage. To do so, innate immunity relies on a limited set of germline-encoded receptors (*pattern recognition receptors*, PRRs) that recognize either conserved structures shared between different microorganisms, known as *pathogen associated molecular patterns* (PAMPs), or common metabolic consequences of infection, known as *damage-associated molecular patterns* (DAMPs). Innate immunity can trigger an inflammatory response just minutes after pathogen recognition (126), whereas adaptive responses require about five days to clonally expand the antigen-specific lymphocytes from a diverse pool of T and B cells expressing somatically recombined receptors (127). These receptors recognize epitopes that are pathogen-specific and generate "immune memory" against subsequent infections. Importantly, the initiation and regulation of adaptive immunity are tightly orchestrated by innate immune mechanisms, primarily through the production of cytokines by innate immune cells (128, 129).

The *Drosophila* immune system has been thoroughly characterized and is composed of cellular and humoral responses, as well as RNA-based immunity. Cellular immunity is mediated by blood cells, called hemocytes, ninety-five percent of which are plasmatocytes that resemble mammalian monocyte/macrophage cells (130). Plasmatocytes circulate in the fly hemolymph phagocytosing cell debris and microbes (131), but their involvement in virus clearance seems modest and virus-specific (132). The hallmark of the humoral response is the rapid production of antimicrobial peptides (AMPs), produced primarily in the fat body and released into the hemolymph. There is an impressive body of literature on the molecular mechanisms of the signaling cascades that trigger AMP production, namely the Toll, immune deficiency (IMD) and JAK-STAT pathways (133-138). Toll responds to fungal and Gram-positive bacterial infections, whereas IMD reacts to Gram-negative bacteria. There is some evidence that the humoral response participates in the defense against virus infection, although the sensing mechanism has been elusive (96, 139-143). Given the scope of this thesis, I will not elaborate on the details of these pathways but instead focus on the best characterized antiviral strategies of *Drosophila*: the cGAS-stimulator of interferon genes (Sting) pathway and RNA interference (RNAi).

The cGAS-Sting pathway has recently emerged as an important mechanism to control viral infections in flies (144). During RNA virus replication, double-stranded RNA (dsRNA) accumulates in the cytosol and is detected by cGAS-like receptors (cGLRs). Upon activation, cGLRs synthesize cyclic dinucleotides (CDNs), which initiate a downstream signaling cascade (145-147). To date, three different cGLRs have been identified in *D. melanogaster*, among these cGLR1 and cGLR2 contribute to antiviral immunity by generating up to four different CDNs: 2'3'-cGAMP, 3'2'-cGAMP, 2'3'-c-di-AMP, and 2'3'-c-di-GMP. These CDNs bind to Sting, promoting the translocation of the NF- κ B transcription factor Relish into the nucleus and triggering the transcriptional activation of Sting-regulated genes. Activation of cGAS-Sting signaling reduces viral loads and extends the survival of flies infected with DCV, CrPV, and several other viruses (145, 148, 149). This response appears to be particularly strong in midguts, Malpighian tubules, central nervous system, and genital ducts. Moreover, activation in the fat body or in the midgut has been proposed to be sufficient to elicit a systemic response in distant tissues (150). Notably, Relish also serves as the transcription factor regulating the IMD pathway, suggesting it is capable of orchestrating at least two distinct transcriptional responses.

This raises the possibility that additional, yet-unidentified transcription factors cooperate with Relish to fine-tune the cGAS-Sting response.

RNAi is a central antiviral pathway in plants, nematodes, and arthropods, acting in a sequence-specific manner to silence viral genes (151). Like the cGAS-Sting pathway, RNAi is activated by dsRNA, which is processed by the endonuclease Dicer-2 into small interfering RNAs (siRNAs) that serve as a guide for the endonuclease Argonaute-2 to cleave viral RNA (152). Effective antiviral immunity via RNAi requires the uptake of exogenous dsRNA (153), and it has been proposed that dsRNA spreads from the site of infection to non-infected cells and tissues, inducing a systemic RNAi response (154). Plants and nematodes amplify the RNAi signal through an RdRp that is encoded in their genome (155, 156). In insects, DNA forms of RNA virus sequences are generated by endogenous transposon-encoded reverse transcriptase enzymes. These viral DNAs (vDNAs) can initiate siRNA production and are proposed to persist after infection, priming the organism against future infections (157-159). Moreover, vDNAs synthesized from viral dsRNA phagocytosed by fly hemocytes can be incorporated and secreted via exosome-like vesicles, providing a vehicle for a systemic antiviral RNAi response (160). This mechanism protects uninfected cells and is reminiscent of the immune memory observed in adaptive immunity (161).

As a viral counter-defense against RNAi, diverse viruses have acquired viral suppressors of RNAi (VSRs). In particular, the DCV and CrPV 1A proteins act by binding to dsRNA and preventing siRNA production (162) and by binding to Argonaute-2 and mediating its degradation (163, 164), respectively. VSRs are not restricted to *Dicistroviridae* and are also described in several RNA virus families, such as *Coronaviridae* (165), *Orthomyxoviridae* (166), *Nodaviridae* (167-169), *Birnaviridae* (170) and *Flaviviridae* (171), and even in DNA viruses from the family of *Ascoviridae* (172) and *Iridoviridae* (173). The wide convergent evolution of counter-defense strategies highlights RNAi as a key antiviral pathway and underscores the host's role in shaping long-term virus evolution. Consistent with this view, both VSRs and components of the fly RNAi machinery rank among the most rapidly evolving genes, reflecting strong and ongoing selective pressure (174, 175). Together, these patterns exemplify antagonistic virus-host coevolution under the Red Queen hypothesis, in which interacting species must continually adapt without achieving a persistent fitness advantage (176).

Novel regulators of virus infection

In addition to RdRp-mediated mutagenesis, viral RNA is subjected to RNA editing, which leads to post-transcriptional sequence changes, and to covalent modifications by host enzymes that change nucleotide chemistry without modifying the sequence. These RNA-altering processes can themselves influence the outcome of infection, as exemplified by the APOBEC enzymes (Apolipoprotein B mRNA editing catalytic polypeptide-like) that protect against retroviruses. APOBECs are cytidine deaminases that introduce C-to-U mutations, creating the characteristic G-to-A hyper-mutation pattern that reduces the production of viable progeny (54, 177), while also acting as a source of genetic variation (178). Likewise, ADARs (Adenosine Deaminases that Act on RNA) can edit and hyper-mutate RNA viruses. They act on long dsRNA, deaminating adenosines into inosines, which base-pair like guanosine and result in A-to-G substitutions that alter mRNA decoding (179). ADARs exert context-dependent effects on virus infection, ranging from pro- to antiviral, depending on the specific virus-host combination (180).

RNA molecules also undergo numerous chemical modifications that alter canonical base-pairing and their physicochemical properties (181). Host enzymes that regulate RNA covalent modifications are commonly classified as “writers”, which modify the RNA, “readers”, which recognize the chemical marks, and “erasers”, which reverse the modifications. Interestingly, the genomes of several human-pathogenic viruses have been found to be modified with diverse chemical marks, including HIV-1 (182-184) and influenza A virus (185). However, the functional relevance of RNA modifications during viral infection remains poorly understood.

Thesis outline

The evolution of viruses is strongly affected by an antagonistic arms race with the host immune system. However, disentangling the specific contribution of individual immune pathways in shaping viral populations remains largely unexplored. This thesis seeks to elucidate the role of major antiviral immune pathways in driving viral molecular evolution and pathogenesis. In **Chapter 2**, we modulated the activity of the *Drosophila* RNAi pathway by either inactivating or overexpressing the endonuclease *Dicer-2*. Using this model, we investigated how RNAi activity influences DCV populations in an experimental evolution

setup. We observed an increase in polymorphisms in populations evolved in *Dicer-2* knock-out flies. Unexpectedly, we also detected the repeated selection of several variants carrying a capsid mutation that enhanced viral fitness *in vivo*, irrespective of the RNAi activity of the host. In **Chapter 3**, we assessed the impact of the cGAS-Sting pathway on the molecular evolution of the fly-adapted DCV and the non-native CrPV. We found limited evidence of immune-driven selection in either virus; however, we identified a critical role for the CrPV 2B protein in facilitating cross-species adaptation. Next, in **Chapter 4**, we explored the involvement of RNA-modifying enzymes in RNA virus pathogenesis. Even though we faced technical limitations, we found two putative proviral factors that warrant further investigation. Finally, in the General discussion (**Chapter 5**), I synthesize these findings, explore the dominant forces driving evolution in our experimental system, and reflect on the broader role of viruses in shaping our ecosystem.

References

1. Koonin EV, Starokadomskyy P. Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question. *Studies in History and Philosophy of Biological and Biomedical Sciences* 2016;59:125-34, doi: 10.1016/j.shpsc.2016.02.016.
2. Wasik BR, Turner PE. On the biological success of viruses. *Annual Review of Microbiology*. 2013;67:519-41, doi: 10.1146/annurev-micro-090110-102833.
3. Lustig A, Levine AJ. One hundred years of virology. *Journal of Virology*. 1992;66(8):4629-31, doi: 10.1128/JVI.66.8.4629-4631.1992.
4. Vinson C, Petre A. Mosaic disease of tobacco. *Botanical Gazette*. 1929;87(1):14-38, doi: 10.1086/333922.
5. Stanley WM. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus. *Science*. 1935;81(2113):644-5, doi: 10.1126/science.81.2113.644.
6. Bernal JD, Fankuchen I. X-ray and crystallographic studies of plant virus preparations: I. Introduction and preparation of specimens II. Modes of aggregation of the virus particles. *The Journal of General Physiology*. 1941;25(1):111-46, doi: 10.1085/jgp.25.1.111.
7. Fraenkel-Conrat H, Singer B, Williams R. Infectivity of viral nucleic acid. *Biochimica et Biophysica Acta*. 1957;25:87-96, doi: 10.1016/0006-3002(57)90422-5.
8. Butler P, Klug A. Assembly of the particle of tobacco mosaic virus from RNA and disks of protein. *Nature New Biology*. 1971;229(2):47-50, doi: 10.1038/newbio229047a0.
9. Forterre P. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Research*. 2006;117(1):5-16, doi: 10.1016/j.virusres.2006.01.010.
10. Robertson MP, Joyce GF. The origins of the RNA world. *Cold Spring Harbor Perspectives in Biology*. 2012;4(5):a003608, doi: 10.1101/cshperspect.a003608.
11. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell*. 1982;31(1):147-57, doi: 10.1016/0092-8674(82)90414-7.
12. Paul N, Joyce GF. A self-replicating ligase ribozyme. *Proceedings of the National Academy of Sciences*. 2002;99(20):12733-40, doi: 10.1073/pnas.202471099.
13. Krupovic M, Dolja VV, Koonin EV. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nature Reviews Microbiology*. 2019;17(7):449-58, doi: 10.1038/s41579-019-0205-6.
14. Koonin EV, Dolja VV, Krupovic M. The logic of virus evolution. *Cell Host and Microbe*. 2022;30(7):917-29, doi: 10.1016/j.chom.2022.06.008.
15. Krupovic M, Koonin EV. Multiple origins of viral capsid proteins from cellular ancestors. *Proceedings of the National Academy of Sciences*. 2017;114(12):E2401-E10, doi: 10.1073/pnas.1621061114.
16. Rahbari R, Wuster A, Lindsay SJ, Hardwick RJ, Alexandrov LB, Turki SA, et al. Timing, rates and spectra of human germline mutation. *Nature Genetics*. 2016;48(2):126-33, doi: 10.1038/ng.3469.
17. Sessions SK. *Genome size*. 2nd ed. San Diego (CA): Academic Press; 2013.
18. Martinez-Gutierrez CA, Aylward FO. Genome size distributions in bacteria and archaea are strongly linked to evolutionary history at broad phylogenetic scales. *PLoS Genetics*. 2022;18(5):e1010220, doi: 10.1371/journal.pgen.1010220.

19. Koonin EV, Krupovic M, Yutin N. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. *Annals of the New York Academy of Sciences*. 2015;1341(1):10-24, doi: 10.1111/nyas.12728.
20. Duffy S, Holmes EC. Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. *Journal of General Virology*. 2009;90(6):1539-47, doi: 10.1099/vir.0.009266-0.
21. Neuman BW, Smart A, Gilmer O, Smyth RP, Vaas J, Boker N, et al. Giant RNA genomes: Roles of host, translation elongation, genome architecture, and proteome in nidoviruses. *Proceedings of the National Academy of Sciences*. 2025;122(7):e2413675122, doi: 10.1073/pnas.2413675122.
22. Takada K, Holmes EC. Genome sizes of animal RNA viruses reflect phylogenetic constraints. *Virus Evolution*. 2025;11(1):veaf005, doi: 10.1093/ve/veaf005.
23. Drake JW. Rates of spontaneous mutation among RNA viruses. *Proceedings of the National Academy of Sciences*. 1993;90(9):4171-5, doi: 10.1073/pnas.90.9.4171.
24. Drake JW, Holland JJ. Mutation rates among RNA viruses. *Proceedings of the National Academy of Sciences*. 1999;96(24):13910-3, doi: 10.1073/pnas.96.24.13910.
25. Kawasaki Y, Abe H, Yasuda J. Comparison of genome replication fidelity between SARS-CoV-2 and influenza A virus in cell culture. *Scientific Reports*. 2023;13(1):13105, doi: 10.1038/s41598-023-40463-4.
26. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics*. 1998;148(4):1667-86, doi: 10.1093/genetics/148.4.1667.
27. Mansky LM. Retrovirus mutation rates and their role in genetic variation. *Journal of General Virology*. 1998;79(6):1337-45, doi: 10.1099/0022-1317-79-6-1337.
28. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiology and Molecular Biology Reviews* 2012;76(2):159-216, doi: 10.1128/MMBR.05023-11.
29. Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon- α therapy. *Science*. 1998;282(5386):103-7, doi: 10.1126/science.282.5386.103.
30. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*. 1996;271(5255):1582-6, doi: 10.1126/science.271.5255.1582.
31. Holmes EC. *The evolution and emergence of RNA viruses*. United States: Oxford University Press Inc., New York; 2009.
32. Domingo E, Holland J. RNA virus mutations and fitness for survival. *Annual Review of Microbiology*. 1997;51(1):151-78, doi: 10.1146/annurev.micro.51.1.151.
33. Johnson T, Barton NH. The effect of deleterious alleles on adaptation in asexual populations. *Genetics*. 2002;162(1):395-411, doi: 10.1093/genetics/162.1.395.
34. Sanjuán R. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2010;365(1548):1975-82, doi: 10.1098/rstb.2010.0063.
35. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*. 2008;9(4):267-76, doi: 10.1038/nrg2323.

36. Acevedo A, Brodsky L, Andino R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature*. 2014;505(7485):686-90, doi: 10.1038/nature12861.
37. Elena SF, Sanjuán R. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *Journal of Virology*. 2005;79(18):11555-8, doi: 10.1128/JVI.79.18.11555-11558.2005.
38. Holmes EC. Error thresholds and the constraints to RNA virus evolution. *Trends in Microbiology*. 2003;11(12):543-6, doi: 10.1016/j.tim.2003.10.006.
39. Belshaw R, Gardner A, Rambaut A, Pybus OG. Pacing a small cage: mutation and RNA viruses. *Trends in Ecology and Evolution*. 2008;23(4):188-93, doi: 10.1016/j.tree.2007.11.010.
40. Fitzsimmons WJ, Woods RJ, McCrone JT, Woodman A, Arnold JJ, Yennawar M, et al. A speed-fidelity trade-off determines the mutation rate and virulence of an RNA virus. *PLoS Biology*. 2018;16(6):e2006459, doi: 10.1371/journal.pbio.2006459.
41. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. Nidovirales: evolving the largest RNA virus genome. *Virus Research*. 2006;117(1):17-37, doi: 10.1016/j.virusres.2006.01.017.
42. Minskaia E, Hertzog T, Gorbalenya AE, Campanacci V, Cambillau C, Canard B, et al. Discovery of an RNA virus 3'→ 5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proceedings of the National Academy of Sciences*. 2006;103(13):5108-13, doi: 10.1073/pnas.0508200103.
43. Moeller NH, Shi K, Demir Ö, Belica C, Banerjee S, Yin L, et al. Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. *Proceedings of the National Academy of Sciences*. 2022;119(9):e2106379119, doi: 10.1073/pnas.2106379119.
44. Baddock HT, Brolih S, Yosaatmadja Y, Ratnaweera M, Bielinski M, Swift LP, et al. Characterization of the SARS-CoV-2 ExoN (nsp14ExoN-nsp10) complex: implications for its role in viral genome stability and inhibitor identification. *Nucleic Acids Research*. 2022;50(3):1484-500, doi: 10.1093/nar/gkab1303.
45. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, Lu X, et al. Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS Pathogens*. 2010;6(5):e1000896, doi: 10.1371/journal.ppat.1000896.
46. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *Journal of Virology*. 2007;81(22):12135-44, doi: 10.1128/JVI.01296-07.
47. Smith EC, Blanc H, Vignuzzi M, Denison MR. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathogens*. 2013;9(8):e1003565, doi: 10.1371/journal.ppat.1003565.
48. Nga PT, Parquet Mdel C, Lauber C, Parida M, Nabeshima T, Yu F, et al. Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. *PLoS Pathogens*. 2011;7(9):e1002215, doi: 10.1371/journal.ppat.1002215.
49. Lauber C, Goeman JJ, Parquet Mdel C, Nga PT, Snijder EJ, Morita K, et al. The footprint of genome architecture in the largest genome expansion in RNA viruses. *PLoS Pathogens*. 2013;9(7):e1003500, doi: 10.1371/journal.ppat.1003500.
50. Darwin C. On the origin of the species. London: John Murray; 1859. Available from https://www.vliz.be/docs/Zeecijfers/Origin_of_Species.pdf.
51. Lin JJ, Bhattacharjee MJ, Yu CP, Tseng YY, Li WH. Many human RNA viruses show extraordinarily stringent selective constraints on protein evolution. *Proceedings of the National Academy of Sciences*. 2019;116(38):19009-18, doi: 10.1073/pnas.1907626116.

52. Holmes EC. Patterns of Intra- and Interhost Nonsynonymous Variation Reveal Strong Purifying Selection in Dengue Virus. *Journal of Virology*. 2003;77(20):11296-8, doi: 10.1128/jvi.77.20.11296-11298.2003.
53. Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. *Nature Reviews Microbiology*. 2004;2(4):279-88, doi: 10.1038/nrmicro863.
54. Sanjuan R, Domingo-Calap P. Genetic diversity and evolution of viral populations. *Encyclopedia of virology*. 2021:53, doi: 10.1016/B978-0-12-809633-8.20958-8.
55. Muller HJ. The relation of recombination to mutational advance. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. 1964;1(1):2-9, doi: 10.1016/0027-5107(64)90047-8.
56. Tamisier L, Fabre F, Szadkowski M, Chateau L, Nemouchi G, Girardot G, et al. Within-plant genetic drift to control virus adaptation to host resistance genes. *PLoS Pathogens*. 2024;20(8):e1012424, doi: 10.1371/journal.ppat.1012424.
57. Jonas A, Taus T, Kosiol C, Schlotterer C, Futschik A. Estimating the Effective Population Size from Temporal Allele Frequency Changes in Experimental Evolution. *Genetics*. 2016;204(2):723-35, doi: 10.1534/genetics.116.191197.
58. Nelson CW, Hughes AL. Within-host nucleotide diversity of virus populations: insights from next-generation sequencing. *Infection, Genetics and Evolution*. 2015;30:1-7, doi: 10.1016/j.meegid.2014.11.026.
59. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. *Nature*. 2008;453(7195):615-9, doi: 10.1038/nature06945.
60. Liu Y, Mittler JE. Selection dramatically reduces effective population size in HIV-1 infection. *BMC Ecology and Evolution*. 2008;8:133, doi: 10.1186/1471-2148-8-133.
61. Bedford T, Cobey S, Pascual M. Strength and tempo of selection revealed in viral gene genealogies. *BMC Ecology and Evolution*. 2011;11:220, doi: 10.1186/1471-2148-11-220.
62. Pérez-Losada M, Arenas M, Galán JC, Palero F, González-Candelas F. Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. *Infection, Genetics and Evolution*. 2015;30:296-307, doi: 10.1016/j.meegid.2014.12.022.
63. McDonald SM, Nelson MI, Turner PE, Patton JT. Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nature Reviews Microbiology*. 2016;14(7):448-60, doi: 10.1038/nrmicro.2016.46.
64. Breyer WA, Matthews BW. A structural basis for processivity. *Protein Science*. 2001;10(9):1699-711, doi: 10.1110/ps.10301.
65. Aaziz R, Tepfer M. Recombination in RNA viruses and in virus-resistant transgenic plants. *Journal of General Virology*. 1999;80(6):1339-46, doi: <https://doi.org/10.1099/0022-1317-80-6-1339>
66. Lai MM. RNA recombination in animal and plant viruses. *Microbiological Reviews*. 1992;56(1):61-79, doi: 10.1128/mr.56.1.61-79.1992.
67. Galetto R, Giacomoni V, Véron M, Negroni M. Dissection of a circumscribed recombination hot spot in HIV-1 after a single infectious cycle. *Journal of Biological Chemistry*. 2006;281(5):2711-20, doi: 10.1074/jbc.M505457200.
68. Baird HA, Galetto R, Gao Y, Simon-Loriere E, Abreha M, Archer J, et al. Sequence determinants of breakpoint location during HIV-1 intersubtype recombination. *Nucleic Acids Research*. 2006;34(18):5203-16, doi: 10.1093/nar/gkl669.

69. Neher RA, Leitner T. Recombination rate and selection strength in HIV intra-patient evolution. *PLoS Computational Biology*. 2010;6(1):e1000660, doi: 10.1371/journal.pcbi.1000660.
70. Savolainen-Kopra C, Blomqvist S. Mechanisms of genetic variation in polioviruses. *Reviews in medical virology*. 2010;20(6):358-71, doi: 10.1002/rmv.663.
71. Taucher C, Berger A, Mandl CW. A trans-complementing recombination trap demonstrates a low propensity of flaviviruses for intermolecular recombination. *Journal of Virology*. 2010;84(1):599-611, doi: 10.1128/JVI.01063-09.
72. Reiter J, Pérez-Vilaró G, Scheller N, Mina LB, Díez J, Meyerhans A. Hepatitis C virus RNA recombination in cell culture. *Journal of Hepatology*. 2011;55(4):777-83, doi: 10.1016/j.jhep.2010.12.038.
73. Chare ER, Gould EA, Holmes EC. Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. *Journal of General Virology*. 2003;84(10):2691-703, doi: 10.1099/vir.0.19277-0.
74. Rice WR, Chippindale AK. Sexual recombination and the power of natural selection. *Science*. 2001;294(5542):555-9, doi: 10.1126/science.1061380.
75. Simon-Loriere E, Holmes EC. Why do RNA viruses recombine? *Nature Reviews Microbiology*. 2011;9(8):617-26, doi: 10.1038/nrmicro2614.
76. Domingo E, Perales C. Viral quasispecies. *PLoS Genetics*. 2019;15(10):e1008271, doi: 10.1371/journal.pgen.1008271.
77. Eigen M, Schuster P. A principle of natural self-organization. *Naturwissenschaften*. 1977;64(11):541-65, doi: 10.1007/BF00450633.
78. Eigen M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*. 1971;58(10):465-523, doi: 10.1007/BF00623322.
79. Domingo E, Sabo D, Taniguchi T, Weissmann C. Nucleotide sequence heterogeneity of an RNA phage population. *Cell*. 1978;13(4):735-44, doi: 10.1016/0092-8674(78)90223-4.
80. Wilke CO. Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*. 2005;5(1):44, doi: 10.1186/1471-2148-5-44.
81. Eigen M. On the nature of virus quasispecies. *Trends in Microbiology*. 1996;4(6):216-8, doi: 10.1016/0966-842X(96)20011-3.
82. Sanjuán R, Cuevas JM, Furió V, Holmes EC, Moya A. Selection for robustness in mutagenized RNA viruses. *PLoS Genetics*. 2007;3(6):e93, doi: 10.1371/journal.pgen.0030093.
83. Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*. 2001;412(6844):331-3, doi: 10.1038/35085569.
84. Berger M, Cerf R. Survival of the flattest in the quasispecies model. *Advances in Applied Probability*. 2025;57(2):566-606, doi: <https://doi.org/10.1017/apr.2024.39>.
85. van Nimwegen E, Crutchfield JP, Huynen M. Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences*. 1999;96(17):9716-20, doi: 10.1073/pnas.96.17.9716.
86. Allocca M, Zola S, Bellosta P. The Fruit Fly, *Drosophila melanogaster*: modeling of human diseases (Part II): InTech; 2018. 131-56 p, doi: 10.5772/66545.
87. Rubin GM, Lewis EB. A brief history of *Drosophila*'s contributions to genome research. *Science*. 2000;287(5461):2216-8, doi: 10.1126/science.287.5461.2216.

88. Anderson KV, Bokla L, Nüsslein-Volhard C. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product. *Cell*. 1985;42(3):791-8, doi: 10.1016/0092-8674(85)90275-2.
89. Lemaitre B, Nicolas E, Michaut L, Reichhart J-M, Hoffmann JA. The dorsoventral regulatory gene cassette *spätzle*/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell*. 1996;86(6):973-83, doi: 10.1016/s0092-8674(00)80172-5.
90. Medzhitov R, Preston-Hurlburt P, Janeway Jr CA. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature*. 1997;388(6640):394-7, doi: 10.1038/41131.
91. Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*. 1993;118(2):401-15, doi: 10.1242/dev.118.2.401.
92. McGuire SE, Le PT, Osborn AJ, Matsumoto K, Davis RL. Spatiotemporal rescue of memory dysfunction in *Drosophila*. *Science*. 2003;302(5651):1765-8, doi: 10.1126/science.1089035.
93. Hao L, Sakurai A, Watanabe T, Sorensen E, Nidom CA, Newton MA, et al. *Drosophila* RNAi screen identifies host genes important for influenza virus replication. *Nature*. 2008;454(7206):890-3, doi: 10.1038/nature07151.
94. Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, et al. A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*. 2007;448(7150):151-6, doi: 10.1038/nature05954.
95. Imler JL, Eleftherianos I. *Drosophila* as a model for studying antiviral defences. In: Rolff J, Reynolds S, editors. *Insect infection and immunity: evolution, ecology, and mechanisms*. Oxford: Oxford University Press; 2009. p. 49-68, doi: 10.1093/acprof:oso/9780199551354.003.0004.
96. Xu J, Cherry S. Viruses and antiviral immunity in *Drosophila*. *Developmental and Comparative Immunology*. 2014;42(1):67-84, doi: 10.1016/j.dci.2013.05.002.
97. Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, Quintana JF, et al. The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. *PLoS Biology*. 2015;13(7):e1002210, doi: 10.1371/journal.pbio.1002210.
98. Kapun M, Nolte V, Flatt T, Schlötterer C. Host range and specificity of the *Drosophila* C virus. *PLoS One*. 2010;5(8):e12421, doi: 10.1371/journal.pone.0012421.
99. Le Gall O, Christian P, Fauquet CM, King AM, Knowles NJ, Nakashima N, et al. Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T=3 virion architecture. *Archives of Virology*. 2008;153(4):715-27, doi: 10.1007/s00705-008-0041-x.
100. Valles SM, Chen Y, Firth AE, Guerin DMA, Hashimoto Y, Herrero S, et al. ICTV Virus Taxonomy Profile: Dicistroviridae. *Journal of General Virology*. 2017;98(3):355-6, doi: 10.1099/jgv.0.000756.
101. Jousset F-X, Bergoin M, Revet B. Characterization of the *Drosophila* C virus. *Journal of General Virology*. 1977;34(2):269-83, doi: 10.1099/0022-1317-34-2-269.
102. Johnson KN, Christian PD. The novel genome organization of the insect picorna-like virus *Drosophila* C virus suggests this virus belongs to a previously undescribed virus family. *Journal of General Virology*. 1998;79(1):191-203, doi: 10.1099/0022-1317-79-1-191.
103. Bonning BC, Miller WA. Dicistroviruses. *Annual Review of Entomology*. 2010;55:129-50, doi: 10.1146/annurev-ento-112408-085457.
104. Hertz MI, Thompson SR. Mechanism of translation initiation by Dicistroviridae IGR IRESs. *Virology*. 2011;411(2):355-61, doi: 10.1016/j.virol.2011.01.005.

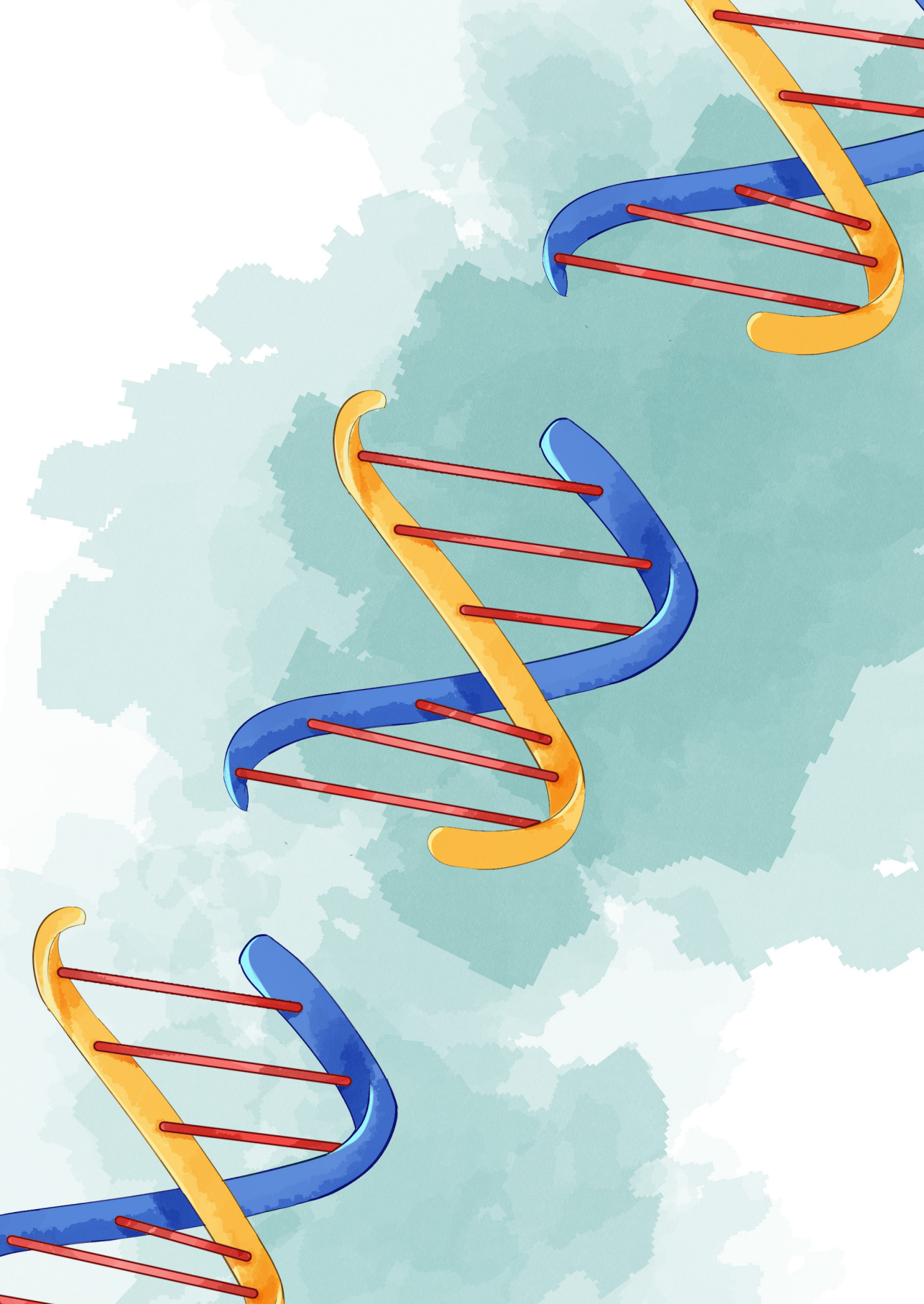
105. Plus N, Croizier G, Reinganum C, Scott PD. Cricket paralysis virus and *Drosophila* C virus: serological analysis and comparison of capsid polypeptides and host range. *Journal of Invertebrate Pathology*. 1978;31(3):296-302, doi: 10.1016/0022-2011(78)90219-7.
106. Petrov A, Grosely R, Chen J, O'Leary SE, Puglisi JD. Multiple parallel pathways of translation initiation on the CrPV IRES. *Molecular Cell*. 2016;62(1):92-103, doi: 10.1016/j.molcel.2016.03.020.
107. Abaeva IS, Young C, Warsaba R, Khan N, Tran LV, Jan E, et al. The structure and mechanism of action of a distinct class of dicistrovirus intergenic region IRESs. *Nucleic Acids Research*. 2023;51(17):9294-313, doi: 10.1093/nar/gkad569.
108. Reinganum C, O'Loughlin G, Hogan T. A nonoccluded virus of the field crickets *Teleogryllus oceanicus* and *T. commodus* (Orthoptera: Gryllidae). *Journal of Invertebrate Pathology*. 1970;16(2):214-20, doi: [https://doi.org/10.1016/0022-2011\(70\)90062-5](https://doi.org/10.1016/0022-2011(70)90062-5).
109. Imrie RM, Roberts KE, Longdon B. Between virus correlations in the outcome of infection across host species: Evidence of virus by host species interactions. *Evolution letters*. 2021;5(5):472-83, doi: 10.1002/evl3.247.
110. Imrie RM, Wallace MA, Longdon B. Positive correlations in susceptibility to a diverse panel of viruses across *Drosophilidae* host species. *Evolution letters*. 2025;9(3):335-45, doi: 10.1093/evlett/qraf002.
111. Ellwanger JH, Chies JAB. Zoonotic spillover: Understanding basic aspects for better prevention. *Genetics and Molecular Biology*. 2021;44(1 Suppl 1):e20200355, doi: 10.1590/1678-4685-GMB-2020-0355.
112. Woolhouse ME, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends in Ecology and Evolution*. 2005;20(5):238-44, doi: 10.1016/j.tree.2005.02.009.
113. Longdon B, Day JP, Alves JM, Smith SCL, Houslay TM, McGonigle JE, et al. Host shifts result in parallel genetic changes when viruses evolve in closely related species. *PLoS Pathogens*. 2018;14(4):e1006951, doi: 10.1371/journal.ppat.1006951.
114. Gutierrez B, Escalera-Zamudio M, Pybus OG. Parallel molecular evolution and adaptation in viruses. *Current Opinion in Virology*. 2019;34:90-6, doi: 10.1016/j.coviro.2018.12.006.
115. Wain LV, Bailes E, Bibollet-Ruche F, Decker JM, Keele BF, Van Heuverswyn F, et al. Adaptation of HIV-1 to its human host. *Molecular Biology and Evolution*. 2007;24(8):1853-60, doi: 10.1093/molbev/msm110.
116. Sauter D, Schindler M, Specht A, Landford WN, Munch J, Kim KA, et al. Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. *Cell Host and Microbe*. 2009;6(5):409-21, doi: 10.1016/j.chom.2009.10.004.
117. Remold SK, Rambaut A, Turner PE. Evolutionary genomics of host adaptation in vesicular stomatitis virus. *Molecular Biology and Evolution*. 2008;25(6):1138-47, doi: 10.1093/molbev/msn059.
118. Liu W, Tang F, Fontanet A, Zhan L, Wang TB, Zhang PH, et al. Molecular epidemiology of SARS-associated coronavirus, Beijing. *Emerg Infect Dis*. 2005;11(9):1420-4, doi: 10.3201/eid1109.040773.
119. Longdon B, Brockhurst MA, Russell CA, Welch JJ, Jiggins FM. The evolution and genetics of virus host shifts. *PLoS Pathogens*. 2014;10(11):e1004395, doi: 10.1371/journal.ppat.1004395.
120. Longdon B, Hadfield JD, Webster CL, Obbard DJ, Jiggins FM. Host phylogeny determines viral persistence and replication in novel hosts. *PLoS Pathogens*. 2011;7(9):e1002260, doi: 10.1371/journal.ppat.1002260.

121. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, et al. Pathways to zoonotic spillover. *Nature Reviews Microbiology*. 2017;15(8):502-10, doi: 10.1038/nrmicro.2017.45.
122. Hoffmann JA, Kafatos FC, Janeway Jr CA, Ezekowitz R. Phylogenetic perspectives in innate immunity. *Science*. 1999;284(5418):1313-8, doi: 10.1126/science.284.5418.1313.
123. Buchmann K. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Frontiers in Immunology*. 2014;5:459, doi: 10.3389/fimmu.2014.00459.
124. Leulier F, Lemaitre B. Toll-like receptors--taking an evolutionary approach. *Nature Reviews Genetics*. 2008;9(3):165-78, doi: 10.1038/nrg2303.
125. Hoffmann JA, Reichhart J-M. *Drosophila* innate immunity: an evolutionary perspective. *Nature Immunology*. 2002;3(2):121-6, doi: 10.1038/ni0202-121.
126. Turvey SE, Brodie DH. Innate immunity. *Journal of Allergy and Clinical Immunology*. 2010;125(2):S24-S32, doi: 10.1016/j.jaci.2009.07.016.
127. Pancer Z, Cooper MD. The evolution of adaptive immunity. *Annual Review of Immunology*. 2006;24(1):497-518, doi: 10.1146/annurev.immunol.24.021605.090542.
128. Iwasaki A, Medzhitov R. Control of adaptive immunity by the innate immune system. *Nature Immunology*. 2015;16(4):343-53, doi: 10.1038/ni.3123.
129. Janeway Jr CA, Medzhitov R. Innate immune recognition. *Annual Review of Immunology*. 2002;20(1):197-216, doi: 10.1146/annurev.immunol.20.083001.084359.
130. Williams MJ. *Drosophila* hemopoiesis and cellular immunity. *The Journal of Immunology*. 2007;178(8):4711-6, doi: 10.4049/jimmunol.178.8.4711.
131. Honti V, Csordás G, Kurucz É, Márkus R, Andó I. The cell-mediated immunity of *Drosophila melanogaster*: hemocyte lineages, immune compartments, microanatomy and regulation. *Developmental and Comparative Immunology*. 2014;42(1):47-56, doi: 10.1016/j.dci.2013.06.005.
132. Lamiable O, Arnold J, de Faria IJdS, Olmo RP, Bergami F, Meignin C, et al. Analysis of the contribution of hemocytes and autophagy to *Drosophila* antiviral immunity. *Journal of Virology*. 2016;90(11):5415-26, doi: 10.1128/JVI.00238-16.
133. Hoffmann JA. The immune response of *Drosophila*. *Nature*. 2003;426(6962):33-8, doi: 10.1038/nature02021.
134. Imler JL. Overview of *Drosophila* immunity: a historical perspective. *Developmental and Comparative Immunology*. 2014;42(1):3-15, doi: 10.1016/j.dci.2013.08.018.
135. Aggarwal K, Silverman N. Positive and negative regulation of the *Drosophila* immune response. *BMB Reports*. 2008;41(4):267-77, doi: 10.5483/bmbrep.2008.41.4.267.
136. Govind S. Innate immunity in *Drosophila*: Pathogens and pathways. *Insect science*. 2008;15(1):29-43, doi: 10.1111/j.1744-7917.2008.00185.x.
137. Myllymäki H, Rämetsä M. JAK/STAT pathway in *Drosophila* immunity. *Scandinavian Journal of Immunology*. 2014;79(6):377-85, doi: 10.1111/sji.12170.
138. Agaisse H, Perrimon N. The roles of JAK/STAT signaling in *Drosophila* immune responses. *Immunological reviews*. 2004;198(1):72-82, doi: 10.1111/j.0105-2896.2004.0133.x.
139. Lamiable O, Imler JL. Induced antiviral innate immunity in *Drosophila*. *Current Opinion in Microbiology*. 2014;20:62-8, doi: 10.1016/j.mib.2014.05.006.
140. Sabin LR, Hanna SL, Cherry S. Innate antiviral immunity in *Drosophila*. *Current Opinion in Immunology*. 2010;22(1):4-9, doi: 10.1016/j.coi.2010.01.007.

141. Mussabekova A, Daeffler L, Imler JL. Innate and intrinsic antiviral immunity in *Drosophila*. *Cellular and Molecular Life Sciences*. 2017;74(11):2039-54, doi: 10.1007/s00018-017-2453-9.
142. Huang Z, Kingsolver MB, Avadhanula V, Hardy RW. An antiviral role for antimicrobial peptides during the arthropod response to alphavirus replication. *Journal of Virology*. 2013;87(8):4272-80, doi: 10.1128/JVI.03360-12.
143. Dostert C, Jouanguy E, Irving P, Troxler L, Galiana-Arnoux D, Hetru C, et al. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. *Nature Immunology*. 2005;6(9):946-53, doi: 10.1038/ni1237.
144. Cai H, Meignin C, Imler JL. cGAS-like receptor-mediated immunity: the insect perspective. *Current Opinion in Immunology*. 2022;74:183-9, doi: 10.1016/j.coi.2022.01.005.
145. Cai H, Li L, Slavik K, Huang J, Yin T, Hedelin L, et al. The virus-induced cyclic dinucleotide 2' 3'-c-di-GMP mediates STING-dependent antiviral immunity in *Drosophila*. *Immunity*. 2023;56(9):1991-2005. e9, doi: 10.1016/j.immuni.2023.08.006.
146. Holleufer A, Winther KG, Gad HH, Ai X, Chen Y, Li L, et al. Two cGAS-like receptors induce antiviral immunity in *Drosophila*. *Nature*. 2021;597(7874):114-8, doi: 10.1038/s41586-021-03800-z.
147. Slavik KM, Morehouse BR, Ragucci AE, Zhou W, Ai X, Chen Y, et al. cGAS-like receptors sense RNA and control 3' 2'-cGAMP signalling in *Drosophila*. *Nature*. 2021;597(7874):109-13, doi: 10.1038/s41586-021-03743-5.
148. Goto A, Okado K, Martins N, Cai H, Barbier V, Lamiable O, et al. The kinase IKK β regulates a STING-and NF- κ B-dependent antiviral response pathway in *Drosophila*. *Immunity*. 2018;49(2):225-34. e4, doi: 10.1016/j.immuni.2019.12.009.
149. Cai H, Holleufer A, Simonsen B, Schneider J, Lemoine A, Gad HH, et al. 2' 3'-cGAMP triggers a STING-and NF- κ B-dependent broad antiviral response in *Drosophila*. *Science Signaling*. 2020;13(660): eabc4537, doi: 10.1126/scisignal.abc4537.
150. Ai X, Deng H, Li X, Wei Z, Chen Y, Yin T, et al. cGAS-like receptors drive a systemic STING-dependent host response in *Drosophila*. *Cell Reports*. 2024;43(12):115081, doi: 10.1016/j.celrep.2024.115081.
151. Guo Z, Li Y, Ding S-W. Small RNA-based antimicrobial immunity. *Nature Reviews Immunology*. 2019;19(1):31-44, doi: 10.1038/s41577-018-0071-x.
152. Gammon DB, Mello CC. RNA interference-mediated antiviral defense in insects. *Current Opinion in Insect Science*. 2015;8:111-20, doi: 10.1016/j.cois.2015.01.006.
153. Saleh MC, Tassetto M, van Rij RP, Goic B, Gausson V, Berry B, et al. Antiviral immunity in *Drosophila* requires systemic RNA interference spread. *Nature*. 2009;458(7236):346-50, doi: 10.1038/nature07712.
154. Karlikow M, Goic B, Saleh M-C. RNAi and antiviral defense in *Drosophila*: setting up a systemic immune response. *Developmental and Comparative Immunology*. 2014;42(1):85-92, doi: 10.1016/j.dci.2013.05.004.
155. Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, et al. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell*. 2001;107(4):465-76, doi: 10.1016/s0092-8674(01)00576-1.
156. Himber C, Dunoyer P, Moissiard G, Ritzenthaler C, Voinnet O. Transitivity-dependent and-independent cell-to-cell movement of RNA silencing. *The EMBO journal*. 2003;22, 4523-33, doi: 10.1093/emboj/cdg431.

157. Goic B, Stapleford KA, Frangeul L, Doucet AJ, Gausson V, Blanc H, et al. Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nature Communications*. 2016;7(1):12410, doi: 10.1038/ncomms12410.
158. Goic B, Vodovar N, Mondotte JA, Monot C, Frangeul L, Blanc H, et al. RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. *Nature Immunology*. 2013;14(4):396-403, doi: 10.1038/ni.2542.
159. Mondotte JA, Gausson V, Frangeul L, Blanc H, Lambrechts L, Saleh M-C. Immune priming and clearance of orally acquired RNA viruses in *Drosophila*. *Nature Microbiology*. 2018;3(12):1394-403, doi: 10.1038/s41564-018-0265-9.
160. Tassetto M, Kunitomi M, Andino R. Circulating immune cells mediate a systemic RNAi-based adaptive antiviral response in *Drosophila*. *Cell*. 2017;169(2):314-25. e13, doi: 10.1016/j.cell.2017.03.033.
161. Bonning BC, Saleh M-C. The interplay between viruses and RNAi pathways in insects. *Annual Review of Entomology*. 2021;66(1):61-79, doi: 10.1146/annurev-ento-033020-090410.
162. van Rij RP, Saleh MC, Berry B, Foo C, Houk A, Antoniewski C, et al. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes and development*. 2006;20(21):2985-95, doi: 10.1101/gad.1482006.
163. Nayak A, Kim DY, Trnka MJ, Kerr CH, Lidsky PV, Stanley DJ, et al. A viral protein restricts *Drosophila* RNAi immunity by regulating Argonaute activity and stability. *Cell Host and Microbe*. 2018;24(4):542-57. e9, doi: 10.1016/j.chom.2018.09.006.
164. Watanabe M, Iwakawa H-o, Tadakuma H, Tomari Y. Biochemical and single-molecule analyses of the RNA silencing suppressing activity of CrPV-1A. *Nucleic Acids Research*. 2017;45(18):10837-44, doi: 10.1093/nar/gkx748.
165. Cui L, Wang H, Ji Y, Yang J, Xu S, Huang X, et al. The nucleocapsid protein of coronaviruses acts as a viral suppressor of RNA silencing in mammalian cells. *Journal of Virology*. 2015;89(17):9029-43, doi: 10.1128/JVI.01331-15.
166. Li Y, Basavappa M, Lu J, Dong S, Cronkite DA, Prior JT, et al. Induction and suppression of antiviral RNA interference by influenza A virus in mammalian cells. *Nature Microbiology*. 2016;2(3):1-9, doi: 10.1038/nmicrobiol.2016.250.
167. Chao JA, Lee JH, Chapados BR, Debler EW, Schneemann A, Williamson JR. Dual modes of RNA-silencing suppression by Flock House virus protein B2. *Nature Structural and Molecular Biology*. 2005;12(11):952-7, doi: 10.1038/nsmb1005.
168. Qi N, Zhang L, Qiu Y, Wang Z, Si J, Liu Y, et al. Targeting of dicer-2 and RNA by a viral RNA silencing suppressor in *Drosophila* cells. *Journal of Virology*. 2012;86(10):5763-73, doi: 10.1128/JVI.07229-11.
169. Sullivan CS, Ganem D. A virus-encoded inhibitor that blocks RNA interference in mammalian cells. *Journal of Virology*. 2005;79(12):7371-9, doi: 10.1128/JVI.79.12.7371-7379.2005.
170. van Cleef KW, van Mierlo JT, Miesen P, Overheul GJ, Fros JJ, Schuster S, et al. Mosquito and *Drosophila* entomobirnaviruses suppress dsRNA-and siRNA-induced RNAi. *Nucleic Acids Research*. 2014;42(13):8732-44, doi: 10.1093/nar/gku528.
171. Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, et al. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells. *Journal of Virology*. 2012;86(24):13486-500, doi: 10.1128/JVI.01104-12.
172. Hussain M, Abraham AM, Asgari S. An Ascovirus-encoded RNase III autoregulates its expression and suppresses RNA interference-mediated gene silencing. *Journal of Virology*. 2010;84(7):3624-30, doi: 10.1128/JVI.02362-09.

173. Bronkhorst AW, Van Cleef KW, Venselaar H, Van Rij RP. A dsRNA-binding protein of a complex invertebrate DNA virus suppresses the *Drosophila* RNAi response. *Nucleic Acids Research*. 2014;42(19):12237-48, doi: 10.1093/nar/gku910.
174. Obbard DJ, Gordon KH, Buck AH, Jiggins FM. The evolution of RNAi as a defence against viruses and transposable elements. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2009;364(1513):99-115, doi: 10.1098/rstb.2008.0168.
175. Obbard DJ, Jiggins FM, Halligan DL, Little TJ. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Current Biology*. 2006;16(6):580-5, doi: 10.1016/j.cub.2006.01.065.
176. Van Valen L. A new evolutionary law. *Evolutionary Theory* 1973;1:1-30.
177. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology*. 2004;4(11):868-77, doi: 10.1038/nri1489.
178. Jonathan M, Ikeda T. APOBEC3 family proteins as drivers of virus evolution. *Frontiers in Virology*. 2023;3:1332010, doi: <https://doi.org/10.3389/fviro.2023.1332010>.
179. Pfaller CK, George CX, Samuel CE. Adenosine deaminases acting on RNA (ADARs) and viral infections. *Annual Review of Virology*. 2021;8(1):239-64, doi: 10.1146/annurev-virology-091919-065320.
180. Piontkivska H, Wales-McGrath B, Miyamoto M, Wayne ML. ADAR editing in viruses: an evolutionary force to reckon with. *Genome Biology and Evolution*. 2021;13(11):evab240, doi: 10.1093/gbe/evab240.
181. Cappannini A, Ray A, Purta E, Mukherjee S, Boccaletto P, Moafinejad SN, et al. MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Research*. 2024;52(D1):D239-D44, doi: 10.1093/nar/gkad1083.
182. Bosmeny MS, Pater AA, Zhang L, Larkai L, Sha BE, Lyu Z, et al. An HIV-1 Reference Epitranscriptome. *bioRxiv*. 2025 [Preprint], doi: 10.1101/2025.01.30.635805.
183. Baek A, Lee GE, Golconda S, Rayhan A, Manganaris AA, Chen S, et al. Single-molecule epitranscriptomic analysis of full-length HIV-1 RNAs reveals functional roles of site-specific m(6)As. *Nature Microbiology*. 2024;9(5):1340-55, doi: 10.1038/s41564-024-01638-5.
184. Courtney DG, Tsai K, Bogerd HP, Kennedy EM, Law BA, Emery A, et al. Epitranscriptomic Addition of m(5)C to HIV-1 Transcripts Regulates Viral Gene Expression. *Cell Host and Microbe*. 2019;26(2):217-27 e6, doi: 10.1016/j.chom.2019.07.005.
185. Courtney DG, Kennedy EM, Dumm RE, Bogerd HP, Tsai K, Heaton NS, et al. Epitranscriptomic Enhancement of Influenza A Virus Gene Expression and Replication. *Cell Host and Microbe*. 2017;22(3):377-86 e5, doi: 10.1016/j.chom.2017.08.004.



Chapter 2

Parallel evolution and enhanced virulence upon *in vivo* passage of an RNA virus in *Drosophila melanogaster*

Oscar M. Lezcano^{1*}, Lara Fuhrmann^{2,3*}, Gayatri Ramakrishnan⁴,
Niko Beerenwinkel^{2,3}, Martijn A. Huynen⁴, Ronald P. van Rij¹

¹ Department of Medical Microbiology, Radboud University Medical Center, P.O. Box 9101, 6500 HB, Nijmegen, the Netherlands

² Department of Biosystems Science and Engineering, ETH Zurich, Klingelbergstrasse 48, Basel, 4056, Switzerland

³ SIB Swiss Institute of Bioinformatics, Quartier Sorge - Bâtiment Amphipôle, Lausanne, 1015, Switzerland

⁴ Department of Medical BioSciences, Radboud University Medical Center, P.O. Box 9101, 6500 HB, Nijmegen, the Netherlands

*These authors contributed equally to this study.

Corresponding authors: niko.beerenwinkel@bsse.ethz.ch; martijn.huijnen@radboudumc.nl;
ronald.vanrij@radboudumc.nl

Published in:

Virus Evol. 2023; 15;9(2):vead074; doi: 10.1093/ve/vead074.

Abstract

Virus evolution is strongly affected by antagonistic co-evolution of virus and host. Host immunity positively selects for viruses that evade the immune response, which in turn may drive counter-adaptations in host immune genes. We investigated how host immune pressure shapes virus populations, using the fruit fly *Drosophila melanogaster* and its natural pathogen Drosophila C virus (DCV), as a model. We performed an experimental evolution study in which DCV was serially passaged for ten generations in three fly genotypes differing in their antiviral RNAi response: wild-type flies and flies in which the endonuclease gene *Dicer-2* was either overexpressed or inactivated. All evolved virus populations replicated more efficiently *in vivo* and were more virulent than the parental stock. The number of polymorphisms increased in all three host genotypes with passage number, which was most pronounced in *Dicer-2* knockout flies. Mutational analysis showed strong parallel evolution, as mutations accumulated in a specific region of the VP3 capsid protein in every lineage in a host genotype-independent manner. The parental tyrosine at position 95 of VP3 was substituted with either one of five different amino acids in 14 out of 15 lineages. However, no consistent amino acid changes were observed in the viral RNAi suppressor gene 1A, nor elsewhere in the genome in any of the host backgrounds. Our study indicates that the RNAi response restricts the sequence space that can be explored by viral populations. Moreover, our study illustrates how evolution towards higher virulence can be a highly reproducible, yet unpredictable process.

Introduction

Viruses are ubiquitous parasites of cellular life and the most abundant biological entities on earth, predicted to appear in any replicator system (1–3). Virus-host coevolution is often understood as an arms race between host defense and viral counter-defense mechanisms that occurs in all host systems (4). For example, the interferon response, one of the most potent antiviral mechanisms in mammals, can be antagonized at the transcriptional, translational, or functional level by many different viruses (5).

In insects, RNA interference (RNAi) plays a critical role in antiviral defense, as viral double stranded RNA (dsRNA) is processed by the endonuclease Dicer-2 (Dcr-2) into small interfering RNAs (siRNAs), serving as guides for the endonuclease Argonaute-2 (AGO2) to mediate viral RNA degradation (6). Several insect viruses have been shown to suppress this pathway, illustrating how immune pressure exerted by RNAi affects the genetic makeup of a virus (7, 8). For instance, the *Drosophila* C virus (DCV) 1A protein binds double-stranded RNA (dsRNA), interacts with Dcr-2, and inhibits Dcr-2 nuclease activity (9,10), whereas proteins from two distinct RNA viruses, the cricket paralysis virus 1A protein and the nora virus VP1 protein bind and inhibit AGO2 function (10–13).

Upon infection, RNA viruses create swarms of closely related within-host genotypes, often referred to as viral quasispecies (14), that contribute to the adaptability of a viral population to changing environments (15, 16). They are the result of the high mutation rate of the viral RNA dependent RNA polymerase (RdRp) and the large population sizes of RNA viruses, which can be as high as 10^{12} infectious particles in an infected host (17, 18). Most mutations in this diverse virus population will be deleterious, some neutral, and very few beneficial (19, 20), and the frequency of these variants will fluctuate according to evolutionary forces, such as the positive and purifying selection imposed by the host immune system. New virus variants that are capable of escaping immune surveillance have an evolutionary advantage and will likely take over the ancestral population. In accordance, plants with a more restrictive immune system have been shown to drive faster evolution of turnip mosaic virus (21).

In this study, we assessed to what extent selection pressure by the immune system drives virus population evolution. We used as a model system the fruit fly *Drosophila melanogaster* infected with its natural pathogen DCV, a

single-stranded positive-sense RNA virus of the *Dicistroviridae* family (genus *Cripavirus*). The virus was serially passaged in wild-type flies and in fly mutants in which *Dcr-2* was either overexpressed or inactivated. Afterwards, the evolved virus populations were characterized using next-generation sequencing (NGS) in combination with phenotypic assays for virus replication and host survival. We found that the increase in the number of polymorphisms during passaging was most pronounced in *Dicer-2* knockout flies. However, the RNAi response had little effect on the accumulation of specific amino acid changes in viral proteins. Instead, we observed parallel evolution of mutations in the VP3 capsid protein in all viral lineages, which markedly enhanced viral replication and virulence *in vivo*.

Results

Experimental DCV evolution increases virulence

To investigate the effect of the host immune system on DCV evolution, we performed an experimental evolution study in which a parental DCV stock was serially passaged over three fly genotypes with different RNAi activity (Fig. 1A). Specifically, we used wild-type flies (w^{1118} , hereafter WT), flies with a frameshift mutation in *Dcr-2* resulting in a premature stop codon (*Dcr-2*^{L811fsX} (22), hereafter *Dcr-2-KO*), and flies overexpressing a *Dcr-2* transgene controlled by the UAS upstream activation sequence and a *Tubulin-Gal4* driver (hereafter *Dcr-2-OE*). As expected (23, 24), *Dcr-2* expression correlated with survival after DCV infection, with *Dcr-2-KO* flies having the shortest life span and *Dcr-2-OE* flies the longest (Fig. 1B).

For each host genotype, we allowed virus populations to independently evolve during ten serial passages, with five independent lineages per genotype. The parental DCV stock was cultured in *Drosophila* S2 cells, and its genetic diversity had been reduced by three rounds of serial dilution prior to the experimental evolution study. To prevent population bottlenecks, flies were infected by intrathoracic injection with a high viral inoculum of 10^4 median tissue culture infectious doses (TCID₅₀), and virus titers were measured after each passage to normalize the inoculation of naive flies for the next passage (Fig. 1C). With the exception of passage 3 and 5, where we observed significantly higher viral titers in *Dcr-2-KO* than in WT flies ($P < 0.05$, *t*-test, Fig. 1C), the differences in viral titers between genotypes were generally small and not consistent along passages, likely because flies were harvested at the plateau of peak titers.

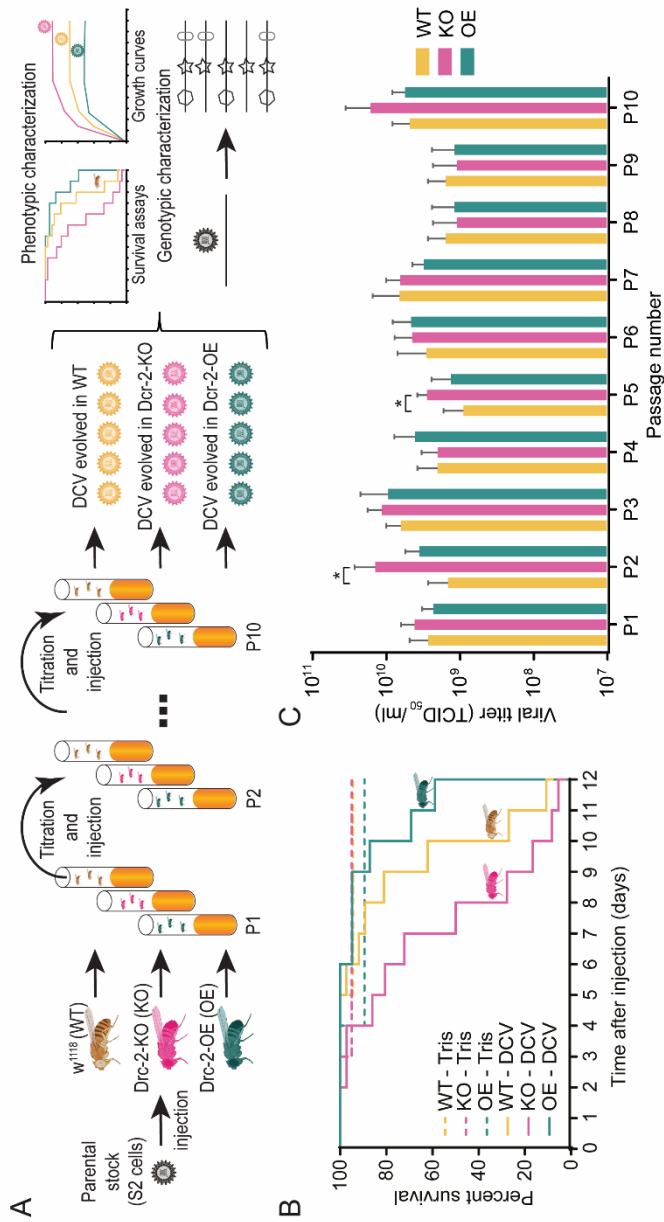


Figure 1. Experimental evolution of DCV in flies differing in their RNAi response. (A) A parental stock of DCV Charolles strain was serially passaged in wild-type flies (*w¹¹¹⁸*; WT), *Dicer-2* knockout flies (*Dcr-2-KO*), and flies overexpressing *Dicer-2* (*Dcr-2-OE*) for ten fly generations with five independent lineages per genotype. Flies were inoculated by intrathoracic injections and titers were determined after every passage. Evolved virus populations were characterized phenotypically using growth curves and survival assays and genotypically by next-generation sequencing. (B) Survival curves of WT, *Dcr-2-KO* and *Dcr-2-OE* flies inoculated with 1,000 TCID₅₀ of DCV. Two to five-day-old female WT flies were inoculated intrathoracically (n = 30) with the parental stock and survival was monitored daily. (C) Viral titers of all lineages at each viral passage. Bars represent means and SD of the five viral lineages. Asterisks indicate *P* < 0.05, t-test. Fly and virus icons were retrieved from BioRender. TCID₅₀, median tissue culture infectious dose.

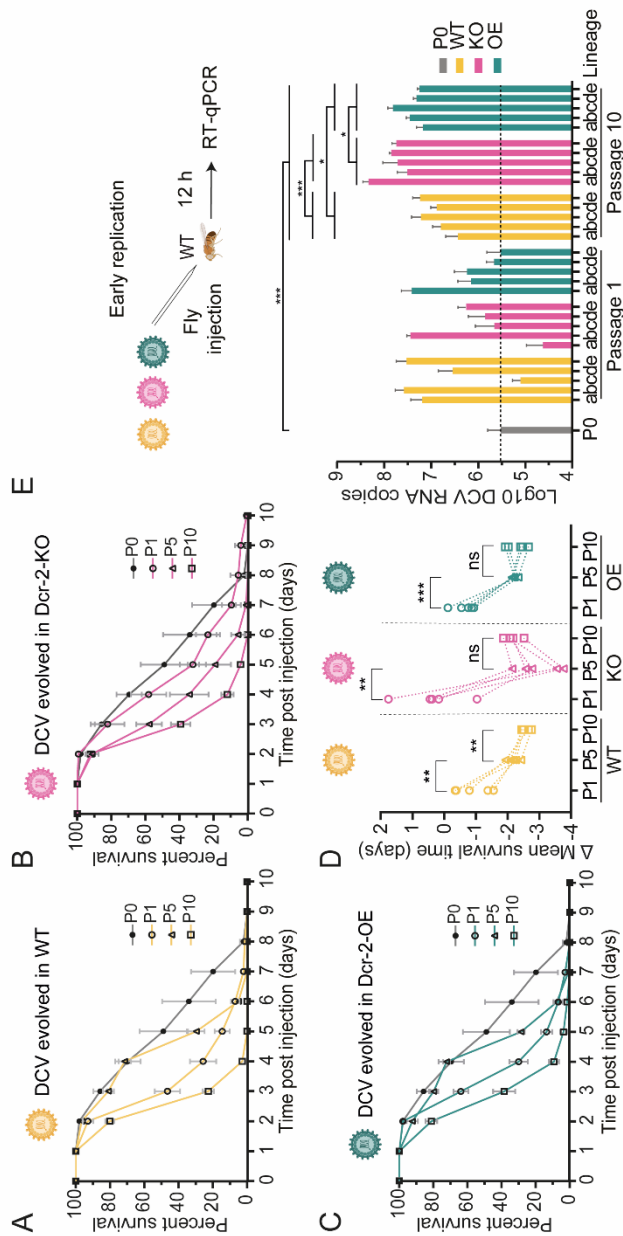


Figure 2. Increased virulence of evolved virus populations independent of the host genotype. (A–C) Survival curves of WT flies inoculated with 10,000 TCID₅₀ of the parental DCV stock (P0) and virus populations from the indicated host genotypes at passage 1, 5 or 10. Two to five-day-old female WT flies (n = 30) were inoculated intrathoracically and survival was monitored daily. Results are shown as means and SEM of all five lineages for virus populations from (A) WT, (B) *Dcr-2-OE*, and (C) *Dcr-2-KO* flies. Data for the parental stock are shown as means and SEM of four independent experiments. Results from Cox regressions are shown in Supplementary Table S1. (D) Difference of the mean survival time for each virus stock relative to the parental stock. Dashed lines connect the values of each lineage. Asterisks indicate significance: **, $P < 0.01$; ***, $P < 0.001$, paired *t*-test. (E) Early replication kinetics of parental and evolved virus populations. WT flies were inoculated with 1,000 TCID₅₀ of each virus stock and three groups of five flies were collected at 12 hours post infection (hpi). Viral RNA copy numbers were quantified by RT-qPCR. The dashed line shows the RNA copies of the parental stock. Results are shown as means and SD of three replicates. Viral RNA levels of the P10 virus populations and the parental stock were compared using a nested ANOVA followed by Tukey's HSD test for multiple comparison. Adjusted *P*-values of Tukey's HSD test are indicated with asterisks: *, $P < 0.05$; ***, $P < 0.001$.

To examine the virulence of the evolved virus populations, we performed survival assays after inoculation of WT flies with 10^4 TCID₅₀ of the parental stock and virus stocks harvested after passage 1, 5 and 10 (P1, P5, P10; Fig. 2A-D). P10 virus populations from all lineages of all host genotypes induced higher mortality in WT flies than the parental stock (Log-rank test, $P < 0.0001$; hazard ratios > 3 , Supplementary Table S1), with mean survival times decreasing from 5.5 days for the parental stock to 3.4 days for the P10 virus populations (average of all genotypes and lineages). The mean survival time of WT flies decreased from P1 to P5 for virus populations from all genotypes (paired t -test, $P < 0.005$) and further decreased from P5 to P10 only for virus populations from WT flies (paired t -test, $P = 0.006$) (Fig. 2D).

We next analyzed whether these survival phenotypes correlated with changes in replication kinetics of the evolved virus populations. To test this, we inoculated WT flies with the parental, P1, and P10 virus stocks and determined viral RNA levels at 12 hours post infection (hpi) (Fig. 2E). We observed a large variation in viral RNA replication between the independent lineages of all genotypes at P1, with the parental stock accumulating 3.4×10^5 RNA copies and the P1 viruses ranging from 4.2×10^4 to 3.9×10^7 RNA copies. Irrespective of the genotype, all lineages of the P10 virus populations reached higher viral RNA levels (range 2.8×10^6 to 2.2×10^8 RNA copies) than the parental stock (nested ANOVA with significant effect of genotypes on the viral RNA level: $P < 0.001$, followed by Tukey's HSD test with adjusted P -value: P -adj. < 0.0001). This phenotype was maintained when the infection was performed orally (Supplementary Fig. S1), indicating that our experimental evolution scheme did not lead to loss of infectivity via the presumed natural infection route. Among genotypes, P10 virus populations from *Dcr-2-KO* flies reached the highest RNA loads, compared to those evolved in WT and *Dcr-2-OE* flies (P -adj. = 0.0003 and P -adj. = 0.0468 Tukey's HSD test), where those evolved in *Dcr-2-OE* flies reached higher RNA loads compared to WT (P -adj. = 0.0374 Tukey's HSD test) (Fig. 2E). Together, these results indicate that *in vivo* serial passage increases replication kinetics and virulence of the evolved virus populations.

Ongoing accumulation of polymorphisms

We next investigated how the virus populations evolved through the serial passage experiment by next-generation sequencing (NGS). Viral genomes were sequenced to high coverage (average coverage of 127,542, Supplementary Fig. S2), allowing us to confidently call low-frequency

mutations, as the median Phred sequencing quality score for the mutation calls was 39, corresponding to an error rate of 0.0001. We observed increasing numbers of mutations during passaging with a mean of 381 polymorphic sites in the evolved virus populations after ten passages, 116 more than in the parental stock (Fig. 3A). Between 179 and 499 single nucleotide variants (SNVs) per evolved population occurred at a frequency below 1%, and between 5 and 21 SNVs at a frequency higher than 1% (Fig. 3B). The number of polymorphic sites increased significantly through the serial passage in *Dcr-2-KO* and *Dcr-2-OE* flies ($P < 0.05$, paired t -test, Fig. 3A). This increase was most pronounced in virus populations from *Dcr-2-KO* flies, which accumulated a mean of 444 polymorphic sites at passage 10 (Fig. 3A), which was significantly different from virus populations from WT and *Dcr-2-OE* flies ($P < 0.02$, t -test). Yet, the vast majority of polymorphic sites were low-frequency mutations (median: 0.1%), and very few SNVs increased in frequency or reached fixation (Fig. 3B).

We next analyzed virus population diversity over time and found that the mean nucleotide diversity did not change significantly over time in any of the host genotypes, nor were there significant differences in population diversity between genotypes ($P > 0.05$, mixed ANOVA, Supplementary Fig. S3A). Mean nucleotide diversity was significantly higher in coding than in non-coding regions of the viral genome ($P < 0.01$, mixed ANOVA, Supplementary Fig. S3A). Specifically, the gene encoding the viral VP3 protein displayed a significant increase in nucleotide diversity compared to the mean diversity of the entire coding region ($P < 0.05$, paired t -test, Supplementary Fig. S3B).

Moderate host genotype-specific adaptation in *Dicer-2* mutants

We next assessed the shared and unique SNVs across genotypes, first considering SNVs with frequencies $\geq 0.01\%$, and found a similar number of unique SNVs between virus populations from WT and *Dcr-2-OE* flies ($n = 813$ and 910 , respectively), both having less SNVs than virus populations from *Dcr-2-KO* flies ($n = 1189$) (Fig. 3C). 15.7% of the SNVs were shared between virus populations from all three fly genotypes, of which 41.07% were non-synonymous mutations. When restricting the analysis to high-frequency SNVs ($\geq 10\%$), the fraction of SNVs shared by the virus populations from all three fly genotypes increased to 21.3%, all of which were non-synonymous mutations (Fig. 3D).

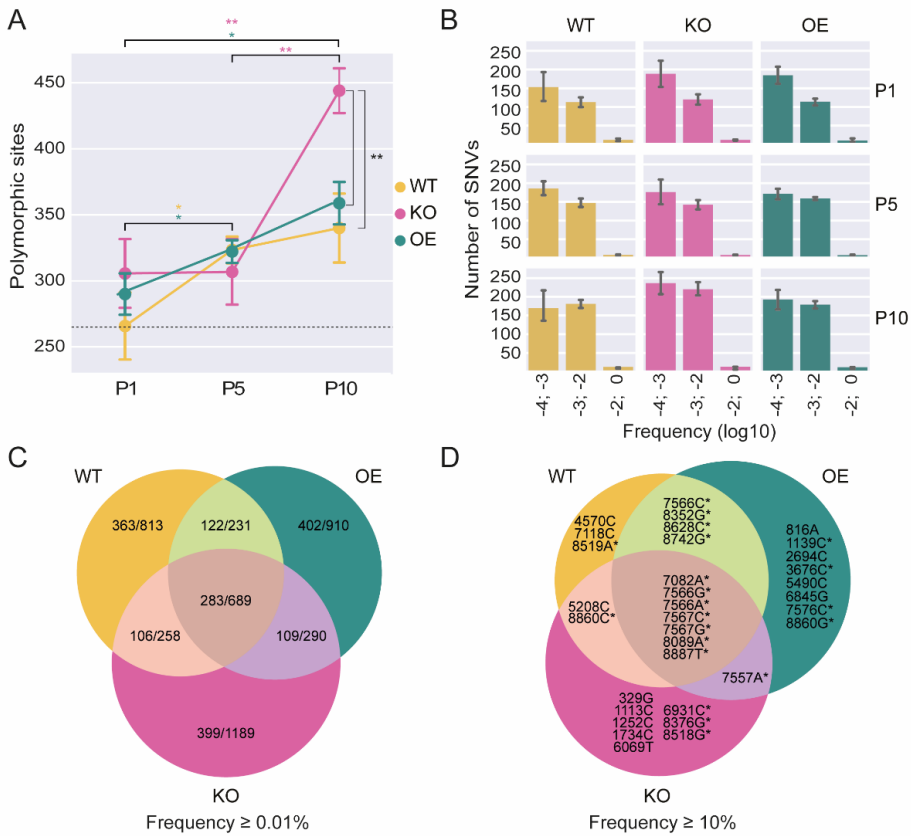


Figure 3. Increasing genetic variation during serial passage. **(A)** Number of polymorphic sites of the virus populations from WT, *Dicer-2-OE* and *Dicer-2-KO* flies across passages. Data are shown as means and SEM across the five lineages for each genotype. The dashed line marks the number of polymorphic sites in the parental stock with respect to the Charolles reference sequence (GenBank Accession Number: MK645242.1). **(B)** Frequencies of single-nucleotide variants (SNVs) in the indicated bins for virus populations from the indicated genotypes. Error bars indicate 95% confidence intervals based on the bootstrap distribution. **(C-D)** Venn diagram of the number of unique and shared SNVs between virus populations from WT, *Dicer-2-OE* and *Dicer-2-KO* flies. Per genotype, SNVs from all lineages and passages with a minimum observed frequency of 0.01% (C) or 10% (D) were combined. The ratios in (C) indicate the number of non-synonymous mutations to all mutations. Asterisks in (D) indicate non-synonymous mutations.

To understand if the observed mutation patterns are the result of host-specific adaptation, we conducted a permutation test, as described by Bons et al. (25). This test assesses whether SNVs occur at higher rates in specific genotypes than expected under a null hypothesis that SNVs are equally likely to occur in each genotype. For virus populations from WT and *Dcr-2-OE* flies, we did not find significant differences between the expected and observed number of

SNVs. In contrast, 131 SNVs were unique to virus populations from *Dcr-2-KO* flies, which was significantly different from the expected number by random allocation ($n = 89$; $P < 0.001$, Supplementary Fig. S4). Together, these data indicate that viral populations are predominantly shaped by host-genotype independent, parallel evolution of shared SNVs as well as moderate host-genotype specific adaptation in the absence of a functional RNAi response.

Deletions in homopolymeric uridine tracts

To explore the evolution of the viral populations in more detail, we visualized the SNVs across the DCV genome in a heatmap (Fig. 4A). We observed some SNVs that were already present in the parental stock and were retained in the viral populations throughout the experiment in most of the lineages (nt 329, 504, 2686, 2736, 7082, 7458, 8089 and 8526), while others appeared *de novo* during the experiment in different lineages (e.g., nt 331, 1026, 4758, 6263 and 7557). Strikingly, we found a ubiquitous deletion of a uridine in a poly(uridine) tract at position 5765–5774, which gives rise to a frameshift and a premature stop codon in the *RdRp* gene. This deletion was also present in the parental stock and it was retained at low, yet fairly constant frequency throughout the experiment in all 15 lineages (around 4%, Supplementary Table S2). Likewise, another uridine deletion in a tract of 8 uridines at position 276–284 in the 5' untranslated region (UTR) occurred in all 15 lineages with a constant frequency of around 3% (Supplementary Table S2). To control for possible sequencing artifacts, we evaluated if the two uridine deletions were located at the edges of the sequence reads containing them, as there is lower confidence in the base calling at read edges. However, in only 7% of reads the deletions were located in proximity to the read edges (within five nucleotides). Furthermore, the mean sequencing quality score in the uridine tracks and the adjacent positions was above 40 both in reads with and without deletion, corresponding to an error rate of 0.0001 (mean quality scores across uridine tracks: 40.1; 10%-quantile: 38.2, 90%-quantile: 41.0; Supplementary Table S3). Further analysis suggested that such deletions also occur in other, published DCV NGS datasets, and we suggest that the viral RdRp is prone to slippage on homopolymeric uridine tracts and that there is purifying selection against the occurrence of these tracts in RNA viruses (Supplementary text S1, Supplementary Table S4–S6).

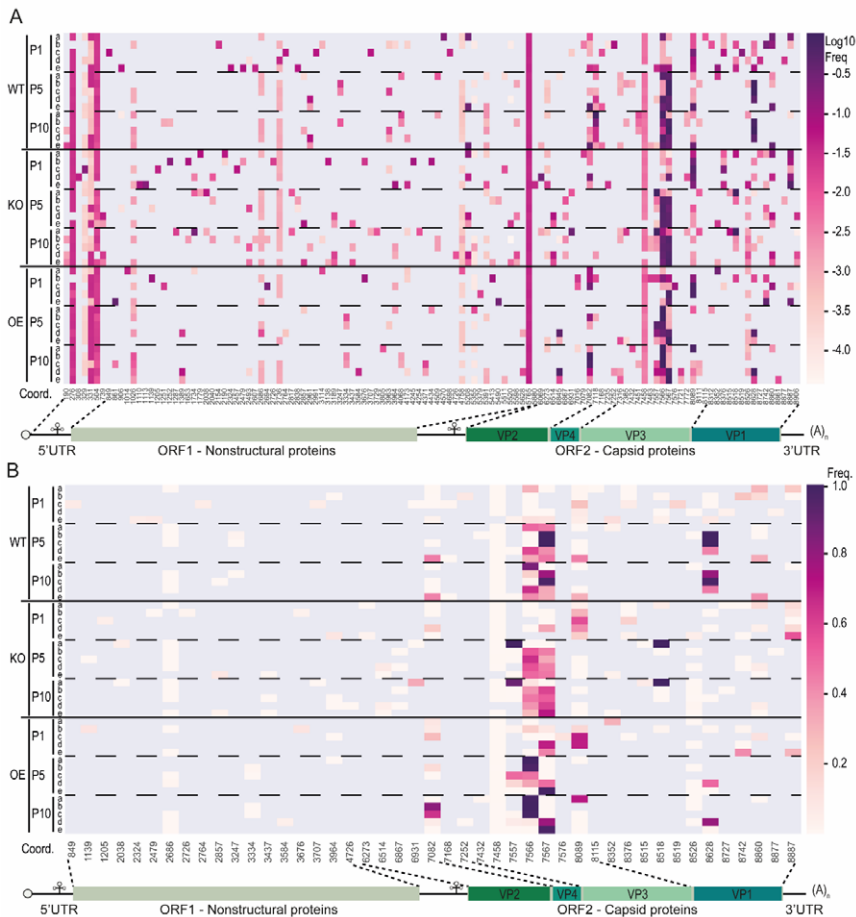


Figure 4. Parallel evolution of virus lineages. Heatmap of SNV frequencies of all SNVs (A) or non-synonymous SNVs (B) for virus populations from WT, *Dicer-2-OE* and *Dicer-2-KO* flies. The x-axis represents genome positions relative to the DCV EB reference strain (NC_001834.1). Only SNVs that occurred in at least one sample with an observed frequency ≥ 0.01 are shown. Deletions were excluded from the analysis in panel (B). A schematic representation of the DCV genome is shown below the heatmaps.

Parallel evolution is driven by positive selection on the capsid protein

Thirteen SNVs became dominant (frequency $> 50\%$) after ten passages, most of which changed the amino acid sequence of the capsid proteins in a host genotype-independent manner (Fig. 3D and Fig. 4B). Of note, two high-frequency SNVs appeared to be host genotype specific: 7118C became dominant only in WT flies (at a frequency of 99% in one lineage and above 20% in two others) and 6845G became dominant in two lineages from *Dcr-2-OE* flies. However, the majority of the dominant SNVs were located at

codon 95 of the capsid VP3 gene (nt 7566–7567), where the tyrosine codon (Y95) in the parental stock was replaced in 14 out of 15 lineages (Fig. 5A). After one passage, the Y95 frequency already declined below 60% in three out of five lineages in *Dcr-2-OE* flies, whereas it remained present at frequencies above 90% in four and three lineages in *Dcr-2-KO* and WT flies, respectively. After five passages, Y95 was undetectable in 12 out of 15 lineages, and at P10, Y95 was only retained in lineage D from *Dcr-2-OE* flies (at a frequency of 6%) and in lineage A from *Dcr-2-KO* flies, which was the only lineage in which Y95 remained dominant. Interestingly, in the latter lineage, two other capsid mutations, VP3 P92T and VP1 N114S, were fixated instead.

Strikingly, from all nine possible SNVs that can be formed from the parental UAU codon coding for Y95, five were found in the evolved virus populations, coding for aspartic acid, asparagine, cysteine, serine and histidine. The emergence of these variants did not correlate with the host genotype in which the virus populations had evolved. The remaining four possible codons were not detected, two of which being stop codons, one coding for a synonymous substitution, and one coding for phenylalanine, the amino acid most similar to tyrosine. Overall, these results suggest that the presence of a benzene ring at residue 95 of VP3 is detrimental to virus infection in flies.

To confirm the selection forces statistically, we computed the population nucleotide diversity per synonymous (π_S) and non-synonymous (π_N) site for each viral gene separately (Supplementary Fig. S5). There were significant main effects of viral gene, site type (synonymous and non-synonymous) and passage number ($P < 0.001$, mixed ANOVA), and significant interactions between viral gene and site type ($P < 0.0001$) and between gene and passage number ($P < 0.05$), but no significant effects of host genotype on the population nucleotide diversity ($P > 0.05$). We found clear indication of positive selection ($\pi_N - \pi_S > 0$) in VP3 across all passages (P1: $P\text{-adj.} = 0.01$; P5: $P\text{-adj.} < 0.0001$; P10: $P\text{-adj.} = 0.004$, Tukey's HSD test with Bonferroni-adjusted P -values for each passage). These results indicate positive selection for amino acid changes in the capsid VP3 protein.

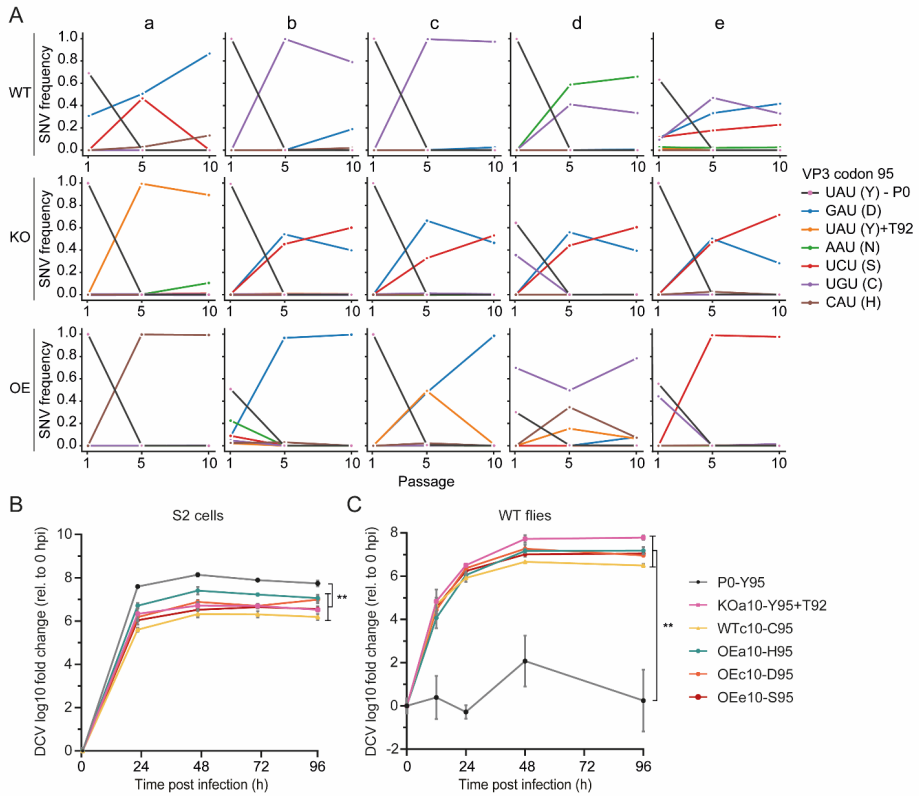


Figure 5. A capsid mutation increases *in vivo* viral fitness. (A) Frequencies of the codon coding for amino acid 95 of VP3 in DCV populations from the indicated host genotypes at passage 1, 5 and 10. UAU represents the parental codon coding for tyrosine (95Y). One variant retained the 95Y residue, but contained a substitution in its vicinity (P92T). **(B)** Growth curves of individual variants in *Drosophila* S2 cells infected at an MOI of 0.1. Variants were named according to the host genotype, lineage and passage number. P0 is the parental virus. Three wells were collected per time point for viral RNA quantification by RT-qPCR. Viral RNA levels are normalized to the 0 hpi time point and presented as geometric mean and SD. **(C)** Growth curves of individual variants in WT flies after inoculation with 100 TCID₅₀ of each variant. Three pools of five flies were collected per time point for viral RNA quantification by RT-qPCR. Viral RNA levels are normalized to the 0 hpi time point and presented as geometric mean and SD. Differences in log₁₀-transformed viral RNA levels between the parental and evolved lineages at 96 hpi were analyzed by *t*-tests (B) and Welch's unequal variances *t*-tests (C). **, $P < 0.01$. Data for lineage OEa10 in (C) did not follow a normal distribution and a Mann-Whitney U test was performed instead ($P = 0.05$).

An exposed amino acid in the capsid is responsible for environment specific fitness gains

We hypothesized that the VP3 mutations were responsible for the observed fitness gains of the passaged virus populations (Fig. 2E). To test this, we selected P10 lineages with one of the fixated mutations (frequency $\geq 90\%$), performed a serial dilution to remove minority variants, and analyzed growth kinetics in the *Drosophila* S2 cell line and in WT flies inoculated with a low dose (multiplicity of infection of 0.1 and 100 TCID₅₀, respectively). All variants efficiently replicated in S2 cells, but the parental stock showed higher replication kinetics than the other variants (unpaired *t*-test, $P < 0.01$, Fig. 5B). In contrast, the parental stock scarcely replicated in adult flies at this low inoculum, whereas the evolved variants replicated efficiently *in vivo* (unpaired Welch's *t*-test, $P < 0.01$, Fig. 5C). These results suggest that Y95 may be beneficial for DCV replication in cell culture, but detrimental for replication and virulence *in vivo*.

To investigate the effects of Y95 on the structural stability of the virion, we modeled the effect of the fixated mutations on the structure of the virus particle. We used the crystal structure of cricket paralysis virus (CrPV) (26), a related dicistrovirus which shares 60% sequence identity with the DCV capsid polyprotein and 79% identity within the region of interest. Given the high identity between these two proteins, it is possible to model the variants in CrPV capsid to predict their probable implications on the DCV capsid. The capsid polyprotein is a large assembly of 240 subunits, composed of copies of four proteins (VP1, VP2, VP3 and VP4) (Fig. 6A). The region of interest is located in a loop close to the protein-protein interface between VP2 and VP3 (Fig. 6A, 6C-6E). While this region is conserved across related dicistroviruses that share more than 50% sequence identity with the DCV capsid, the residues at positions 92 and 95 are relatively poorly conserved (Fig. 6B). Moreover, given their location in the loop, these residues are predicted to be solvent-exposed, suggesting that they may not be involved in maintaining the structural integrity of the capsid. Indeed, we do not find drastic differences in the predicted folding free energy of the capsid structure between the parental Y95 and its mutants (ranging from 0.71 kcal/mol to 3.33 kcal/mol; Supplement Table S7), suggesting that the observed fitness gains cannot be explained by structural effects of the acquired mutations.

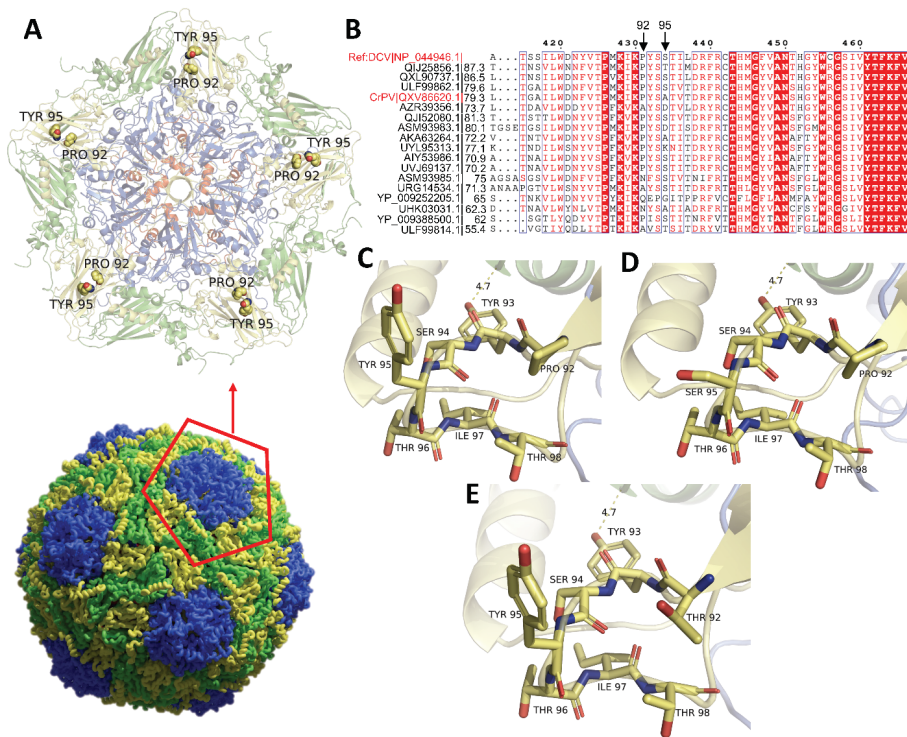


Figure 6. Structural analysis of capsid mutations selected during serial passage. **(A)** The CrPV capsid structure (PDB: 1B35) is shown with VP1 in blue, VP2 in green, VP3 in yellow, and VP4 in red (bottom). View at the fivefold symmetry axis of the capsid in cartoon representation with Tyr95 and Pro92 highlighted as spheres (top). **(B)** Multiple sequence alignment of the capsid of DCV and related viruses in the VP3 region of interest. Accession numbers are followed by their amino acid sequence identities to the DCV reference sequence (NP_044946.1). Completely conserved residues are highlighted in red blocks. **(C)** The parental residues Tyr95 and Pro92 are located in a solvent-exposed loop connecting two beta strands. Amino acids surrounding Tyr95 and Pro92 within a distance of 6 Å are shown as sticks. **(D)** A similar view showing the Tyr95 to Ser95 substitution selected upon serial passage. **(E)** A marginally stabilizing mutant Pro92Thr (-1.79 kcal/mol) is shown. Proximity of the loop to VP2 (green) at a distance of 4.7 Å is depicted in panels (C–E).

Discussion

In this study, we have explored the effect of immune pressure on virus evolution using *Drosophila* as a model host. We passaged DCV in flies differing in Dcr-2 activity, hypothesizing that variations in RNAi immune pressure would affect virus evolution at the phenotypic and genotypic level, for instance by removing selection from the viral 1A protein that interferes with the host RNAi pathway. However, the effect of the RNAi machinery on DCV evolution was moderate. No

specific polymorphisms accumulated in the DCV 1A gene in virus populations from *Dcr-2-KO* flies, nor in the other genotypes. However, we did observe that virus populations from *Dcr-2-KO* flies accumulated more polymorphisms than *Dcr-2* expressing flies, in line with previous observations (27), suggesting that the absence of immune pressure releases selective constraints, reducing purifying selection and allowing virus populations to further explore the mutational landscape.

It will be of interest to use our experimental approach to study the evolution of viruses that differ in their mode of RNAi suppression. Although DCV 1A interacts with several proteins including Dicer-2 (10), it relies on its dsRNA binding activity to suppress RNAi (9). As such, DCV may be under different selection pressure than viruses that suppress RNAi via direct protein-protein interactions, such as CrPV and Nora virus (10–13). It will likewise be interesting to study DCV evolution in flies bearing alleles of the *pastrel* restriction factor, which has strong effects on viral replication and virulence and may encode a protein that directly interacts with viral proteins (28).

Even though we observed some host genotype-specific effects on virus evolution, the strongest evolutionary pattern was shared among virus populations from all host genotypes, suggesting parallel and reproducible evolution. Specifically, the parental tyrosine at position 95 of VP3 was substituted with either one of five different amino acids, resulting in strongly enhanced replication and virulence *in vivo*. The predictability and reproducibility of evolutionary trajectories across independent populations has been a long-standing problem (29). Examples of parallel evolution exist, e.g., a mutation giving rise to antibiotic resistance was observed in multiple isolates of *Escherichia coli* independently (30), and specific drug resistance-associated mutations appeared repeatedly in human immunodeficiency virus-1 (HIV-1) in individuals receiving antiretroviral therapy (31–33). Likewise, the same subset of mutations has arisen independently in HIV-1 in two different cell lines during a long-term experimental evolution study (34). In contrast, in the long-term evolutionary study of Lenski *et al.* (35), the phenotype allowing *E. coli* to use citrate in aerobic conditions occurred in multiple independent lineages, but the underlying molecular mechanism was different (36). In our experiment, VP3 gene mutations occurred in parallel in all virus lineages, concomitant with a fitness gain and higher virulence *in vivo*. These mutations were fixated in 14 out of 15 lineages independently, indicating that loss of Y95 is a highly reproducible, yet unpredictable process, as five different variants have taken over the population.

Another instance of polymorphisms occurring in all lineages were deletions of a single uridine in homopolymeric tracts of at least 6 uridines in the 5' UTR and in the RdRp gene. These uridine deletions also occurred in NGS data of DCV from other studies, and we propose that they are caused by RdRp slippage on homopolymeric uridine tracts. In contrast, thymidine deletions did not occur in NGS data from a DNA virus, herpes simplex virus 2, likely as a result of the proofreading activity of its viral DNA polymerase. In agreement, *E. coli* strains lacking mismatch repair mechanisms show increased numbers of deletions in simple sequences repeats (37).

It is well described that insertions and deletions preferentially occur in homopolymeric tracts in viral genomes (38). Some viruses make use of this feature, such as viruses of the *Potyviridae* family that use transcriptional slippage to insert an adenosine in a specific GAAAAAA sequence in viral transcripts to induce a frameshift and expression of an alternative viral protein (39). Likewise, the measles and Sendai viruses undergo a slippage event that adds an extra guanidine in transcripts coding for the P protein, creating a frameshift transcript that codes for the V protein involved in antagonizing host immunity (40). However, transcriptional slippage also occurs in picornaviruses on homopolymeric stretches of more than six uridines and these sequences are underrepresented in their genome, suggesting that there are also fitness costs associated with poly(uridine) tracts (41). The deletion in the *RdRp* gene in DCV introduces a premature stop codon in the RdRp and likely results in the production of defective viral genomes (DVGs). DVGs have been reported in many virus families and may affect pathogenesis via different mechanisms, like interfering with antiviral immunity or facilitating viral persistence (42). Whether DCV likewise produces DVGs and their involvement in the infection cycle remains to be established.

We found that the parental DCV stock replicated at low levels in flies receiving a low inoculum, but that it efficiently replicated in cell culture. This suggests that the parental virus is unfit because it lacks a basic function that is essential *in vivo*, but not in cell culture. Considering that the virus replicates efficiently in cells, we expect cell entry not to be impaired. Using structural models, we predicted that Y95 substitutions do not affect virus particle stability and that the residue is solvent exposed. We hypothesize that the observed phenotype is due to immune recognition of the parental stock containing Y95 in VP3, and escape thereof by mutants in VP3. This scenario is not unanticipated, as a single nucleotide substitution has been found to mediate immune evasion before. For example, blood clearance of alphaviruses like chikungunya virus is

mediated mainly by resident macrophages in the liver, but a single amino acid substitution in the E2 glycoprotein disrupts this process and increases virus dissemination (43). We speculate that the parental virus is recognized by an immune receptor to mediate virus clearance, thus explaining the differences between *in vitro* versus *in vivo* replication.

To conclude, we have found that elimination of RNAi-based immune pressure during virus evolution led to a higher accumulation of polymorphisms, perhaps due to a reduction in purifying selection. Moreover, we have observed how a single amino acid change in the capsid of an RNA virus can have dramatic environment-specific effects on the growth rate. Our study illustrates how evolution towards higher virulence can be a highly reproducible, yet unpredictable process.

Materials and Methods

Fly strains and husbandry

Flies were maintained at 25 °C in standard fly food. Eggs were treated with bleach and, subsequently, flies were treated with tetracycline for two fly generations, as described before (44). Absence of nora virus, Drosophila C virus, Drosophila X virus and cricket paralysis virus was confirmed by RT-PCR using random hexamers for cDNA synthesis and the following primers for PCR:

- NoraFor, ATGGCGCCAGTTAGTGCAGACCT;
- NoraRev, CCTGTTGTTCCAGTTGGGTTCGA;
- DCVFor, AAAATTTTCGTTTTAGCCCAGAA;
- DCVRev, TTGGTTGTACGTCAAAATCTGAG;
- DXVFor, TCGGAAGAACCAAAAGGATG;
- DXVRev, GTCCTCTCCACGCACTCTTC;
- CrPVFor, ACGAGGAAGCAACTCAAGGA;
- CrPVRev, GAGCCCGCTGAGATGTAAAG.

Absence of *Wolbachia* in the fly stocks was confirmed by PCR using primers wspFor TTTGCAAGTGAAACAGAAGG, wspRev GCTTTGCTGGCAAAATGG. Isogenized *Dicer-2-KO* (*Dcr-2*^{L811fsX}) have been described before (27). Flies overexpressing *Dicer-2* were generated by crossing *UAS-Dicer-2* virgin females (VDRC ID 60008, (45)) with *Tubulin-Gal4/TM3 Sb* male flies, after which flies lacking the *Sb* marker were selected from the F1 offspring. *w*¹¹¹⁸ flies were used as wild-type. Two to five-day-old female flies were used for the serial passage and the RNA extractions.

Virus stock and titration

The parental DCV stock (Charolles strain) was produced in *Drosophila* S2 cells, which were maintained in Schneider's medium supplemented with 10% fetal calf serum and 50 U/ml penicillin and 50 µg/ml streptomycin (Gibco) medium at 27 °C. The stock was bottlenecked through three rounds of serial dilutions, by selecting the highest dilution that still showed cytopathic effects. Subsequently, a T25 flask of confluent S2 cells containing 5 ml of medium was infected with the bottlenecked virus. After 72 hours, the supernatant was collected, centrifuged at 300 *g* for 5 min, aliquoted, and stored at -80 °C. Titers were determined by end-point dilution in 96-well plates, as described before (44), and expressed as median tissue culture infectious dose (TCID₅₀), calculated with the Reed-Muench method.

Experimental evolution

Females of each genotype were intrathoracically inoculated with 10,000 TCID₅₀ of DCV in 50 nl of 10 mM Tris-HCl, pH 7.3 using a Nanoject II microinjector (Drummond Scientific). For each of the five lineages per genotype, forty flies were inoculated. For the first passage, flies were inoculated with the parental stock and from P1 onwards, all 15 lineages were kept independently. After each passage, flies were harvested for preparation of virus stocks. Pools of ten flies were homogenized in 220 µl of PBS, using 1 mm silica beads in a Precellys homogenizer two times for 10 seconds, centrifuged at 16,000 *g* for 10 min to discard fly debris, and the supernatant was transferred to fresh tube, aliquoted, stored at -80°C, and titrated. For RNA extraction and NGS library of the parental stock, 1 ml of RNA-Solv (Omega) was added to 100 µl of virus stock, while for the viral lineages, pools of five flies were homogenized in 1 ml of RNA-Solv (Omega) using 1 mm silica beads in a Precellys homogenizer.

Virus infections and survival assays

To study virulence, 10,000 TCID₅₀ of all lineages from P1, P5, P10 and the parental DCV stock were injected separately into 35 female *w*¹¹¹⁸ flies. Survival was checked daily and analyzed using the Kaplan-Meier estimator. Early replication was assessed after inoculation of two to five-day-old female *w*¹¹¹⁸ flies with 1000 TCID₅₀ of all lineages of P1, P10 or the parental stock. Three pools of five flies per lineage were collected at 12 hpi and processed for RT-qPCR. For the oral infections, 30 female *w*¹¹¹⁸ flies were inoculated with 1000 TCID₅₀ by intrathoracic injection, placed in tubes for three days, after which these flies were discarded and 30 naive female *w*¹¹¹⁸ flies were added to the contaminated vial. Three pools of five flies were collected at 48 hpi and processed for RT-qPCR.

To analyze replication of the different capsid variants, P10 virus lineages containing the variant of interest at a frequency $\geq 90\%$ were bottlenecked by serial dilution. Supernatant from the last well showing cytopathic effect in an end-point dilution was used to grow stocks in a T25 flask, followed by titration and Sanger sequencing to confirm the presence of the mutation. 100 w^{1118} female flies were inoculated with 100 TCID₅₀ of each viral stock, and three pools of five flies per variant were collected at the given time points and processed for RT-qPCR.

RT-qPCR

Total RNA was extracted, reverse transcribed using the TaqMan reverse transcription kit (Thermo Fisher) using oligo (dT) as a primer and quantitative PCR analysis was performed using the GoTaq qPCR SYBR master mix (Promega) on a LightCycler 480 instrument (Roche), according to the manufacturers' recommendations. DCV was quantified using primers DCVFor (TCAAGAAAAGTTGCGTGGGT) and DCVRev (CAGAGCGTCCTTGGAGAGTG), and expression was normalized to expression of *Ribosomal protein 49*, amplified using primers Rp49For (ATGACCATCCGCCAGCATAC) and Rp49Rev (CTGCATGAGCAGGACCTCCA). A standard curve of 10-fold dilutions of plasmid pAc5.1-V5-His-A (Invitrogen) containing the DCV ORF-2 sequence was used to convert Ct values to genome copy numbers.

RNA extraction, cDNA synthesis and NGS library preparation

Total RNA was isolated using RNA-Solv (Omega Bio-tek), and reverse transcribed using Superscript IV (Thermo Fisher) and oligo (dT) as a primer, following the manufacturers' instructions. Viral cDNA was PCR amplified using CloneAmp HiFi (TaKaRa) in four overlapping amplicons of about 2.2 kb, using the following primers:

- For1, TGTACACACGGCTTTTAGGTAGA;
- Rev1, GGAAAAGTGTTGCAAGAGCGA;
- For2, TGTTCTTCGGGAAATGGGGA;
- Rev2, TGCCTGTCCACACGAATAG;
- For3, GTTACGCGTGTCTTTGACG;
- Rev3, AACTTCCGTACCAACGCTCA;
- For4, AAAGTTGCGTGGGTTTGTGG;
- Rev4, CGTGTAAGCAGGGCAGATAGT.

Amplicons were purified with the NucleoSpin kit (MACHEREY-NAGEL), sheared in a Bioruptor Pico (Diagenode) by 12 cycles of 30 sec of sonication and 30 sec of cooling in 1.5 ml Bioruptor tubes (Diagenode), mixed in equimolar

proportions, and used to prepare the library with the Next ultra II DNA library prep kit (E7645, NEB) using 1 µg of DNA per sample as input. Samples were multiplexed with barcodes (E7335, E7500, E7710 and E7730, NEB), and the libraries were sequenced on an Illumina HiSeq 4000 sequencer as paired-end 100 base reads by the GenomEast platform, a member of the 'France Genomique' consortium. Image analysis and base calling were performed using RTA version 2.7.7 and bcl2fastq version 2.20.0.422.

NGS data processing

NGS data were analyzed using a custom computational workflow (Supplementary Fig. S6). For processing the raw read data, the bioinformatics pipeline V-pipe (46) was integrated into the workflow. Burrows-Wheeler Alignment Tool (BWA-MEM) (47) was used to align reads from the parental stock to the DCV EB reference strain that contains complete UTR sequences (GenBank accession number: NC_001834.1). Bcftools was used to generate a consensus sequence from the parental sequence, which was then used to align the raw reads from the evolutionary samples from passage 1, 5, and 10 using BWA-MEM. Mutations with respect to the parental stock consensus sequence were called using ShoRAH (Short Read Assembly into Haplotypes) (48) with default parameters (posterior threshold of 0.9, alpha 0.1 and shift 3). The posterior threshold was chosen conservative to ensure that only mutations of high confidence are included in the subsequent analyses. Moreover, the Strand Bias Filter (49) was applied to filter out further variation of low confidence. Phred sequencing quality scores for mutation calls were accessed using pysamstats (<https://github.com/alimanfoo/pysamstats>). Further, to evaluate the differences of our parental stock to its ancestor, the Charolles strain, the parental stock sample was processed with respect to the Charolles reference sequence (GenBank Accession Number: MK645242.1). An adapted version of vcf_annotator (<https://github.com/rpetit3/vcf-annotator>) was used to annotate non-synonymous mutations. Population nucleotide diversity and number of polymorphisms were computed considering only mutations with a minimum frequency of 0.0001. SNPGenie's within-pool analysis was used in the default settings and minimum frequency 0.0001 (50) to compute the population nucleotide diversity per synonymous (piS) and non-synonymous (piN) site.

Permutation test for genotype specific adaptation

A permutation test as described before (51) was used to test for host genotype-specific adaptation. Mutations that reached the frequency threshold of 0.0001 in at least one evolutionary line were considered for this analysis. For

each mutation, the number of lineages containing the mutation were counted. Then, each mutation occurrence was randomly reallocated to a lineage by using the number of mutations that occurred in each lineage as weights. In this manner, 1000 samples were drawn. The observed number of mutations that were unique or shared among genotypes was then compared to the number expected from the permutation samples, and a two-sided test was used to assess the difference. For each category, P -values were computed as $\frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \geq O)$ where n is the number of samples drawn, X_i is the i -th random realization and O is the observed number of mutations in the respective category.

Randomization test

A randomization test was performed to test whether the observed number of two consecutive phenylalanine (FF) residues encoded as 'TTTTTT' was different from the expected number based on random mixing of the TTT and TTC codons. For each viral genome, the total number of FF occurrences, the number of FF encoded by TTTTTT (O), and the occurrences of the TTC and TTT codon were counted. For each viral genome, samples were drawn, by randomly pairing the TTC and TTT codons based on their codon frequencies. To test whether the number TTTTTT encodings from the randomization exceeds the observed number, a two-sided test was performed. P -values were computed as $\frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \geq O)$ where n denotes the number of TTTTTT realizations in the i -th randomization.

Quality analysis of deletion calls

To evaluate the quality of reads supporting the two uridine deletions at positions 276–284 and 5765–5774, an additional quality control was performed on the alignment files of all samples. The mean position-wise base calling quality scores were assessed in the homopolymeric regions of the two deletions and the values were compared between reads with deletions to the values of reads without the deletions. In addition, the proportion of reads where the deletion track was located within a range of five nucleotides from the read edges was quantified.

Capsid structural modeling

In the absence of structural information of the DCV capsid protein, the crystal structure of the capsid of CrPV (PDB: 1B35) was used to model the impact of mutations on capsid structure. Mutations of interest, located at sites 92 and 95 in VP3, were modeled for all VP3 subunits in the capsid assembly. Structural models were constructed for each mutation using the functions RepairPDB and BuildModel from the FoldX suite (v.5.0) with five iterations of sidechain

rotamer adjustments. For each mutant constructed, the average differences in capsid structural stabilities, i.e., free energy of protein folding (kcal/mol), between the parental amino acid and the mutants were calculated.

To assess sequence conservation of the DCV capsid polyprotein (NP_044946.1), a multiple sequence alignment was generated of closely related viruses (> 50% sequence identity), identified using a BLAST search: Warroolaba Creek virus (QIJ25856.1), Rondani's wasp virus 1 (QXL90737.1), Cricket paralysis virus (QXV86620.1), Hypsignathus monstrosus dicistrovirus (AZR39356.1), Millport beadlet anemone dicistro-like virus 1 (ASM93983.1), Maanshan Dicis tick virus 1 (UYL95313.1), Nilaparvata lugens C virus (AIY53986.1), Arma chinensis virus 1 (UVJ69137.1), Caledonia beadlet anemone dicistro-like virus 2 (ASM93985.1), Anopheles C virus (YP_009252205.1), Guiyang dicistrovirus 2 (UHK03031.1), Apis dicistrovirus (YP_009388500.1), as well as other, uncharacterized members of the *Cripavirus* genus (ULF99862.1) and the *Dicistroviridae* family (QJI52080.1, AKA63264.1, URG14534.1, ULF99814.1).

Statistical tests

Two-sided *t*-tests on log₁₀-transformed DCV titers for each passage was used for pairwise comparisons in viral titers in the different genotypes (Python package SciPy (52), `ttest_ind` function). Data were log₁₀-transformed to ensure that the residuals were Gaussian distributed (Python package SciPy, `normaltest` $P > 0.05$ for each group), and equal variance was confirmed using a Levene's test. Mean survival times of WT flies upon infection with virus populations from the different passages were compared across passages using a paired two-sided *t*-test. The Shapiro-Wilk test was applied to validate that the data were Gaussian distributed (Python package Pingouin (53), `normality` function). Mortality of WT flies infected with P10 viruses and the parental stock viruses was compared using a log-rank test and Cox regression (IBM SPSS Statistics, version 29.0.0.0). Viral RNA levels between the parental stock and the evolved viral populations at passage 10 were compared using a nested ANOVA, after which a Tukey's HSD test with *P*-value adjustment was applied for multiple comparison between genotypes. The data were first log₁₀-transformed, and equal variance was confirmed using a Levene's test ($P > 0.05$). The factor 'technical replicate' is nested within each biological lineage, and the parental stock was classified as a separate genotype category. For each genotype, the difference in the number of polymorphic sites across passages was analyzed using paired two-sided *t*-tests. To compare differences in the number of polymorphic sites among the host genotypes at passage 10, two-sided *t*-tests were used (Python package SciPy, `ttest_rel` function). The data

were Gaussian distributed (Python package, SciPy normaltest $P > 0.05$ for each group), and equal variance was confirmed using a Levene's test. To analyze viral nucleotide diversity within passages and between fly genotypes, a mixed ANOVA was used with passages 1, 5 and 10 as within-subject factor and fly genotype as between-subject factor (Python package Pingouin, `mixed_anova` function). The residuals were Gaussian distributed, and equal variance was confirmed using a Levene's test (Python package Pingouin, `homoscedasticity` function). To compare the mean nucleotide diversity between coding and non-coding regions, a mixed ANOVA was used with passages as the within-subject factor and coding or non-coding region as the between-subject factor (Python package Pingouin, `mixed_anova` function). A mixed ANOVA was performed to analyze the effect of genotype, passage, protein and site type (non-synonymous and synonymous) on the population nucleotide diversity. Passage was used as within-subject factor in the model. The data were cube-root-transformed to ensure approximate normal distribution and equal variance among groups, which was confirmed using a Levene's test ($P > 0.05$). Subsequently, per passage, Tukey's HSD test with Bonferroni adjusted P -value was used for multiple comparison of viral gene and site type. Viral RNA loads between different VP3 variants at 96 hpi in S2 cells were log-10 transformed and upon validation of the normal distribution using the Saphiro-Wilk test (function `shapiro.test`, R Stats package) and equal variance by the F test (`var.test` function, R Stats package) and analyzed using two-sided t -tests. Welch's unequal variances t -tests were used to compare log10-transformed viral RNA loads of VP3 variants at 96 hpi in WT flies. Unless noted otherwise, statistical analyses were performed using Rstudio 2022.02.1+461.

Data availability

All raw NGS data have been deposited in NCBI BioProject under the accession number PRJNA993483.

Code availability

The data processing pipeline can be found at <https://github.com/cbg-ethz/DCV-Evolution-Data-Processing>.

Acknowledgements

We thank members of the laboratory for discussions and Carla Saleh (Pasteur Institute) for providing *Dicer-2* knockout flies. This work was financially supported by the European Union's Horizon 2020 research and innovation program, under the Marie Skłodowska-Curie Actions Innovative Training Networks grant agreement no. 955974 (VIROINF).

References

1. Koonin EV, Dolja VV. A virocentric perspective on the evolution of life. *Curr Opin Virol.* 2013;3(5):546-57, doi: 10.1016/j.coviro.2013.06.008.
2. Iranzo J, Puigbo P, Lobkovsky AE, Wolf YI, Koonin EV. Inevitability of genetic parasites. *Genome Biol Evol.* 2016;8(9):2856-69, doi: 10.1093/gbe/evw193.
3. Krupovic M, Dolja VV, Koonin EV. Origin of viruses: primordial replicators recruiting capsids from hosts. *Nat Rev Microbiol.* 2019;17(7):449-58, doi: 10.1038/s41579-019-0205-6.
4. Finlay BB, McFadden G. Anti-immunology: evasion of the host immune system by bacterial and viral pathogens. *Cell.* 2006;124(4):767-82, doi: 10.1016/j.cell.2006.01.034.
5. Garcia-Sastre A, Biron CA. Type 1 interferons and the virus-host relationship: a lesson in detente. *Science.* 2006;312(5775):879-82, doi: 10.1126/science.1125676.
6. Gammon DB, Mello CC. RNA interference-mediated antiviral defense in insects. *Curr Opin Insect Sci.* 2015;8:111-20, doi: 10.1016/j.cois.2015.01.006.
7. Guo Z, Li Y, Ding SW. Small RNA-based antimicrobial immunity. *Nat Rev Immunol.* 2019;19(1):31-44, doi: 10.1038/s41577-018-0071-x.
8. Bronkhorst AW, van Rij RP. The long and short of antiviral defense: small RNA-based immunity in insects. *Curr Opin Virol.* 2014;7:19-28, doi: 10.1016/j.coviro.2014.03.010.
9. van Rij RP, Saleh MC, Berry B, Foo C, Houk A, Antoniewski C, et al. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes Dev.* 2006;20(21):2985-95, doi: 10.1101/gad.1482006.
10. Nayak A, Kim DY, Trnka MJ, Kerr CH, Lidsky PV, Stanley DJ, et al. A viral protein restricts *Drosophila* RNAi immunity by regulating argonaute activity and Saability. *Cell Host Microbe.* 2018;24(4):542-57 e9, doi: 10.1016/j.chom.2018.09.006.
11. Nayak A, Berry B, Tassetto M, Kunitomi M, Acevedo A, Deng C, et al. Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in *Drosophila*. *Nat Struct Mol Biol.* 2010;17(5):547-54, doi: 10.1038/nsmb.1810.
12. van Mierlo JT, Overheul GJ, Obadia B, van Cleef KW, Webster CL, Saleh MC, et al. Novel *Drosophila* viruses encode host-specific suppressors of RNAi. *PLoS Pathog.* 2014;10(7):e1004256, doi: 10.1371/journal.ppat.1004256.
13. van Mierlo JT, Bronkhorst AW, Overheul GJ, Sadanandan SA, Ekstrom JO, Heestermans M, et al. Convergent evolution of argonaute-2 slicer antagonism in two distinct insect RNA viruses. *PLoS Pathog.* 2012;8(8):e1002872, doi: 10.1371/journal.ppat.1002872.
14. Llaure AS, Andino R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* 2010;6(7):e1001005, doi: 10.1371/journal.ppat.1001005.
15. Duffy S. Why are RNA virus mutation rates so damn high? *PLoS Biol.* 2018;16(8):e3000003, doi: 10.1371/journal.pbio.3000003.
16. Llaure AS, Frydman J, Andino R. The role of mutational robustness in RNA virus evolution. *Nat Rev Microbiol.* 2013;11(5):327-36, doi: 10.1038/nrmicro3003.
17. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet.* 2008;9(4):267-76, doi: 10.1038/nrg2323.
18. Moya A, Holmes EC, Gonzalez-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. *Nat Rev Microbiol.* 2004;2(4):279-88, doi: 10.1038/nrmicro863.
19. Hughes AL. Adaptive evolution of genes and genomes. New York, NY, USA: Oxford University Press; 1999.

20. Acevedo A, Brodsky L, Andino R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature*. 2014;505(7485):686–90, doi: 10.1038/nature12861.
21. Navarro R, Ambros S, Butkovic A, Carrasco JL, Gonzalez R, Martinez F, et al. Defects in plant immunity modulate the rates and patterns of RNA virus evolution. *Virus Evol*. 2022;8(2):veac059, doi: 10.1093/ve/veac059.
22. Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, et al. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell*. 2004;117(1):69–81, doi: 10.1016/s0092-8674(04)00261-2.
23. Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler JL. Essential function in vivo for Dicer-2 in host defense against RNA viruses in *Drosophila*. *Nat Immunol*. 2006;7(6):590–7, doi: 10.1038/ni1335.
24. Spellberg MJ, Marr MT, 2nd. FOXO regulates RNA interference in *Drosophila* and protects from RNA virus infection. *Proc Natl Acad Sci U S A*. 2015;112(47):14587–92, doi: 10.1073/pnas.1517124112.
25. Bons E, Leemann C, Metzner KJ, Regoes RR. Long-term experimental evolution of HIV-1 reveals effects of environment and mutational history. *PLoS Biol*. 2020;18(12):e3001010, doi: 10.1371/journal.pbio.3001010.
26. Tate J, Liljas L, Scotti P, Christian P, Lin T, Johnson JE. The crystal structure of cricket paralysis virus: the first view of a new virus family. *Nat Struct Biol*. 1999;6(8):765–74, doi: 10.1038/11543.
27. Mongelli V, Lequime S, Kousathanas A, Gausson V, Blanc H, Nigg J, et al. Innate immune pathways act synergistically to constrain RNA virus evolution in *Drosophila melanogaster*. *Nat Ecol Evol*. 2022;6(5):565–78, doi: 10.1038/s41559-022-01697-z.
28. Magwire MM, Fabian DK, Schweyen H, Cao C, Longdon B, Bayer F, et al. Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS Genet*. 2012;8(11):e1003057, doi: 10.1371/journal.pgen.1003057.
29. Gould SJ. *Wonderful life: the burgess shale and the nature of history*. New York, NY, USA: W. W. Norton & Company; 1989.
30. Coolen JPM, den Drijver EPM, Verweij JJ, Schildkraut JA, Neveling K, Melchers WJG, et al. Genome-wide analysis in *Escherichia coli* unravels a high level of genetic homoplasmy associated with cefotaxime resistance. *Microb Genom*. 2021;7(4):000556, doi: 10.1099/mgen.0.000556.
31. Flynn WF, Chang MW, Tan Z, Oliveira G, Yuan J, Okulicz JF, et al. Deep sequencing of protease inhibitor resistant HIV patient isolates reveals patterns of correlated mutations in Gag and protease. *PLoS Comput Biol*. 2015;11(4):e1004249, doi: 10.1371/journal.pcbi.1004249.
32. Crandall KA, Kelsey CR, Imamichi H, Lane HC, Salzman NP. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate ratio to detect selection. *Mol Biol Evol*. 1999;16(3):372–82, doi: 10.1093/oxfordjournals.molbev.a026118.
33. Beerenwinkel N, Daumer M, Sing T, Rahnenfuhrer J, Lengauer T, Selbig J, et al. Estimating HIV evolutionary pathways and the genetic barrier to drug resistance. *J Infect Dis*. 2005;191(11):1953–60, doi: 10.1086/430005.
34. Bertels F, Leemann C, Metzner KJ, Regoes RR. Parallel Evolution of HIV-1 in a Long-Term Experiment. *Mol Biol Evol*. 2019;36(11):2400–14, doi: 10.1093/molbev/msz155.

35. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*. 1991;138(6):1315-41, doi: 10.1086/285289.
36. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2008;105(23):7899-906, doi: 10.1073/pnas.0803151105.
37. Kumar P, Nagarajaram HA. A study on mutational dynamics of simple sequence repeats in relation to mismatch repair system in prokaryotic genomes. *J Mol Evol*. 2012;74(3-4):127-39, doi: 10.1007/s00239-012-9491-6.
38. Domingo E. *Molecular Basis of Genetic Variation of Viruses: Error-Prone Replication*. Cambridge, MA, USA: Academic Press; 2016, doi: 10.1016/B978-0-12-800837-9.00002-2.
39. Olsper A, Chung BY, Atkins JF, Carr JP, Firth AE. Transcriptional slippage in the positive-sense RNA virus family Potyviridae. *EMBO Rep*. 2015;16(8):995-1004, doi: 10.15252/embr.201540509.
40. Atkins JF, Loughran G, Bhatt PR, Firth AE, Baranov PV. Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res*. 2016;44(15):7007-78, doi: 10.1093/nar/gkw530.
41. Stewart H, Olsper A, Butt BG, Firth AE. Propensity of a picornavirus polymerase to slip on potyvirus-derived transcriptional slippage sites. *J Gen Virol*. 2019;100(2):199-205, doi: 10.1099/jgv.0.001189.
42. Vignuzzi M, Lopez CB. Defective viral genomes are key drivers of the virus-host interaction. *Nat Microbiol*. 2019;4(7):1075-87, doi: 10.1038/s41564-019-0465-y.
43. Carpentier KS, Davenport BJ, Haist KC, McCarthy MK, May NA, Robison A, et al. Discrete viral E2 lysine residues and scavenger receptor MARCO are required for clearance of circulating alphaviruses. *Elife*. 2019;8:e49163, doi: 10.7554/eLife.49163.
44. Merkle SH, van Rij RP. Analysis of resistance and tolerance to virus infection in *Drosophila*. *Nat Protoc*. 2015;10(7):1084-97, doi: 10.1038/nprot.2015.071.
45. Dietzl G, Chen D, Schnorrrer F, Su KC, Barinova Y, Fellner M, et al. A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*. 2007;448(7150):151-6, doi: 10.1038/nature05954.
46. Posada-Céspedes S, Seifert D, Topolsky I, Jablonski KP, Metzner KJ, Beerenwinkel N. V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data. *Bioinformatics*. 2021;37(12):1673-80, doi: 10.1093/bioinformatics/btab015.
47. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60, doi: 10.1093/bioinformatics/btp324.
48. Zagordi O, Bhattacharya A, Eriksson N, Beerenwinkel N. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinformatics*. 2011;12:119, doi: 10.1186/1471-2105-12-119.
49. McElroy K, Zagordi O, Bull R, Luciani F, Beerenwinkel N. Accurate single nucleotide variant detection in viral populations by combining probabilistic clustering with a statistical test of strand bias. *BMC Genomics*. 2013;14:501, doi: 10.1186/1471-2164-14-501.
50. Nelson CW, Moncla LH, Hughes AL. SNPGenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. *Bioinformatics*. 2015;31(22):3709-11, doi: 10.1093/bioinformatics/btv449.
51. Bons E, Bertels F, Regoes RR. Estimating the mutational fitness effects distribution during early HIV infection. *Virus Evol*. 2018;4(2):vey029, doi: 10.1093/ve/vey029.

52. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-72, doi: 10.1038/s41592-019-0686-2.
53. Vallat R. Pingouin: statistics in Python. *J Open Source Softw*. 2018;3(31):1026, doi: 10.21105/joss.01026

Supplemental text

Uridine deletions in homopolymeric tracts

Single nucleotide deletions were observed at two locations in the DCV genome at low, yet fairly constant levels during passaging. Since both deletions occurred in homopolymeric uridine tracts, these may be due to error-prone RNA replication by the viral RdRp or to an artifact of library preparation and sequencing. To assess this, we first called variants in other, published DCV NGS data sets (1–3). Uridine deletions in those data sets occurred at the same positions as in our data, and at similar frequencies in the RdRp, but at variable frequencies in the 5'UTR (ranging from 3.6 to 87.8%) (Supplementary Table 4), suggesting that the deletion in the 5'UTR does not negatively impact virus fitness. We next called variants in Illumina based NGS data in a published dataset of a DNA virus, herpes simplex virus 2 (HSV2) (4). The HSV2 genome contains 37 tracts with more than six consecutive thymidines, but no deletions could be detected in those regions (Supplementary Table 6), suggesting that the occurrence of uridine deletions in our DCV data are not due to a sequencing artifact.

If uridine deletions are indeed due to error-prone RNA replication, we would expect negative selection against the occurrence of poly(U) tracts in viral genomes. We examined the coding regions of different RNA viruses for two consecutive phenylalanine (F) residues, which are encoded by either UUC or UUU and may cause the appearance of a hexameric uridine tract. With a randomization test, we then computed the expected number of FF to be encoded by UUUUUU, taking codon frequencies into account. For eight out of nine RNA viruses from different taxa analyzed (*Picornavirales*, *Mononegavirales*, and *Nidovirales* order), the observed number of UUUUUU was lower than the expected number, in three of these cases the underrepresentation reached statistical significance ($P < 0.05$; Supplementary Table 5). In contrast, for DNA viruses the expected number of poly-thymidine tracts matched the observed number for four out of six analyzed viruses (from the order *Herpesvirales*, *Lefavirales* and *Pimascovirales*), while the other two show a mild underrepresentation. Together, we suggest that the viral RdRp is prone to slippage on homopolymeric uridine and that there is purifying selection against the occurrence of these tracts in RNA viruses.

Supplemental tables

Table S1. Hazard ratios (RH) of the evolved virus population compared to the parental stock.

	WT				KO				OE			
	95.0% CI				95.0% CI				95.0% CI			
	RH	Lower	Upper	P value	RH	Lower	Upper	P value	RH	Lower	Upper	P value
P1	2,12	1,63	2,75	1,76E-08	1,15	0,89	1,49	2,79E-01	1,87	1,44	2,42	2,56E-06
P5	1,63	1,23	2,16	5,80E-04	2,37	1,77	3,19	8,53E-09	1,59	1,19	2,12	1,55E-03
P10	4,37	3,21	5,94	7,45E-21	3,13	2,32	4,22	7,73E-14	3,38	2,51	4,55	8,30E-16

Table S2. Table of all SNVs with minimum frequency 0.0001 per evolutionary lineage. Available from: [vead074_Supp.](#)

Table S3. Detailed quality control for deletion calls in uridine tracts at position 276-285 and position 5765-5774. Available from: [vead074_Supp.](#)

Table S4. Frequencies of uridine deletions at position 276-285 and position 5765-5774 in DCV sequencing data from the literature. Data were analyzed using the same data processing pipeline as described in ‘Methods’.

Reference	Sample Accession Number	Position 276-285		Position 5765-5774	
		Coverage	Frequency	Coverage	Frequency
(1)	SRR17044453	18182	15.34%	10108	2.15%
(1)	SRR6156525	1304	0.00%	2469	3.81%
(2)	ERR3510613	805	15.78%	1126	2.13%
(2)	ERR3510616	589	3.57%	1310	0.00%
(3)	SRR17044471	29353	41.74%	16762	2.2%
(3)	SRR17044453	14470	16.48%	7933	2.28%
(3)	SRR17044623	19169	47.43%	6067	2.46%
(3)	SRR17044490	24236	87.81%	10541	3.19%

Table S5. Underrepresentation of FF amino acid encoded by UUUUUU in viral genomes.

Virus	Reference genome	Virus type	Number of FF	Observed number of UUUUUU encoding FF	Expected number of UUUUUU encoding FF, based on randomization			P-values
					Mean	25%-quantile	75%-quantile	
Drosophila C virus	NC_001834.1	RNA	4	0	2.328	1	4	0.046
Cricket paralysis virus	NC_003924.1	RNA	6	1	3.051	1	5	0.182
Nora virus	NC_007919.3	RNA	3	1	1.418	0	3	0.904
Sigma virus	JX403934.1	RNA	4	0	0.687	0	2	0.932
Acute bee paralysis virus	NC_002548.1	RNA	5	0	2.542	1	5	0.046
Deformed wing virus	NC_004830.2	RNA	4	0	2.187	0	4	0.07
SARS-CoV-2	NC_045512.2	RNA	18	1	8.817	5	13	0.002
Herpes simplex virus 2	NC_001798.2	DNA	56	11	11.344	6	17	0.924
Human cytomegalovirus	OK000912.1	DNA	116	28	28.056	20	37	0.918
Kallithea virus	NC_033829.1	DNA	43	24	21.926	15	28	0.422
Invertebrate iridescent virus 6	NC_003038.1	DNA	269	158	196.661	183	210	<0.001
Alphabaculovirus	NC_024625.1	DNA	86	18	30.485	22	39	<0.001

Table S6. Frequencies of deletions in homopolymeric T regions in the Herpes simplex virus serotype 2 genome in sequencing data (accession number: ERR3278849) (2). Available from: vead074_Supp.

Table S7. List of folding free energy differences in the capsid structure calculated between the wildtype Y95 and its mutants D95, S95, N95, C95 and H95, and for P92T.

Variant	Avg. energy difference between mutant and wildtype*	Total energy of wildtype	Avg. energy of mutant
Tyr95Asp	0.71 kcal/mol	1768.31 kcal/mol	1769.02 kcal/mol
Tyr95Ser	2.1 kcal/mol	1768.31 kcal/mol	1770.41 kcal/mol
Tyr95Asn	2.64 kcal/mol	1768.31 kcal/mol	1770.95 kcal/mol
Tyr95Cys	3.07 kcal/mol	1768.31 kcal/mol	1771.38 kcal/mol
Tyr95His	3.33 kcal/mol	1768.31 kcal/mol	1771.64 kcal/mol
Pro92Thr	-1.79 kcal/mol	1768.31 kcal/mol	1766.52 kcal/mol

*ddG= dG(mut) - dG(wildtype)

Supplemental figures

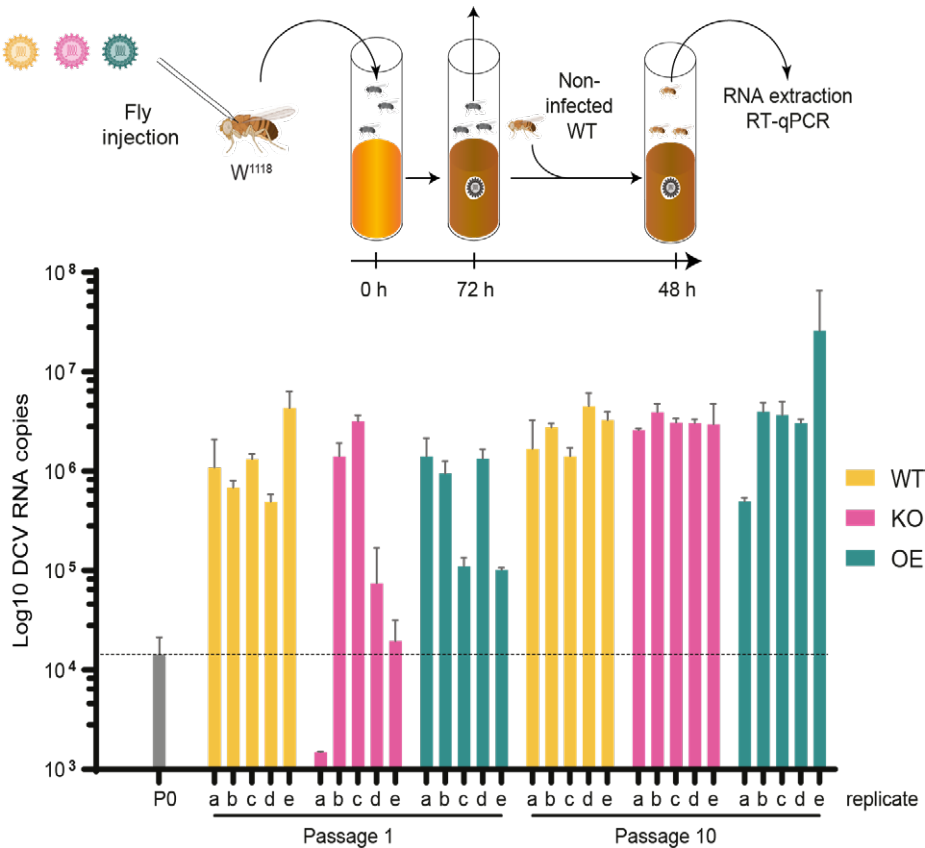


Figure S1. Increased replication of evolved virus populations after oral inoculation. 1,000 TCID₅₀ of each virus stock was injected into WT flies and incubated in separate vials for 3 days. Those flies were discarded and uninfected flies were added to the vials and allowed to feed for 2 days. Viral RNA copy numbers were quantified by RT-qPCR from three pools of five flies. The dashed line shows the RNA copies of the parental stock (P0). Results are shown as means and SD of the three replicates of five flies each.

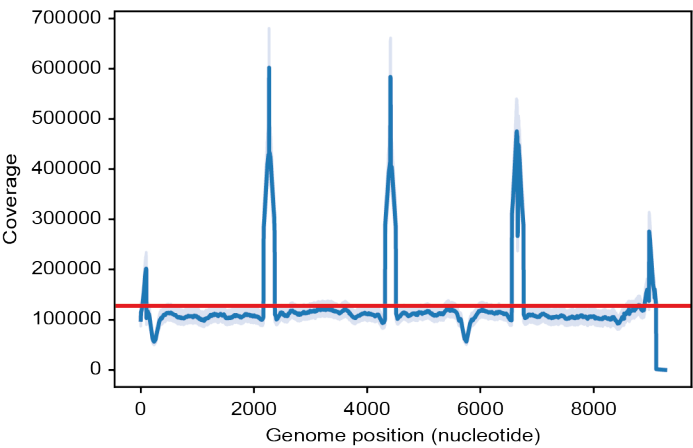


Figure S2. Mean sequence coverage across the genome over all experimental lineages and parental stock. The red line indicates the mean coverage (127,542 reads per position), blue shading indicates the 95% confidence intervals based on the bootstrap distribution. Peaks in sequence coverage correspond to regions where amplicons overlap.

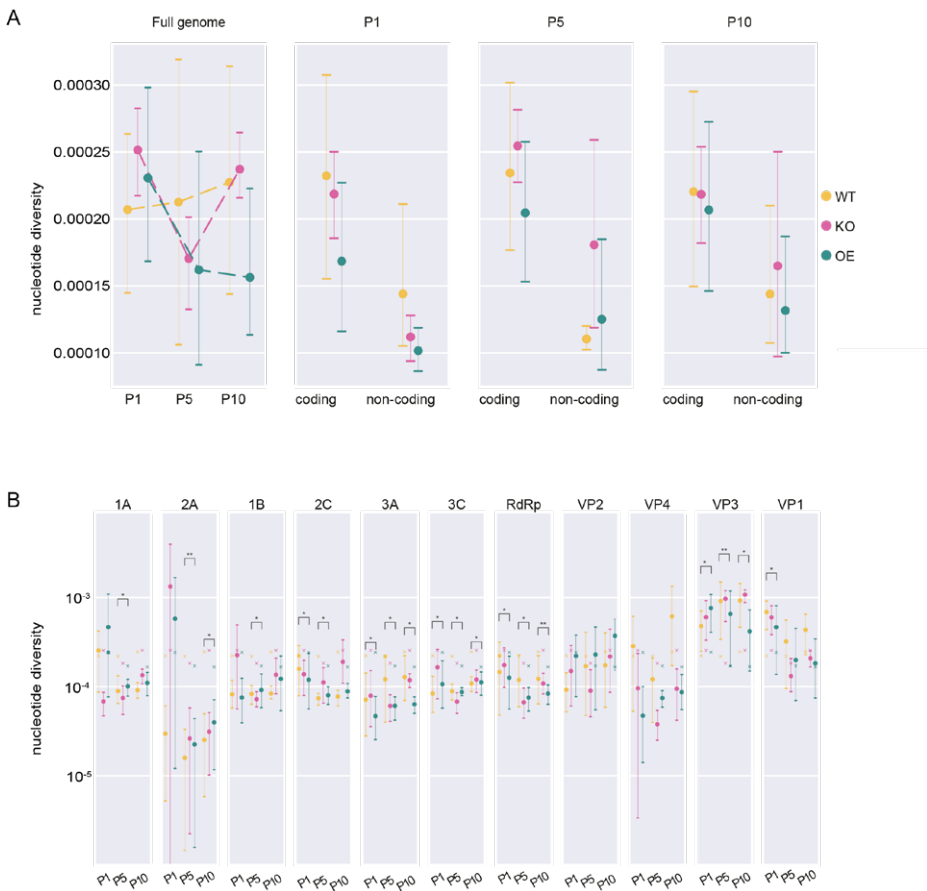


Figure S3. Nucleotide diversities of virus populations during serial passage. **(A)** Mean population nucleotide diversity across the full genome (left panel) and in coding and non-coding regions of the genome (right panels) of virus populations evolved in the three host genotypes. Data are shown as means and standard errors across the five lineages for each genotype. **(B)** Mean nucleotide diversity per viral gene marked with dots, the error bars indicating the standard error across the five lineages. Crosses mark the nucleotide diversity across the coding region for each genotype and passage as a baseline. Statistical significance was tested with paired t-tests comparing baseline to the values of each gene. P-values are indicated with asterisks: *, $P < 0.05$; **, $P < 0.01$.

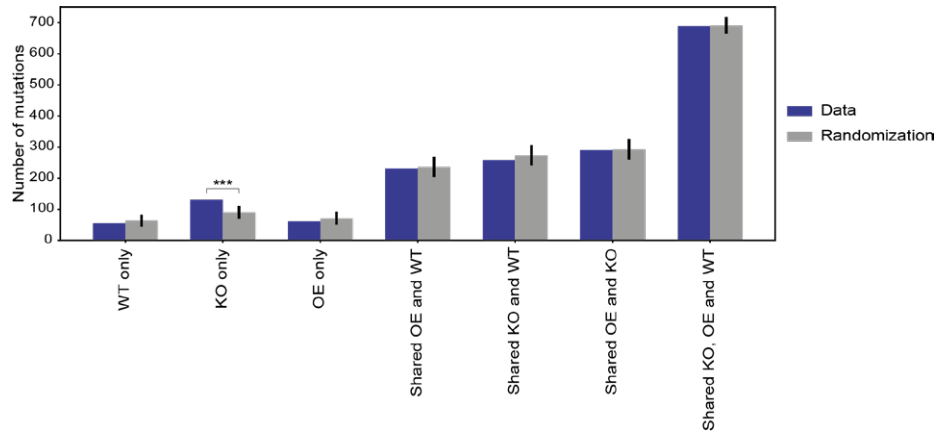


Figure S4. Host genotype specific adaptation in Dcr-2-KO flies. A permutation test with 1000 randomizations was used to analyze adaptation to specific host genotypes. Label 'only' refers to mutations that reach the frequency threshold 0.01% in at least two lineages in the same host genotype. Label 'shared' refers to mutations that reach the frequency threshold in more than one genotype. P-values are indicated with asterisks: ***, $P < 0.001$. Error bars mark the 95% highest posterior density (HPD) intervals around the mean across the randomizations.

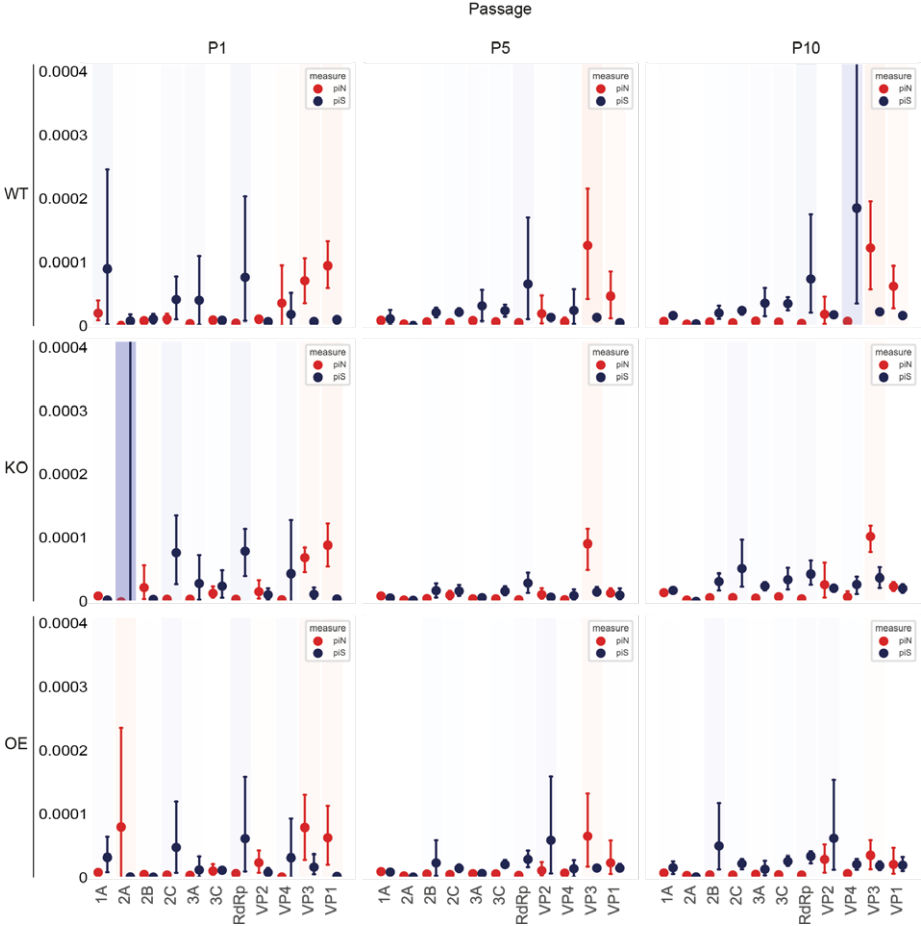


Figure S5. Positive selection on the VP3 capsid protein. Comparison of population nucleotide diversity per synonymous (piS) and non-synonymous (piN) site for each viral gene. Red and blue shading indicate $\pi_N - \pi_S > 0$ and $\pi_N - \pi_S < 0$, respectively. Error bars indicate 95% confidence intervals based on the bootstrap distribution.

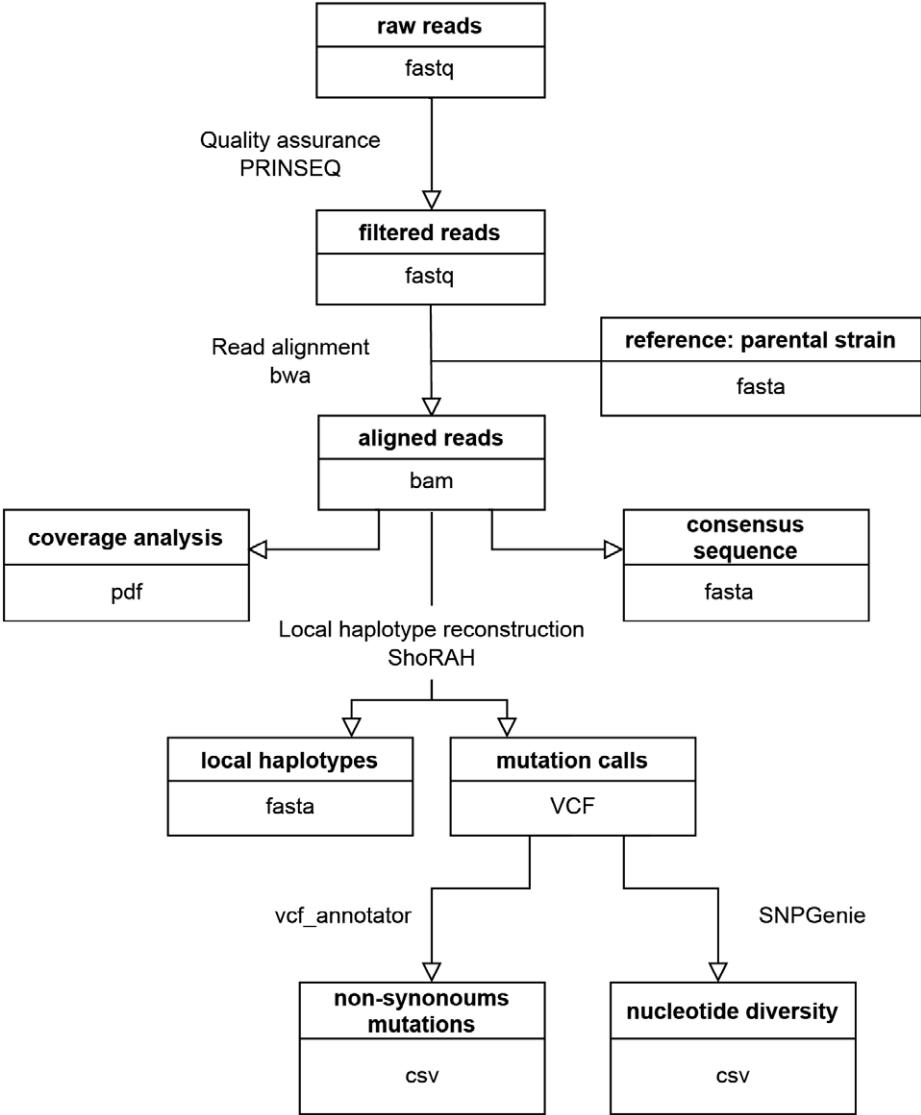
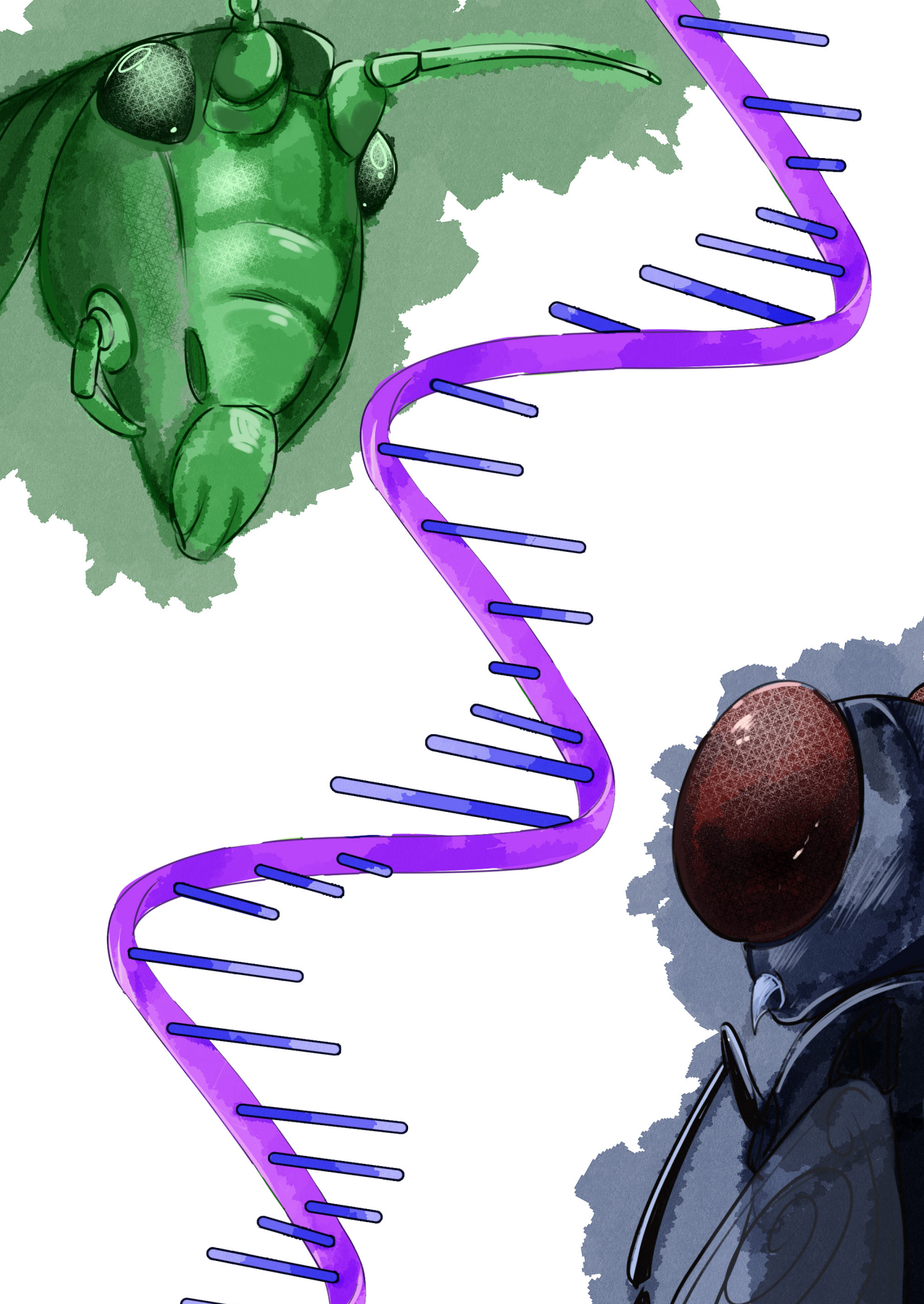


Figure S6. Data processing pipeline overview: Quality assurance, Multiple sequence alignment, Coverage analysis, Consensus sequence calling, Mutation and local haplotype calling, Annotation of mutations calls, Computation of nucleotide diversity.

Supplementary references

1. Longdon B, Day JP, Alves JM, Smith SCL, Houslay TM, McGonigle JE, et al. Host shifts result in parallel genetic changes when viruses evolve in closely related species. *PLoS Pathog.* 2018;12;14(4):e1006951.
2. Martinez J, Bruner-Montero G, Arunkumar R, Smith SCL, Day JP, Longdon B, et al. Virus evolution in *Wolbachia*-infected *Drosophila*. *Proc Biol Sci.* 2019;286(1914):20192117.
3. Mongelli V, Lequime S, Kousathanas A, Gausson V, Blanc H, Nigg J, et al. Innate immune pathways act synergistically to constrain RNA virus evolution in *Drosophila melanogaster*. *Nat Ecol Evol.* 2022;(5):565-578.
4. López-Muñoz AD, Rastrojo A, Kropp KA, Viejo-Borbolla A, Alcamí A. Combination of long- and short-read sequencing fully resolves complex repeats of herpes simplex virus 2 strain MS complete genome. *Microb Genom.* 2021;7(6):000586.



Chapter 3

A recurrent adaptive mutation in the transmembrane 2B protein of an insect picorna-like virus in a nonnative host

Oscar M. Lezcano^{1*}, Lara Fuhrmann^{2,3*}, Reinder T. Bos¹, Haitao Wang^{1,4},
Milou Stevens¹, Niko Beerenwinkel^{2,3}, Martijn A. Huynen⁵, Ronald P. van Rij¹

¹ Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, 6500 HB, Nijmegen, the Netherlands

² Department of Biosystems Science and Engineering, ETH Zurich, Klingelbergstrasse 48, Basel, 4056, Switzerland

³ SIB Swiss Institute of Bioinformatics, Quartier Sorge - Bâtiment Amphipôle, Lausanne, 1015, Switzerland

⁴ Institute of Plant Protection, Key laboratory of Food Quality and Safety of Jiangsu Province, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

⁵ Department of Medical BioSciences, Radboud University Medical Center, 6500 HB, Nijmegen, the Netherlands

* These authors contributed equally to this study.

Corresponding authors: niko.beerenwinkel@bsse.ethz.ch; martijn.huijnen@radboudumc.nl; ronald.vanrij@radboudumc.nl

Published in:

J Virol. 2026; 100(1):e0123925; doi: 10.1128/jvi.01239-25.

Abstract

Virus emergence is often due to cross-species transmission and adaptation to the new host. We studied the effect of innate immune responses on shaping virus populations in native and nonnative virus-host combinations, using as a model system *Drosophila melanogaster* infected with either *Drosophila* C virus (DCV) or cricket paralysis virus (CrPV). In this host, the cGAS-like receptor 1 senses viral double-stranded RNA and produces cyclic dinucleotides (CDNs) to activate the STING protein and induce an antiviral response. Both viruses were serially passaged in three host conditions differing in their cGAS-STING response: wild-type (WT) flies, *Sting* knock-out (KO) flies, and flies with a primed immune response by CDN injection. We found no immune-related effects on virus evolution, but we uncovered the CrPV nonstructural 2B protein as a key regulator of cross-species transmission. Nucleotide diversity specifically accumulated in the 2B gene during passage of CrPV in its nonnative *Drosophila* host, while 2B of the fly-adapted DCV displayed markedly lower and constant nucleotide diversity. In particular, the CrPV 2B D29N variant was selected in all six virus lineages evolved in WT and KO flies, with an estimated selection coefficient greater than 0.2. This variant replicated faster and was more lethal than the parental virus in all three host backgrounds. 2B is a predicted transmembrane protein, which we found to be associated with cellular endomembranes and may be involved in replication organelle formation. Our findings suggest a role for the 2B protein in adaptation to a new host independent of the cGAS-STING pathway.

Importance

The forces driving virus evolution are central to understanding cross-species transmission and virus emergence. It is well established that the adaptive immune system drives virus evolution in mammals, but whether innate responses likewise drive virus evolution upon host shifts is less well understood. In this manuscript, we used *Drosophila melanogaster* as a model to study the evolution of a native and a nonnative pathogen under conditions in which innate antiviral immunity is either abolished or enhanced. Using an experimental evolution approach, we find little evidence for adaptive evolution of the natural pathogen Drosophila C virus. In contrast, we observed a recurrent adaptive mutation in the viral nonstructural 2B protein in the nonnative cricket paralysis virus, independent of the antiviral cGAS-STING pathway. Our work provides insights into viral adaptation to new hosts and the characteristics of the 2B protein of dicistroviruses, a family comprising important model insect viruses.

Introduction

Emerging viral diseases are often the consequence of cross-species transmission. Host shifting is restricted by the phylogenetic distance between the hosts and often requires molecular adaptation to unrecognized host factors such as entry receptors and the immune system of the new host (1). Indeed, host shifts appear to occur most frequently among RNA viruses (2), likely due to their error-prone replication and adaptability. Also within natural hosts, the immune system is an important driver of virus-host coevolution, which is especially well established for the adaptive immune response of mammals. Models suggest that immune memory selects new surface antigens unrecognized by the host immune system, forcing the virus into uncharted antigenic space (3). Other simulations proposed that the immune response accelerates viral evolution during persistent infections (4). Indeed, in hepatitis C virus and human immunodeficiency virus-1 (HIV-1) infection, significant associations were observed between the emergence of specific viral mutants and specific human leukocyte antigen (HLA) types, likely a consequence of immune-driven selection of variants with decreased antigen-HLA binding affinity (5,6). However, also in organisms without adaptive immunity, host resistance may drive the evolution of escape variants. This is exemplified by an experimental evolution study in *Escherichia coli*, where the

bacteriophage λ repeatedly evolved the ability to use an alternative surface receptor in response to the evolution of host resistance due to reduced expression of the original LamB receptor (7).

In mammals, the main innate antiviral pathway is the interferon response (8), which can be activated by the cGAS-STING signalling pathway, an evolutionarily conserved pathway that originated in bacteria to control bacteriophage infections (9). The cGAS protein (cyclic GMP-AMP synthase) acts as a *pattern recognition receptor* that senses double-stranded DNA (dsDNA) in the cytosol and in response synthesizes cyclic di-nucleotides (CDNs), in particular 2'3'-cyclic GMP-AMP (cGAMP). CDNs are soluble molecules that act as second messengers, activating the conserved STING protein (stimulator of interferon genes). STING is an adapter protein embedded in the endoplasmic reticulum (ER). Upon activation, it translocates to the Golgi apparatus through an unknown mechanism that is essential for downstream signalling (reviewed in reference 10). This in turn leads to translocation of the transcription factor IRF3 (interferon regulatory factor 3) into the nucleus and the expression of type I interferons and interferon stimulated genes.

Despite lacking the interferon response, insects have a functional cGAS-STING pathway, albeit with marked differences compared to the mammalian pathway. For instance, *Drosophila melanogaster* encodes two cGAS-like receptors (cGLRs) capable of producing CDNs (11,12) and different *Drosophila* species may even encode up to seven cGLRs (13). Of these, cGLR1 is activated by double-stranded RNA (dsRNA), an intermediate in RNA virus replication and a hallmark of RNA virus infection (13). cGLR1 then produces 3'2'-cGAMP, a strong STING agonist that activates *Sting*-dependent signaling, leading to the translocation of the NF-KB transcription factor Relish to the nucleus and the transcriptional activation of STING-regulated genes (SRGs) (14). This response has been shown to limit viral RNA accumulation and extend survival of flies infected with several viruses, including the dicistroviruses *Drosophila* C virus (DCV) and Cricket paralysis virus (CrPV) (15), both positive-sense RNA viruses from the genus *Cripavirus* (*Dicistroviridae*).

Here, we aimed to gain a quantitative understanding of the specific effects of host shifts and innate immune adaptation on virus evolution. We performed an experimental evolution study of a native and a nonnative RNA virus in three *D. melanogaster* lines, differing in the activity of the cGAS-STING pathway. Along with the natural fly pathogen DCV, we used the nonnative CrPV,

originally isolated from Australian field crickets (*Teleogryllus oceanicus* and *Teleogryllus commodus*) (16). While CrPV has also been found in metagenomic data from Australian honey bees (17), it has to the best of our knowledge never been reported in wild-caught flies. We passaged the viruses for 10 fly generations, keeping 3 parallel lineages per virus and condition, after which the evolved lineages were characterized by RNA-sequencing. We found limited immune-related effects on virus evolution but observed high genetic diversity in the CrPV 2B gene. In particular, the conserved aspartic acid (D) at position 29 was substituted by an asparagine (N) in all virus lineages evolved in wildtype and *Sting* null mutant flies. We found that the CrPV D29N variant replicated faster and was more virulent than the wildtype CrPV strain in all fly conditions (virulence defined here as the capacity to cause host damage [18]). Our results suggest that viruses may not readily evolve resistance to cGAS-STING immunity and highlight the importance of the nonstructural 2B protein in viral adaptation to a non-native host species.

Results

Experimental CrPV and DCV evolution

To investigate the influence of the cGAS-STING pathway on host adaptation and virus evolution, we experimentally evolved DCV and CrPV in flies differing in their immune response: wild-type flies (w^{1118} , hereafter WT), *Sting* null mutant flies generated by CRISPR/Cas9 editing ($dSTING^{L76GfsTer11}$ (15), hereafter KO), and WT flies whose immune response was ectopically activated by injection of 3'2'-cGAMP (hereafter referred to as immune primed, IP) (Fig 1A). As anticipated (13–15), modulation of cGAS-STING pathway activity correlated with fly survival after DCV infection. *Sting* KO flies were more sensitive to infection than WT flies ($P < 0.0001$, Log-rank test; mean survival times 6.7 days for WT and 3.7 days for KO), and IP flies were less sensitive than WT flies ($P < 0.0001$, Log-rank test; mean survival time of 8.8 days for IP flies; Fig 1B). The protective effect of cGAMP injection was *Sting*-dependent as cGAMP injection in *Sting* KO flies did not affect survival after DCV infection (Fig 1B, KO-cGAMP), thereby also excluding that a possible direct interaction between cGAMP and the virus inoculum caused the protective effect of cGAMP in WT flies. Likewise, cGAMP injection extended fly lifespan after CrPV infection, although *Sting* inactivation did not affect survival ($P < 0.001$, Log-rank test; mean survival times of 6.8, 7.2, and 8.7 days for WT, KO, and IP, respectively; Fig 1C).

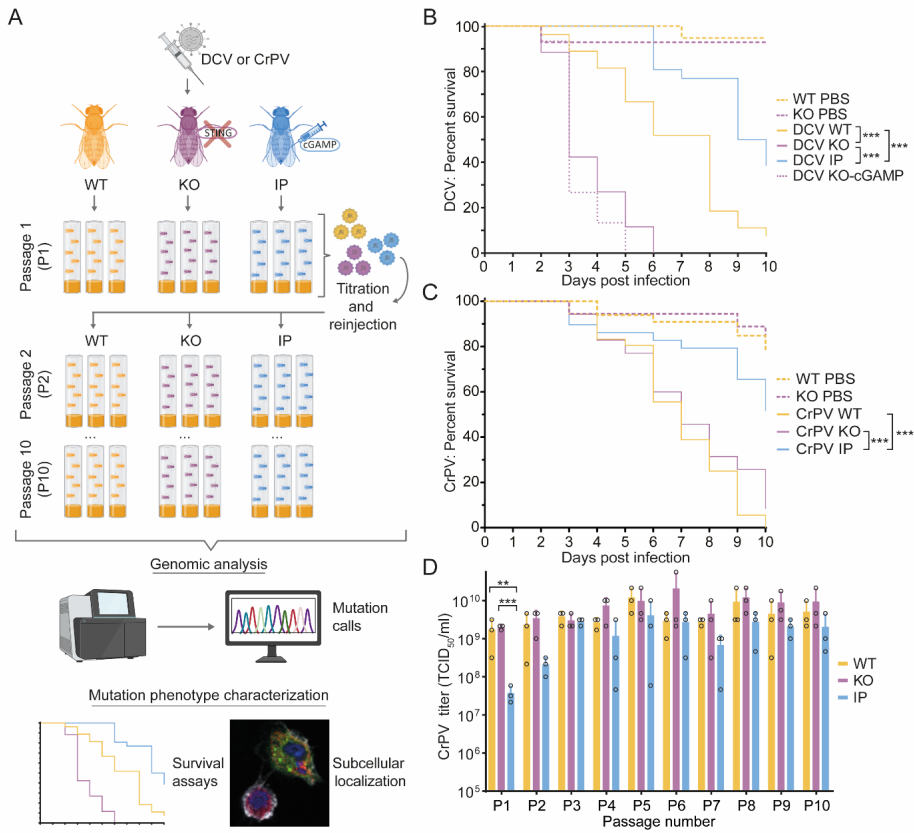


Figure 1. Experimental evolution of CrPV and DCV in *Drosophila melanogaster* with differing cGAS-STING response. **(A)** DCV or CrPV parental stocks were serially passaged in three distinct fly conditions: WT, *Sting* knockout (KO), and cGAMP immune primed (IP). This process was carried out with three virus lineages per condition for ten fly generations. Flies were intrathoracically inoculated, and viral titers were determined after each passage. The evolved virus populations were subjected to genomic analysis using next-generation sequencing, and the phenotype of the fixated mutations was studied. **(B–C)** Survival curves of WT, *Sting* KO, and IP flies infected with either 100 TCID₅₀ of DCV (B) or 1,000 TCID₅₀ of CrPV (C) ($n = 40$ flies per condition). PBS was taken along as a mock infection control. As an additional control, cGAMP was co-injected with DCV in *Sting* KO flies in (B). Log rank test P -values are indicated with asterisks: *** $P < 0.0001$. **(D)** CrPV titers in lysates of flies after each passage of all evolving lineages. Bars represent the means and SD of the three replicate viral lineages. Individual data points are shown as circles. Bonferroni adjusted P -values of the estimated marginal means test are indicated with asterisks: ** $P < 0.01$, *** $P < 0.001$.

We allowed virus populations to evolve during ten fly generations (passages), with three independent lineages per host condition (Fig 1A). The DCV parental stock was a host-adapted strain generated by serial dilution of the virus evolved in WT flies for 10 passages from our previous study (19). The CrPV parental

stock was generated from an infectious molecular clone (20). Virus titers were measured after each passage to standardize the inoculum for subsequent passages and to minimize population bottlenecks and differences in the rate of genetic drift. Viral titers of CrPV and DCV after each passage generally showed small differences between fly conditions across most passages (CrPV: Fig 1D; DCV: Fig S1A), likely due to virus titers reaching a plateau at this time point (three days post infection). The exception was the first passage of CrPV evolution, where the viral titers in IP flies were significantly lower than in WT and KO flies (61 and 71-fold, respectively; $P\text{-adj} < 0.005$, estimated marginal means for pairwise comparison; Fig 1D). In fact, viral titers were barely high enough to inoculate the next generation with the desired 1,000 TCID₅₀ inoculum and we therefore decided to increase the CrPV inoculum to 5,000 TCID₅₀ from passage three onwards.

Increasing DCV synonymous nucleotide diversity

We studied molecular evolution during the experiment by RNA-sequencing at passages 1, 2, 3, 5, and 10, and computed viral nucleotide diversity across passages. We obtained a mean read depth across the genomes of 47,501 and 29,539 for the DCV and CrPV samples, respectively, which was consistent across samples (Fig S2). For DCV populations evolved in WT flies, we incorporated the diversity values from our previous study (19). As expected, serial dilution used to prepare the parental virus (P0) reduced nucleotide diversity, consistent with population bottlenecks (Fig 2A). By weighing nucleotide diversity per site according to the fraction of possible synonymous or non-synonymous substitutions, we found that the serial dilution specifically reduced non-synonymous nucleotide diversity (Fig S1C). Afterwards, we observed a significant increase in nucleotide diversity as passages progressed across all DCV populations, regardless of their host condition (Fig 2A, $P < 0.0001$, Linear mixed model, Table S1). Increasing nucleotide diversity was predominantly explained by the accumulation of synonymous single-nucleotide variants (SNVs) (Fig S1B–S1D). Only a single SNV in the 3' untranslated region accumulated in all lineages, but this SNV preexisted in the parental virus. However, as we observed no indications for adaptive DCV evolution during the experiment, such as the condition-specific fixation of SNVs or the accumulation of nucleotide diversity, we focused our subsequent analyses on the evolved CrPV populations.

Transient increase and subsequent collapse of CrPV nucleotide diversity

In the CrPV parental stock, we identified a total of 244 SNVs with respect to the consensus sequence of its 9.2-kb genome, of which four occurred at frequencies

above 10% (Fig S3A). As the parental stock was produced from an infectious clone, these SNVs must have accumulated during *in vitro* transcription or the 48 hours of replication in *Drosophila* S2 cells used to prepare the stock.

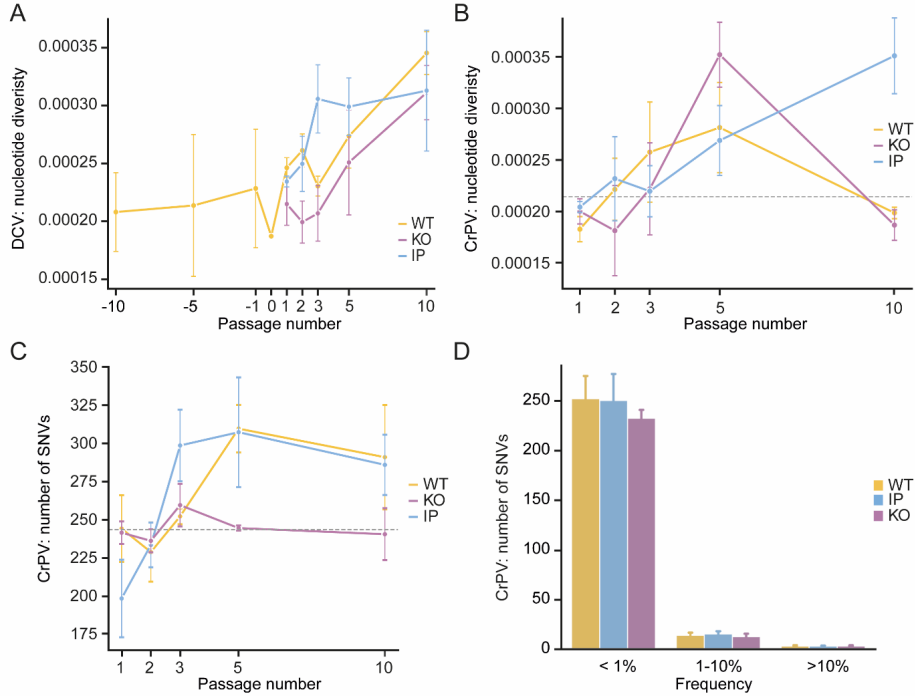


Figure 2. Host condition specific diversity trends in viral populations during experimental evolution. **(A)** Mean nucleotide diversity of evolved DCV populations. Diversity values from DCV evolved in WT flies from our previous study (19) are integrated and denoted as passages -10, -5, and -1. Passage 0 is the parental stock and passages 1 to 10 are the evolved DCV populations of the current study. Data are shown as means and standard error across the three replicate lineages for each genotype. **(B-C)** The mean nucleotide diversity (B) and number of single-nucleotide variants (SNVs) (C) of evolved CrPV populations from WT, *Sting* KO, and IP flies across passages. Data are shown as means and SEM across the three replicate lineages for each genotype. The dashed lines mark the number of SNVs and mean nucleotide diversity in the parental CrPV stock. **(D)** Frequency distribution of SNVs for CrPV virus populations from the indicated host conditions across all passages and replicates. Bars indicate means and error bars indicate 95 percent confidence intervals based on the bootstrap distribution.

As observed for DCV, nucleotide diversity significantly increased for CrPV with passage number ($P < 0.001$, Estimate = 0.05, Linear mixed model, Table S2). However, there was a significant decrease in diversity between passage 5 and 10 for populations evolved in WT and *Sting* KO flies ($P = 0.001$, Estimate = -0.39, Linear mixed model). A significant interaction was observed

between the IP background and passage 10 ($P = 0.01$, Estimate = 0.25, Linear mixed model), indicating a sustained increase in nucleotide diversity of viral populations evolved in IP flies at later passages, in contrast with the sharp decline observed in WT and KO flies (Fig 2B, Fig S3B). At the same time, we observed increasing numbers of SNVs from passage 1 onwards in the populations evolved in WT and IP flies (Fig 2C) with a mean of 244 and 199 SNVs at passage 1, respectively, which increased to a mean of 291 and 286 at passage 10. In contrast, for CrPV populations evolved in *Sting* KO flies, the mean number of SNVs remained roughly at the same level as the parental stock. The vast majority of SNVs occurred at low frequency (median: 0.07%, Fig 2D), whereas only between 9 and 28 SNVs occurred at a frequency above 1%, and between 1 and 5 SNVs at a frequency above 10%, across all host conditions, passages and replicates (Fig 2D).

Specific accumulation of nucleotide diversity in the CrPV 2B and RdRp genes

We calculated the nucleotide diversity for each viral gene and each condition (Fig 3A, Fig S4), and found that the mean nucleotide diversity varied significantly between viral genes ($P\text{-adj.} < 0.001$, two-way mixed ANOVA), but not between host conditions ($P\text{-adj.} = 0.202$). Post hoc pairwise comparisons (methods, Table S3) revealed that the 3D gene encoding the RNA dependent polymerase (RdRp) and the 2B gene showed significantly increased diversity compared to seven and ten other genes, respectively (Table S3). Additionally, 2B was the only gene that showed higher diversity in the evolved populations across all three host conditions than in the parental stock (Fig S4). Moreover, the diversity in 2B was significantly higher compared to the remaining coding region (Fig 3B, $P\text{-adj.} = 0.0004$, Estimated marginal means with Tukey's adjustment), indicating that the dynamics in 2B dominate the diversity in the whole genome. Conversely, in evolved DCV populations, we observed the opposite trend: the diversity in 2B was significantly lower than in the remaining coding region (Fig 3B, $P\text{-adj.} = 0.0024$, estimated marginal means with Tukey's adjustment). Furthermore, a direct comparison of 2B diversity between DCV and CrPV populations revealed significantly lower diversity in DCV (Fig 3B, $P\text{-adj.} < 0.0001$, Estimated marginal means with Tukey's adjustment), suggesting the importance of the 2B gene in CrPV adaptation to the *Drosophila* host and stringent selection on 2B in DCV.

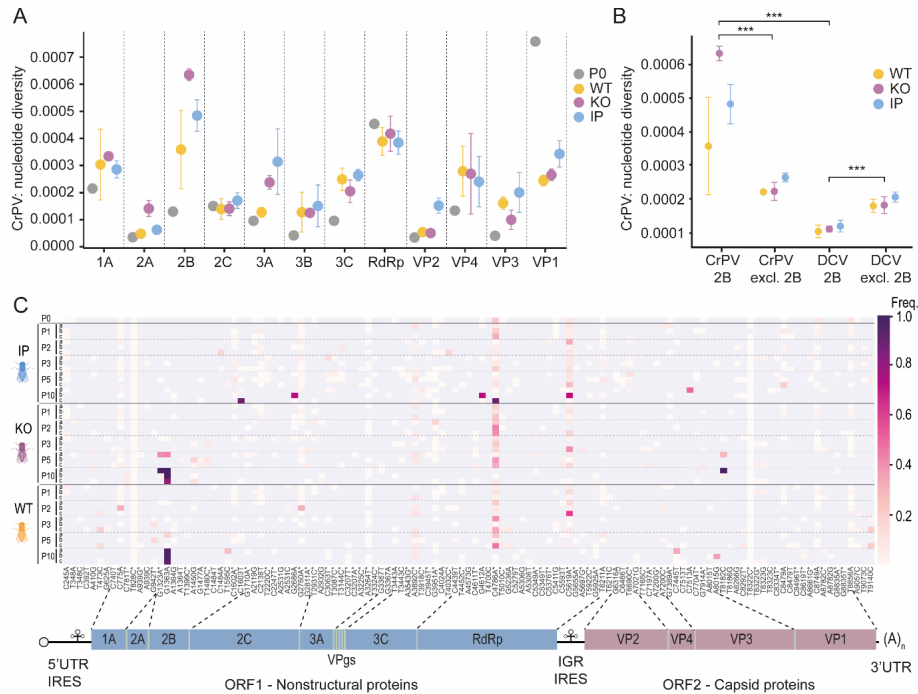


Figure 3. Nucleotide diversity accumulates specifically in the CrPV 2B gene. **(A)** Nucleotide diversity for each CrPV gene. For each host background, the mean nucleotide diversity of the three replicate lineages over all passages was computed. Data are shown as means and standard errors across the three replicate lineages for each condition. **(B)** Nucleotide diversity in the viral 2B gene and the remaining coding region in the evolved DCV and CrPV populations. Data are shown as means and standard error across the three replicate lineages for each host background. Bonferroni adjusted *P*-values of the estimated marginal means test comparing the four categories are indicated with asterisks: *** *P* < 0.001. **(C)** Heatmap of SNV frequencies in CrPV populations evolved in WT, *Sting* KO, and IP flies across the viral genome. Lowercase letter a, b, c indicate the three independent lineages. Non-variable sites are indicated in gray. Non-synonymous mutations are marked with an asterisk. Only SNVs with a frequency greater than 0.01 in at least two samples are shown. Lower panel: schematic representation of the CrPV genome. Gene coordinates follow Uniprot annotation (Replicase polyprotein, Q9IJX4 and structural polyprotein, P13418); capsid protein names are according to ICTV (21), which differs from the P13418 entry. IRES, internal ribosome entry site; IGR, intergenic region.

Independent fixation of CrPV 2B D29N in six lineages

A complex pattern of SNV retention and emergence was observed along the CrPV genome across different lineages and passages (Fig 3C). Many SNVs present in the parental stock were retained in most lineages (e.g., C775A, T828C, C1602A, G2760A, A3264T, C4786A), while others showed temporal variations. For instance, C8749A and T8856A persisted until passage 5 but were absent in passage 10, and G4612A was retained only in populations from

IP flies. We observed the *de novo* appearance of A3892C, a silent mutation of Arg154 in the 3C protease (AGA to CGA), in all host conditions across most passages and replicates at frequencies of 2.0–14.9%. Another interesting *de novo* SNV was G1363A resulting in an aspartic acid to asparagine substitution at position 29 of the 2B protein (D29N). This substitution emerged at passage 2 in WT flies for the first time at a frequency of 0.03% and in *Sting* KO flies at passage 3 at a frequency of 0.05%. In all replicate lineages in WT and *Sting* KO flies, this mutation was fixed at passage 10 at frequencies above 82%. In contrast, in viral populations from IP flies, G1363A was only detectable in one replicate lineage at a frequency of 0.2% at passage 3 but rather than increasing in frequency, it became extinct in the following passages. In one of the lineages from *Sting* KO flies (KO-a), we detected an additional synonymous mutation, G1323A in close proximity to and at similar frequencies as the G1363A substitution. Local haplotype analysis confirmed the co-occurrence of G1363A mutations at passages 5 and 10 (2101/2501 and 4792/4802, and G1323A in a high proportion of reads that carry at least one of those respectively), suggesting that G1323A is a hitchhiking mutation.

D29N causes a marked increase in viral virulence

The recurrent fixation of D29N in multiple independent lineages suggests that it confers an evolutionary benefit to the virus. To experimentally study the phenotype of D29N, we introduced the mutation into the infectious clone and compared the growth kinetics of WT CrPV and the D29N mutant in WT, *Sting* KO and IP flies. The D29N mutation significantly increased viral RNA levels ($P < 0.0001$, Linear mixed-effects model, Table S4), with an estimated 5 to 10-fold increase in all three host conditions compared to the WT virus (Fig 4A). Stimulation of the cGAS-STING pathway by co-injection of 3'2'-cGAMP significantly reduced CrPV RNA levels by a magnitude of 26.7 for both variants in WT flies ($P = 0.0014$, Linear mixed-effects model, Table S4), indicating that D29N did not affect sensitivity to cGAS-STING mediated immunity. In line with the increased viral replication rate, survival of flies infected with the D29N mutant virus was significantly reduced compared to WT CrPV infected flies in all host backgrounds ($P < 0.0001$, Log-rank test; mean survival times of 8.2, 8.2, and 9.0 days for WT, KO, and IP flies infected with WT CrPV and 4.8, 5.1 and 5.7 for WT, KO, and IP infected with the CrPV D29N), although 3'2'-cGAMP co-injection still extended fly survival (Fig 4B).

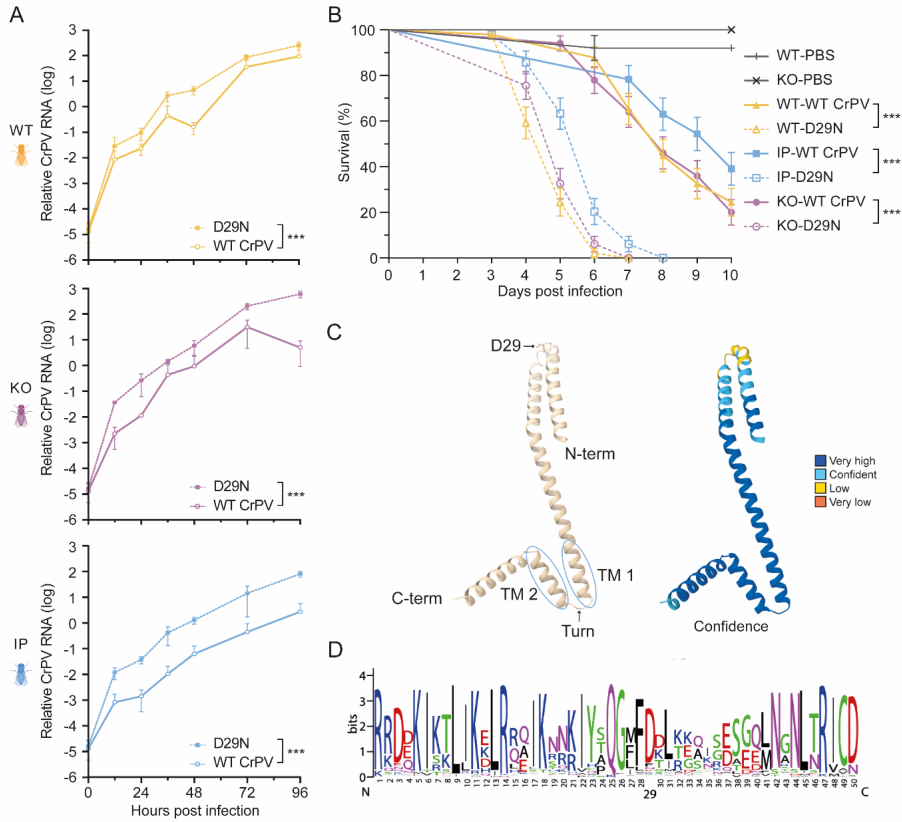


Figure 4. Higher replication kinetics and virulence of the CrPV D29N variant. **(A)** Growth curves of WT CrPV and D29N mutant in WT, KO, and IP flies inoculated with 1,000 TCID₅₀. Three pools of five flies were harvested per time point, and viral RNA was quantified by RT-qPCR. The mean and SD of CrPV RNA levels relative to *rp49* are shown. The difference between WT and D29N CrPV was statistically significant in a linear model ($P < 2.75 \times 10^{-10}$). **(B)** Survival curves of WT, *Sting* KO, and IP flies inoculated with PBS or 1,000 TCID₅₀ of either WT or D29N mutant CrPV. Two to five-day-old females ($n = 50$) were intrathoracically inoculated and monitored daily for survival. Log rank test P -values are indicated with asterisks: *** $P < 0.0001$. **(C)** AlphaFold 3 prediction of the CrPV 2B structure as a monomer. The position of the D29 residue and both predicted transmembrane α -helices (TM) are indicated. Color coding in the right panel indicates the confidence level of the prediction: yellow, low confidence ($70 > \text{pLDDT} > 50$); light blue, confident ($90 > \text{pLDDT} > 70$); dark blue, very high confidence ($\text{pLDDT} > 90$). **(D)** Sequence logo of the first 50 amino acids of the 2B protein and 131 homologous proteins identified by BLASTp.

We approximated the selection coefficient of the D29N mutant relative to WT CrPV using two independent approaches: tracking changes in nucleotide frequencies during the experimental evolution study (Fig 3C) and comparing the replication rates of WT and D29N mutant CrPV that we estimated assuming exponential growth (Fig 4A, Methods). Both methods yield similar results,

with mean estimates of 0.287 and 0.267 for the frequency and RNA-based estimates, respectively, indicating increased fitness of the D29N mutant virus (Table S5). Overall, these data indicate that D29N causes an increase in viral replication and virulence.

2B is a predicted transmembrane protein that localizes to endomembranes

No experimental data exist about the function of CrPV 2B or any of its homologs including DCV 2B that we identified using PSI-BLAST (22) or HHPred (23). Picornaviruses such as coxsackievirus and poliovirus (*Enterovirus* genus) also encode a protein named 2B that, as in CrPV, lies directly upstream of the 2C RNA helicase. To find evidence for homology between CrPV 2B and enterovirus 2B, we ran HHPred in the *pairwise profile alignment* mode, in which we compared CrPV 2B with the complete coxsackievirus A10 (CVA10) and poliovirus polyproteins. CrPV 2B specifically aligned with the poliovirus 2B protein and the CVA10 2B protein (Fig S5A–S5B), albeit with high E-values ($E = 0.55$ and $E = 0.24$, respectively). The region that is most similar between the 2B proteins is a stretch of 49 amino acids that corresponds to the two predicted transmembrane (TM) helices in CrPV 2B (Fig S5A) (24) and two experimentally determined TM helices in the coxsackievirus and poliovirus 2B proteins (25,26). The D29 residue lies upstream of these helices and appears well conserved among homologs that could be detected with BLASTp (Fig 4C). AlphaFold structure prediction (27), whose modelling accuracy improved by the addition of lipids, as well as HHpred (Fig 4C, Fig S5C), suggested that all three 2B proteins are predominantly α -helical. They likely adopt a helix-turn-helix motif that traverses the organelle membranes and would expose D29 towards the cytoplasm (26). Based on its tentative homology with the enterovirus 2B proteins, we hypothesized that the CrPV 2B protein localizes in the endomembrane system of the host cell (25,26). To experimentally validate this, we tagged the WT and D29N 2B proteins with eGFP at either the N or C terminus and studied their subcellular localization, using *Drosophila* proteins known to localize in the ER (Calnexin 99a), cis-Golgi (GM130) and lysosomes (Rab7) as markers (Fig 5). We observed that 2B localized in the ER (red), but also showed punctuate patterns corresponding to the cis-Golgi (Fig 5A, magenta), lysosomes (Fig 5B), and other membranous organelles, whose identity remains to be established. The D29N mutation did not seem to drastically affect the subcellular localization of the 2B protein.

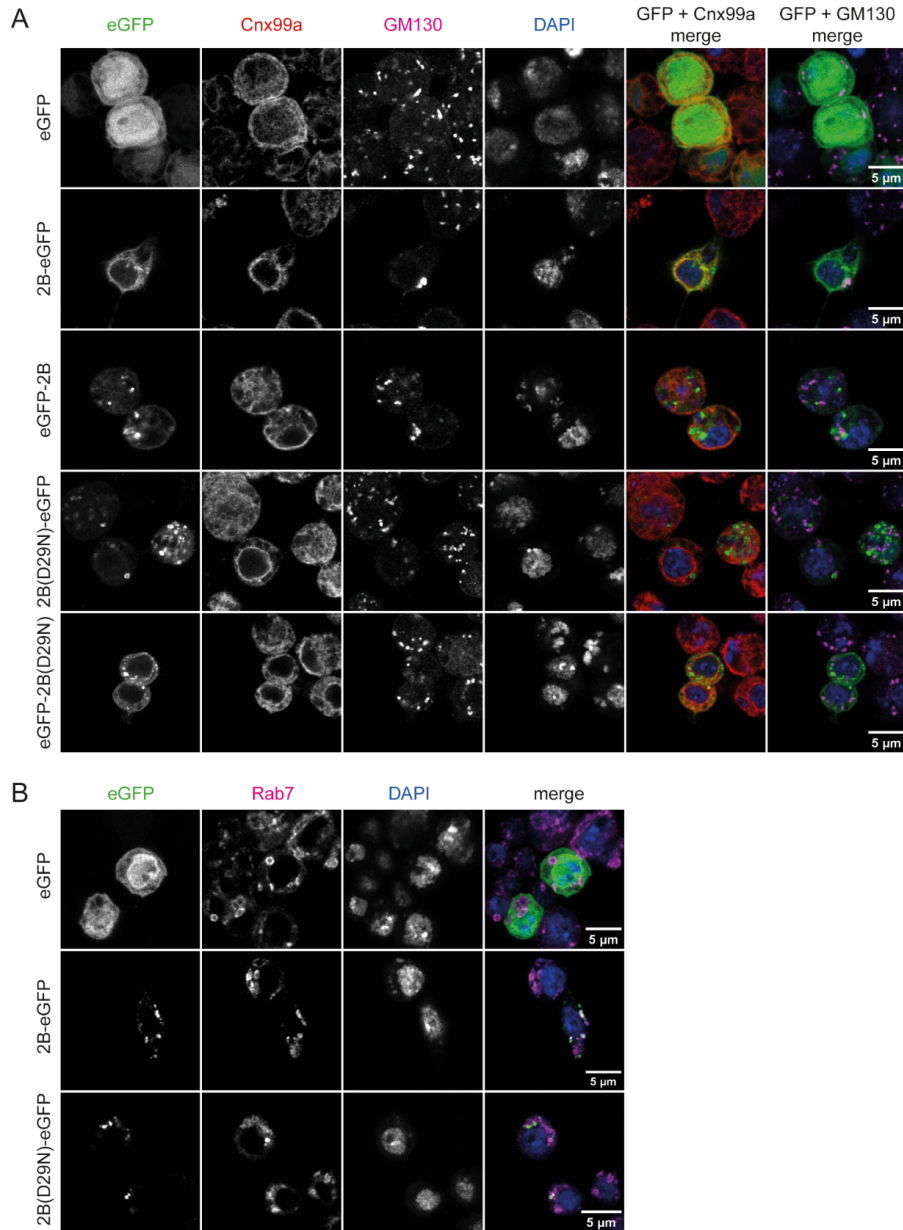


Figure 5. Subcellular localization of CrPV 2B WT and D29N mutant. Confocal microscopy images of *Drosophila* S2 cells transfected with plasmids encoding the indicated proteins. Cells were fixed 48 h after transfection and stained with **(A)** antibodies against the endoplasmic reticulum protein Calnexin 99A (Cnx99a) and the cis Golgi protein GM130, and **(B)** antibodies against the lysosome protein Rab7 (magenta). Nuclei were stained with DAPI. In the merged images in (A) eGFP is shown in green, Cnx99a in red, GM130 in magenta, and DAPI in blue.

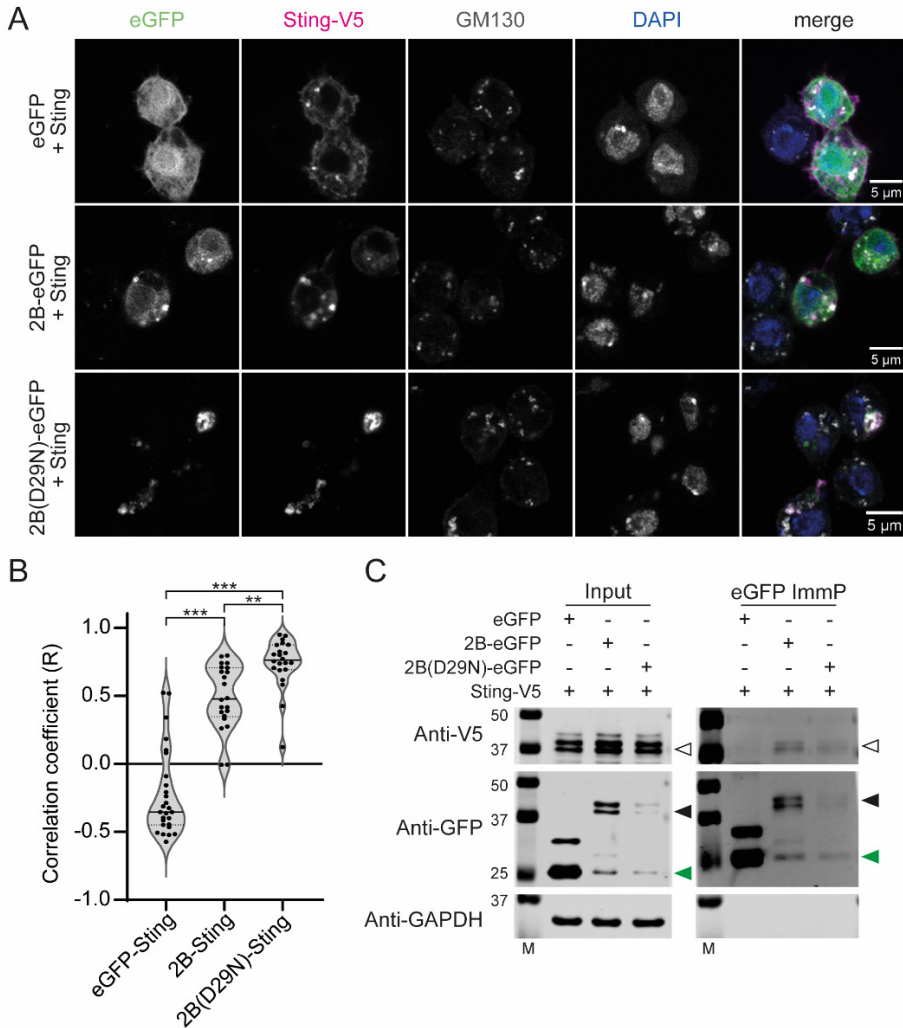


Figure 6. Sting colocalizes with the CrPV 2B protein. **(A)** Confocal microscopy images of *Drosophila* S2 cells transfected with a combination of the indicated plasmids. Cells were fixed 48 h after transfection and stained with antibodies against the V5 tag and the cis Golgi protein GM130. **(B)** Violin plot of Pearson's correlation coefficient (R) of Sting with eGFP (n = 27), 2B-eGFP (n = 22) or 2B(D29N)-eGFP (n = 22) in individual cells, calculated in Fiji using the Costes method. The median is represented with a solid black line and the quartiles with dashed lines. Adjusted *P*-values from Tukey's multiple comparison test are indicated with asterisks: ** *P* < 0.01, *** *P* < 0.001. **(C)** Co-immunoprecipitation of Sting with 2B and 2B(D29N). S2 cells were transfected with the indicated plasmids and the cell lysates (input) and eGFP immunoprecipitates (eGFP ImmP) were analyzed by western blot using eGFP, V5, and GAPDH antibodies. The white arrow heads indicate Sting-V5 (expected size 43.1 KDa), black arrow heads indicate 2B-eGFP and 2B(D29N)-eGFP (43.3 KDa) and green arrow heads eGFP. As previously reported for human STING, *Drosophila* Sting showed different migrating bands that might be due to post-translational modifications (28). M, size marker.

Considering that human STING is embedded into the ER membrane, we analyzed whether *Drosophila* Sting would interact with CrPV 2B. We epitope-tagged Sting and co-expressed the fusion construct with either the WT and mutant 2B-GFP fusion proteins or an eGFP control in S2 cells. The punctate pattern of 2B was highly similar to the pattern produced by the Sting protein (Fig 6A), which was validated by colocalization analysis (median Pearson's correlation coefficient $R = 0.48$; Fig 6B). Strikingly, an even stronger correlation was observed for the 2B D29N mutant ($R = 0.76$; Fig 6B). The interaction between 2B and Sting was further validated by a co-immunoprecipitation assay. We found that the Sting protein was recovered in immunoprecipitations of both the WT and D29N 2B-GFP proteins, but did not coprecipitate with the eGFP control alone (Fig 6C). Together, these data suggest that the 2B protein localizes in cellular endomembranes, where it interacts with Sting. However, given that the D29N mutation was also selected in flies lacking Sting, it is unlikely to represent an adaptation to optimize 2B interactions with Sting.

Discussion

In this study, we have explored the impact of the cGAS-STING pathway on virus evolution using *D. melanogaster* infected with DCV or CrPV as a model. We allowed the viruses to adapt to flies with a modulated cGAS-STING response, hypothesizing that variations in the pathway activity would affect virus evolution at the genomic and phenotypic level. We found no effect of the STING pathway on DCV evolution, but instead noted a gradual increase in synonymous nucleotide diversity in all host conditions. Likewise, for CrPV we found no evidence for adaptation that was specific to different cGAS-STING regimes. In contrast, we observed that nucleotide diversity accumulated especially in the CrPV 2B gene in all conditions, while DCV exhibited reduced diversity in 2B compared to the remaining coding regions, consistent with findings from our previous study (19). Specifically, for CrPV, we detected an SNV resulting in a D29N amino acid substitution in the 2B protein that was selected in all viral lineages from WT and *Sting* KO flies, while its frequency did not increase in viral populations from IP flies, despite being detected in one lineage at an early passage. CrPV mutants harboring D29N replicated faster and were more virulent than the WT virus in all host backgrounds. The appearance of parallel adaptive mutations driven by cross-species transmission has been previously observed in experimental evolution studies of DCV adapting to different *Drosophila* species (29) and vesicular stomatitis virus adapting to either

human or dog cells (30). As *Drosophila* is not the natural host of CrPV, our data suggest that the 2B mutation likewise emerged to facilitate adaptation to its nonnative host.

We found that DCV nucleotide diversity continued to increase with passage number, suggesting that the virus populations had not reached equilibrium yet. Equilibrium is reached when the gain of genetic diversity is equal to the loss caused by purifying selection and random drift (31). This state seems difficult to attain in virus populations, as nucleotide diversity not only increases during previous relatively short-term evolutionary experiments (32–34), but also in long-term experiments, for example, HIV-1 experimental evolution over more than three years, during which mutations continue to accumulate (38). Likewise, the number of polymorphic sites also continued to accumulate in anellovirus over a period of 30 years in chronically infected patients (36). Therefore, the time scale necessary for a virus to achieve equilibrium is likely too long to be achieved in our experiment.

In 6 of the 18 virus lineages in our study, nucleotide diversity decreased from passage 5 to passage 10. A common factor of those virus populations was the fixation of the adaptive D29N mutation in the CrPV 2B gene. We hypothesize that the diversity decrease was caused by a selective sweep (37), as previously observed during the fixation of drug resistance variants in HIV-1 (38). The observation that the D29N mutant was not selected in IP flies despite its higher replication rates across all host conditions, might be explained by altered evolutionary dynamics, due to a reduction of the viral effective population size upon the stimulation of cGAS-STING pathway. Under these conditions, evolution is expected to be dominated by random drift rather than deterministic selection occurring under large population sizes (39,40). While we attempted to estimate variance effective population size from temporal changes in synonymous allele frequencies, our experiment involved selection effects and population fluctuations that violate key assumptions of standard effective population size estimation methods (40), making such estimates unreliable in our system (41–43). We did observe, nonetheless, a significantly reduced viral RNA load in IP flies, suggesting a smaller total population size, which may contribute to the observed evolutionary patterns.

It is currently unclear why the 2B mutation facilitated CrPV host adaptation. The function of the 2B protein of members of the *Dicistroviridae* family has not been studied. However, their corresponding relative positions in the viral

genome, putative homology based on sequence-profile comparisons, and the location of (predicted) transmembrane regions suggests that CrPV 2B is homologous to 2B of enteroviruses, although direct evidence is lacking. Enterovirus 2B is an ion channel-forming protein, typically 50–120 amino acids in length. It containing two hydrophobic regions that adopt a helix-turn-helix conformation, enabling insertion into cellular membranes (reviewed in reference 44). The protein also has at least one amphipathic α -helical structure that can oligomerize to form transmembrane hydrophilic pores (45,46), possibly as a tetrameric complex (47–50). Enterovirus 2B has viroporin activity, which is associated with increased membrane permeabilization, disrupted calcium homeostasis, and induction of apoptosis and autophagy (51). In addition, enterovirus 2B is associated with viral replication organelles, thought to be derived from modified ER and Golgi apparatus membranes (52). We indeed observed localization of CrPV 2B in cellular endomembranes, but whether the protein is also a viroporin and how mutations in such a protein would facilitate host adaptation remains to be studied.

The 2B D29N mutation was selected in both WT and *Sting* KO flies and the mutant conferred a growth advantage to CrPV also in *Sting* KO flies, indicating that 2B D29N-mediated host adaptation is independent of the cGAS-STING pathway. It was therefore rather unexpected to find that CrPV 2B colocalizes with *Sting*, and that the colocalization was even stronger for the D29N 2B mutant. While this conundrum awaits to be solved, it is interesting to note that enterovirus 2B has been shown to affect STING localization in mammalian cells. Specifically, the 2B protein of rhinovirus was found to reduce Ca^{2+} levels in the ER, triggering the relocation of STING to the replication organelles (53), where it acts as a proviral factor to promote viral replication (28).

Overall, our work provides new insights into viral adaptation to a new host and the characteristics of the 2B protein of dicistroviruses, a family comprising important model viruses to dissect insect virus-host interactions.

Materials and Methods

Fly strains and husbandry

Flies were maintained at 25°C in standard fly food. Eggs were bleached and treated with tetracycline as described before (54), and the absence of Nora virus, DCV, *Drosophila* X virus, CrPV and *Wolbachia* was verified as described

earlier (19). Flies containing the *dSTING*^{L76GfsTer11} null allele (15) (referred to as *Sting* KO in this study) were a kind gift from Jean-Luc Imler (CNRS-Université de Strasbourg). *w*¹¹¹⁸ flies were used as wild-type (WT), as the *Sting* KO flies had been isogenized in this genetic background according to the procedure described for cGLR1 and cGLR2 in (11). Immune-primed flies were generated by the intrathoracic injection of WT flies with 69 nl of a 1 µg/µl solution of c[A(2',5') pG(3',5')p] (3'2'-cGAMP) (BioLog). Two to five-day-old female flies, likely mated as they were obtained from a vial containing unsorted flies, were used for the experimental evolution study, the virus infection and the survival assays.

Plasmids

The CrPV 2B coding sequence was amplified by PCR from cDNA of infected *w*¹¹¹⁸ flies and cloned into fusion vectors expressing eGFP under the control of the Actin promoter using In-Fusion cloning (TAKARA). For the construction of pAct-2B-eGFP plasmid, the 2B sequence was amplified using the forward primer 2B-eGFP-For (GGTACCTACTAGTCCcaccATGCGCCGAGATGAGAAAATTTCAACCTT), containing the Kozak sequence and the start codon, and the reverse primer 2B-eGFP-Rev (GCCCTTGCTCACCAtggatccacctgatccgccTTGGGTTGTTGTCTGCAAAATTTTGT) containing a linker peptide (GGSGGS). The PCR product was cloned into the pAct-eGFP-C backbone linearized with primers vectorFor (ATGGTGAGCAAGGGCGAG) and vectorRev (GGACTAGTAGGTACCCGATCC). For the construction of pAct-eGFP-2B plasmid, the 2B sequence was amplified using the forward primer eGFP-2B-For (GACGAGCTGTACAAGGGCGGATCAGGTGGATCCCGCCGAGATGAGAAAATTTCAACCTT) containing the linker peptide and the reverse primer eGFP-2B-Rev (AGCTCAGGCCTTAGAttaTTGGGTTGTTGTCTGCAAAATTTTGT) containing a stop codon. The PCR product was cloned into the pAct-eGFP-N backbone linearized with primers vectorFor (TCTAAGGCCTGAGCTCGCT) and vectorRev (CTTGTA CAGCTCGTCCATGCC). The mutant fusion constructs pAct-2B(D29N)-eGFP and pAct-eGFP-2B(D29N) were generated by site directed mutagenesis by amplifying either pAct-2B-eGFP or pAct-eGFP-2B with primer For (GGTTTCTTTAATGATCTCAAAGGAGCAAAAGG) and Rev (TAAAGAAACCTTG GGTGTAAATTCTGTTG). The linearized PCR product was recombined using In-Fusion cloning (TAKARA).

The pCrPV(D29N) infectious clone was generated by amplifying the 2B(D29N) sequence from the plasmid pAct-2B(D29N)-eGFP using primers For (GAGAAAATTTCAACCTTGATTAAGAAG) and Rev (AATCAAGGCGGCACGATAC) and the backbone from pCrPV-3 (20) with primers For (CGTGCCGCCTTGAT

TGTAAT) and Rev (GTTGAAATTTTCTCATCTCGGCG) using In-Fusion cloning (TAKARA).

The *Sting* coding sequence was amplified by PCR from cDNA of *w*¹¹¹⁸ flies and cloned into pAc5.1-V5-His-A (Invitrogen) using In-Fusion cloning (TAKARA) to produce Sting-V5. The forward primer (GATCGGGGTACCTAC ccaccATGGCAATCGCTAGCAACGT) contained the Kozak sequence and the reverse primer (AGGGATAGGCTTACCggatccacctgatccgccGTTGGAAATTTTCG TCAATAGTTTTGGTTTTGTTT) contained a sequence encoding a linker peptide (GGS GGS).

***In vitro* transcription and transfection of the CrPV infectious clone**

The plasmid encoding the CrPV molecular clone (pCrPV-3 [20]) was a kind gift from Eric Jan (University of British Columbia). *In vitro* transcription and transfection were performed as previously described (20). Briefly, the plasmids were amplified in Stellar Competent *E. coli*, purified, linearized with Ecl136II (Invitrogen), and used for *in vitro* transcription at 30 °C for 3 h using the RiboMAX kit (Promega). The RNA was treated with DNase RG1 (15 min at 37 °C) and purified using RNeasy columns (Qiagen). The viral RNA was transfected into *Drosophila* S2 cells (Invitrogen) in 6-well plates containing 2 ml of Schneider's medium. After 48 h, the supernatant was collected, centrifuged at 300 *g* for 5 min, aliquoted, titered, and stored at –80 °C.

Cells and virus stocks

The parental DCV and CrPV stocks were produced in *Drosophila* S2 cells (Invitrogen), which were maintained in Schneider's medium supplemented with 10% fetal calf serum (Sigma) and 50 U/ml penicillin and 50 µg/ml streptomycin (Gibco) at 27 °C. DCV previously passaged in WT flies for 10 generations (19) was bottlenecked by serial dilution in a 24-well plate, selecting the highest dilution that still showed cytopathic effects to inoculate a T25 flask of confluent S2 cells containing 7 ml of medium. After 72 h, the supernatant was collected, centrifuged at 300 *g* for 5 min, aliquoted, and stored at –80 °C. Titers were determined by end-point dilution in 96-well plates, as described before (54), and expressed as median tissue culture infectious dose (TCID₅₀), calculated using the Reed–Muench method.

Experimental evolution

Thirty females of each host condition were intrathoracically inoculated with 1,000 TCID₅₀ of DCV or CrPV in 69 nL of PBS, pH 7.3, using a Nanoject II

microinjector (Drummond Scientific). For IP flies, the 3'2'-cGAMP was co-injected with the virus inoculum into WT flies. Three lineages per host condition were used as independent replicates. For the first passage, flies were inoculated with the respective parental stock, and from P1 onwards, all eighteen lineages were kept independently (nine lineages per virus). After each passage, pools of ten flies were snap-frozen and homogenized two times for 10 s in 220 μ l of PBS, using 1 mm silica beads in a Precellys homogenizer, centrifuged at 16,000 g for 10 min to discard fly debris, and the supernatant was transferred to a fresh tube, aliquoted, stored at -80°C , and directly titrated. The next generation of flies was then inoculated through intrathoracic inoculation. In the case of CrPV, the virus inoculum was increased to 5,000 TCID₅₀ from passage three onwards, as the titers after the first and second passage were barely high enough to reach the desired inoculum for the next passage in the IP flies. For RNA extraction and preparation of next-generation sequencing (NGS) library of the parental stock, 1 ml of TRI Reagent (Sigma) was added to 100 μ l of virus stock, while for the viral lineages, pools of five flies were homogenized in 1 ml of TRI reagent (Sigma) using 1 mm silica beads in a Precellys homogenizer. RNA was isolated according to the manufacturer's instructions.

Virus infections and survival assays

To study virus growth kinetics, 100 TCID₅₀ of DCV or 1,000 TCID₅₀ of CrPV WT or D29N mutant were injected into two to five-day-old WT, *Sting* KO and IP flies. For the IP condition, 3'2'-cGAMP was co-injected with the virus. Three pools of five flies per condition were harvested at the indicated time points and processed for RT-qPCR. To study virulence, the same dose and procedure were applied to fifty flies per condition. Survival was monitored daily and analyzed using the Kaplan-Meier estimator. The difference between the curves was compared using the Log rank test in GraphPad Prism 10.

RT-qPCR

RT-qPCR was performed as previously described (19), using primers CrPVFor (ACGAGGAAGCAACTCAAGG) and CrPVRev (GAGCCCGCTGAGATGTAAAG) for CrPV RNA quantified.

NGS library preparation

Total RNA from passages 1, 2, 3, 5, and 10 was extracted using TRI Reagent (Sigma) and reverse transcribed with superscript IV (Thermo Fisher) and oligo (dT). Viral cDNA was amplified in four overlapping amplicons of about 2.2 kb. The primers used for CrPV were as follows:

For1, 5'-CTCCCCGTGAGAAACCTTGTTT-3';
 Rev1, 5'-GTGTTTGTAAAGCGTCGGGTTTG-3';
 For2, 5'-GATCCCGGACCGAGACATTG-3';
 Rev2, 5'-TTACCGCCTGACCAACCTTG-3';
 For3, 5'-ACGGATATGCTTGCCCCCTAAC-3';
 Rev3, 5'-TGGGGTGAAACATAGGGAATTCTC-3';
 For4, 5'-CTTCGCGCCACACTTGTTG-3';
 Rev4, 5'-AAAACCTGTTAGCCCCGATG-3'.

Primers to amplify DCV are provided in (19). Amplicons were pooled, purified with NucleoSpin kit (MACHEREY-NAGEL), sheared in a Bioruptor Pico (Diagenode) in 1.5 ml Bioruptor tubes (Diagenode) to an average size of 200 bp following 20 cycles of 30 s of sonication and 30 s of cooling. The library was prepared using NEB next ultra II DNA Library Prep kit (NEB) according to the manufacturer's instructions and multiplexed with barcodes (E7335, E7500, E7710 and E7730, NEB). Library size was determined using the DNA 1000 Bioanalyzer kit (Agilent), the concentration was measured with Qubit dsDNA high sensitivity (Thermo Fisher), and libraries were sequenced on an Illumina NextSeq200 as paired-end 55 bp long and demultiplexed with bcl2fastq (RRID:SCR_015058).

NGS data processing

Raw Illumina read data were processed using the bioinformatics pipeline V-pipe 3.0 (55), which was integrated into a custom Snakemake workflow. The CrPV and DCV parental stock datasets were aligned to their respective reference sequences (CrPV, NC_003924.1 and DCV EB, NC_001834.1), using the Burrows-Wheeler Alignment Tool BWA-MEM (56). Consensus sequences were generated for both parental stock datasets using SmallGenomeUtilities (57). These consensus sequences were used as references for aligning raw reads from the samples at passages 1, 2, 3, 5, and 10. Sequencing errors were corrected and mutations called relative to the parental stock consensus sequence using the tool VILOCA (version 1.1.0). VILOCA achieves this by clustering sequencing reads into local haplotypes approximately as long as the read length through a finite Dirichlet process mixture model, which integrates read quality scores to reliably distinguish true variants from sequencing noise (58). Non-synonymous mutations were annotated using an adapted version of vcf_annotator (<https://github.com/rpetit3/vcf-annotator>). Population nucleotide diversity was calculated using SNPGenie (59), which estimates average pairwise nucleotide differences directly from the observed

allele frequencies at each position after error correction. SNPGenie computes overall diversity and diversity per synonymous (piS) and non-synonymous (piN) sites, weighting each site by the fraction of possible substitutions that are synonymous or nonsynonymous. The minfreq parameter in SNPGenie was set to zero. The complete computational workflow and notebooks for generating figures are available on GitHub (<https://github.com/cbg-ethz/DCV-CrPV-cGAS-STING-pathway-data-analysis>).

Mean nucleotide diversity

A linear mixed-effects model was fitted to analyze the $\log_{10}$ -transformed DCV mean nucleotide diversity (lme, nlme package in R [60]). First, a model was constructed including passage, condition, and their interaction as fixed effects, along with evolutionary lineage replicate as a random effect to account for non-independence within replicates. To allow for heteroscedasticity between passages, a variance structure (VarIdent) was included, and since slight autocorrelation across passages within lineages was detected, an autoregressive moving average correlation structure (corARMA, $q = 1$) was incorporated. This model was then compared to a simpler model without the interaction term using ANOVA. Since ANOVA indicated no significant difference between the full and simplified model, the simpler model was used due to its fewer degrees of freedom. Residuals versus fitted values plots, as well as Levene's tests for homogeneity of variances across passage, genotype, and their interaction, did not indicate significant heteroscedasticity. Q-Q plots of the residuals showed approximate normality.

A similar mixed-effects model was fitted to analyze the $\log_{10}$ -transformed CrPV mean nucleotide diversity. To account for the piecewise linear structure of the data (initially increasing linearly, then decreasing), passage number was included as fixed effects along with an indicator variable for passage 10. This approach allowed modelling a potential change in the effect of passage after passage 5. An interaction term between condition and this indicator variable was also included to test for condition-specific changes after passage 10. Additionally, to account for within-lineage autocorrelation across passages, an AR(1) autocorrelation structure was incorporated in the model residuals using the corAR1 function from the nlme package. Model assumptions were confirmed using check_model from the performance package in R [61].

CrPV gene diversity was analyzed using ANOVA (R rstatix, anova_test [62]) with the dependent variable being the mean diversity over the passages

for each replicate line, the between-subject factor being condition, and the within-subject factor being gene. Normality was confirmed using qq-plots, homogeneity of variances was confirmed using Levene's test (R, `levene_test`), and homogeneity of covariances of the between-factors was tested using Box's M-Test (R, `box_m`). As post hoc analysis, pairwise comparisons between all gene levels were performed using estimated marginal means with Bonferroni adjustment for multiple testing (`emmeans` package in R [63]).

Nucleotide diversity in 2B and remaining coding region in DCV and CrPV populations

Nucleotide diversity in the 2B gene was compared to the remaining coding region by computing the mean diversity over the passages and then using a linear mixed-effects model (`lme4` package in R [64]), with the dependent variable being the log-transformed diversity for each evolutionary lineage. The fixed effects in the model were virus (DCV or CrPV), condition (included as a control variable), and genome region, while the evolutionary lineage was included as a random effect. A sensitivity analysis confirmed that removing condition from the model did not alter the significance or effect sizes of virus or genome region, supporting its exclusion. The normality of residuals was assessed using Q-Q plots and the Shapiro-Wilk test (`shapiro.test`). The homogeneity of variance was evaluated using plots of residuals versus fitted values and Levene's test (`car` package in R). The independence of residuals was checked through plots of residuals and autocorrelation function plots. Post hoc pairwise comparisons were performed using estimated marginal means with Tukey's adjustment for multiple comparisons between virus and genome region (`emmeans` package in R).

Statistical analysis for titers and relative RNA levels

A mixed ANOVA with within-subject factor passage and independent between-subject factor host condition was used to assess variation among log-transformed viral titers between condition and passages. Subsequently, a pairwise comparison of host conditions at each passage level was performed using estimated marginal means (R package `emmeans`, v1.10.7 [63]), adjusting P-values with Bonferroni correction. ANOVA assumptions were confirmed using the R-package *performance* (61) with the function `check_homogeneity` (Leven's test, $P > 0.05$), `check_sphericity` ($P > 0.05$), and `check_normality`, where the points remained within the confidence interval, suggesting that the residuals are approximately normally distributed. Residual diagnostics, including Durbin-Watson tests (DCV: $P = 0.998$, CrPV: $P = 0.07$)

and autocorrelation function (ACF) plot of residuals, indicated no significant autocorrelation, supporting the use of a model without an autocorrelation structure. Nevertheless, given the temporal structure of the data, a model incorporating an AR(1) autocorrelation structure was also fitted as a sensitivity analysis. This alternative model exhibited some collinearity and some deviations from normality in residuals. However, the significance of *P*-values remained unchanged, suggesting that the main model inference is robust to potential autocorrelation. This analysis was performed for each virus separately.

To compare the relative RNA levels between WT and D29N CrPV, a linear mixed-effects model was used on the log-transformed values (lme, nlme package in R [60]). A model was fitted that included host condition, time post infection and variant (D29N or WT) as additive effects. To account for measurements from the same lineage over time, a random intercept was included for time post-infection within each lineage. Additionally, to model the correlation of the measurements over time post-infection within each lineage, a first-order autoregressive (AR(1)) correlation structure was incorporated. Model assumptions were confirmed using `check_model` from the performance package in R (61), revealing only mild deviations from the model assumptions. Residual autocorrelation was assessed using the autocorrelation function plot (`acf`), confirming that AR(1) correlation structure successfully eliminated temporal dependencies

Approximation of selection coefficient of D29N mutation

Two complementary approaches were used to estimate the fitness effect of the D29N mutation: replication rate analysis and frequency data analysis. The replication rates (*r*) of the wild-type and mutant (D29N) were computed based on the fold change (*x*) representing RNA levels at 96 hours post-infection (hpi) relative to 0 hpi, assuming an exponential growth model:

$$x = e^{\frac{t}{10}r}$$

where *t* is time in hours and assuming an 10-hour generation time (65). Using the calculated replication rates, the selection coefficient (*s_{RNA}*) was determined by comparing the growth advantage of D29N relative to P0 over one generation (66):

$$s_{RNA} = \frac{r_{mutant} - r_{wildtype}}{r_{wildtype}}$$

The selection coefficient (S_{freq}) was also estimated using changes in D29N frequencies over time (67):

$$S_{freq} = \frac{\ln(\frac{f_{mutant}}{1-f_{mutant}})_{final} - \ln(\frac{f_{mutant}}{1-f_{mutant}})_{initial}}{n},$$

with n being the number of total generations. For this analysis, only evolutionary lineages with multiple time points in which D29N occurred were included, as single-time-point data do not allow for the determination of frequency change over time. Positive selection coefficients indicate that the mutant grows faster than the wild type, while negative coefficients indicate slower growth.

Protein structure prediction

AlphaFold 3 (27) was used to predict the CrPV 2B structure, which was simulated with the addition of oleic acid, myristic acid, palmitic acid and Ca^{2+} as ligands and ions.

2B sequence logo and homology detection

The CrPV 2B protein sequence was used in a homology search using BLASTp (68) against the non-redundant databases (NRDB) with an E value cutoff of 0.05, resulting in 131 homologs that were mainly from *Riboviria*. A multiple alignment with all the obtained proteins was created on the NCBI site with Cobalt (69) and the first fifty amino acids of the 2B protein in the alignment were used to construct a sequence logo (70). Homology of the 2B CrPV proteins with other proteins was established with PSI-BLAST (22) using default parameters ($E < 0.005$ for inclusion in the next iteration) against the NRDB database and iterating until convergence. To test the homology of CrPV 2B with poliovirus (Q1PHW2 in UniProt) and CVA10 (A0A6M2Z865 in UniProt) polyproteins HHPred was used in the pairwise alignment mode using default parameters.

Immunofluorescence assay

Drosophila S2 cells were seeded on 0.01% poly-D-Lysine coated coverslips in 24-well plates and incubated for 3 h at 27°C. Cells were then transfected with 0.5 µg of the indicated plasmids using 1 µL of X-tremeGENE HP (Roche), and the medium was refreshed after 24 h. For co-transfection of two plasmids, 0.25 µg of each plasmid were used instead. All following steps were performed at room temperature and in the dark when possible. At 48 h after transfection, cells were fixed in 4% paraformaldehyde for 15 min, permeabilized with

0.1% Triton-X100 in PBS for 15 min, and washed with PBS containing 0.1% Tween-20. Cells were blocked in 2% Normal Goat Serum (NGS) for 30 min and incubated for 45 min with the primary antibody diluted in 2% NGS in PBS. Calnexin 99A antibody (Developmental Studies Hybridoma Bank, Cnx99A6-2-1; RRID:AB_2722011) was diluted 1:10, GM130 antibody (Abcam, ab30637) was diluted 1:250, Rab7 antibody (Developmental Studies Hybridoma Bank, RRID:AB_2722471) was diluted 1:10, and V5 antibody (Invitrogen, 46-0705) was diluted 1:200. After washing, cells were stained with secondary antibody goat anti-rabbit conjugated with Alexa Fluor 594 (Thermo Fisher, A11012) and goat anti-mouse with Alexa Fluor 647 (Thermo Fisher, A21235), used at a 1:500 dilution. Nuclei were stained with 1 µg/mL DAPI (Sigma) for 10 min and coverslips were mounted on slides in ProLong Glass (Thermo Fisher). Cells were imaged on a Zeiss LSM900 Airyscan and 63x oil objective using the 405, 488, 561 and 633 excitation lasers. Images were Airyscan processed before analysis with Fiji software (71). Brightness and contrast were adjusted for visualization purposes. Raw images were used for colocalization analysis using the JACoP plug-in in Fiji (72). Images from individual cells were cropped and the Pearson's correlation coefficient between the eGFP and V5 signal was determined by applying the Costes automatic threshold. The coefficients were compared using a one-way ANOVA with Tukey's multiple comparison test. Shapiro-Wilk normality test confirmed normality of residuals and Levene's test confirmed equal variances.

Immunoprecipitation

The immunoprecipitation was performed following the instruction of the ChromoTek GFP-Trap Magnetic Agarose Kit (Proteintech). Briefly, S2R+ cells were co-transfected with plasmids encoding eGFP, 2B-eGFP or 2B(D29N)-eGFP in combination with a plasmid encoding Sting-V5. Cells were rinsed with PBS and lysed in ice-cold lysis buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % NP40, 1 mM PMSF). The cell lysates were centrifuged at 17,000 *g* for 15 min at 4°C. The supernatants were diluted with lysis buffer, equilibrated GFP-Trap beads were added, and samples were incubated with end-over-end rotation for 1 h at 4°C. The beads were then washed for four times with washing buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.05 % NP40, 1 mM PMSF) and resuspended in Laemmli sample buffer. The samples were heated for 5 min at 95°C, resolved on 12% SDS-PAGE gels, and transferred to the nitrocellulose membrane by wet transfer. Membranes were stained with antibodies against GFP (rat, 1:1000, ChromoTek, 3h9-150), V5 (mouse, 1:2000, Invitrogen, 46-0705), or

GAPDH (rabbit, 1:1000, Bioss, bs-8778R), followed by secondary antibodies IRDye 800CW goat anti-rat IgG (LICORBio, 926-32219), IRDye 800CW goat anti-mouse IgG (LICORBio, 926-32210), and IRDye 800CW goat anti-rabbit IgG (LICORBio, 926-32211), respectively. All secondary antibodies were used at a 1:2,000 dilution. The membranes were imaged using an Odyssey M Imaging System.

Data availability

All raw NGS data have been deposited in NCBI BioProject under the accession number PRJNA1257840. All other relevant data are within the manuscript and its Supporting Information files.

Acknowledgments

We thank members of the laboratory for discussion, Jean-Luc Imler (Université de Strasbourg) for kindly providing *Sting* knockout flies, and Eric Jan (University of British Columbia) for providing the cricket paralysis virus infectious clone. The CNX99A and Rab7 monoclonal antibodies developed by Sean Munro, MRC Laboratory of Molecular Biology, was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa.

Funding

This study was financially supported by the European Union's Horizon 2020 research and innovation program, under the Marie Skłodowska-Curie Actions Innovative Training Networks grant agreement no. 955974 (VIROINF) to RPeR and NB, and the China Scholarship Council program (Grant No. 202208320155) to HW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

1. Longdon B, Brockhurst MA, Russell CA, Welch JJ, Jiggins FM. The evolution and genetics of virus host shifts. *PLoS Pathog*. 2014 Nov 6;10(11):e1004395.
2. Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol (Amst)*. 2005 May;20(5):238–44.
3. Marchi J, Lässig M, Walczak AM, Mora T. Antigenic waves of virus-immune coevolution. *Proc Natl Acad Sci USA*. 2021 Jul 6;118(27):e2103398118.
4. Korobeinikov A. Immune response and within-host viral evolution: Immune response can accelerate evolution. *J Theor Biol*. 2018 Nov 7;456:74–83.
5. Merani S, Petrovic D, James I, Chopra A, Cooper D, Freitas E, et al. Effect of immune pressure on hepatitis C virus evolution: insights from a single-source outbreak. *Hepatology*. 2011 Feb;53(2):396–405.
6. Neuner-Jehle N, Zeeb M, Thorball CW, Fellay J, Metzner KJ, Frischknecht P, et al. Using viral diversity to identify HIV-1 variants under HLA-dependent selection in a systematic viral genome-wide screen. *PLoS Pathog*. 2024 Aug 8;20(8):e1012385.
7. Burmeister AR, Lenski RE, Meyer JR. Host coevolution alters the adaptive landscape of a virus. *Proc Biol Sci*. 2016 Sep 28;283(1839):20161528.
8. Seth RB, Sun L, Chen ZJ. Antiviral innate immunity pathways. *Cell Res*. 2006 Feb;16(2):141–7.
9. Cai H, Imler J-L. cGAS-STING : insight on the evolution of a primordial antiviral signaling cassette. *Fac Rev*. 2021 Jun 8;10:54.
10. Ablasser A, Chen ZJ. cGAS in action: Expanding roles in immunity and inflammation. *Science*. 2019 Mar 8;363:eaat8657.
11. Holleufer A, Winther KG, Gad HH, Ai X, Chen Y, Li L, et al. Two cGAS-like receptors induce antiviral immunity in *Drosophila*. *Nature*. 2021 Sep;597(7874):114–8.
12. Slavik KM, Morehouse BR, Ragucci AE, Zhou W, Ai X, Chen Y, et al. cGAS-like receptors sense RNA and control 3'2'-cGAMP signalling in *Drosophila*. *Nature*. 2021 Sep;597(7874):109–13.
13. Cai H, Li L, Slavik KM, Huang J, Yin T, Ai X, et al. The virus-induced cyclic dinucleotide 2'3'-c-di-GMP mediates STING-dependent antiviral immunity in *Drosophila*. *Immunity*. 2023 Sep 12;56(9):1991–2005.e9.
14. Goto A, Okado K, Martins N, Cai H, Barbier V, Lamiable O, et al. The Kinase IKK β Regulates a STING-and NF- κ B-Dependent Antiviral Response Pathway in *Drosophila*. *Immunity*. 2018 Aug 21;49(2):225–234.e4.
15. Cai H, Holleufer A, Simonsen B, Schneider J, Lemoine A, Gad HH, et al. 2'3'-cGAMP triggers a STING- and NF- κ B-dependent broad antiviral response in *Drosophila*. *Sci Signal*. 2020 Dec 1;13(660):eabc4537
16. Reinganum C, O'Loughlin GT, Hogan TW. A nonoccluded virus of the field crickets *Teleogryllus oceanicus* and *T. commodus* (Orthoptera: Gryllidae). *J Invertebr Pathol*. 1970 Sep;16(2):214–20.
17. Roberts JMK, Anderson DL, Durr PA. Metagenomic analysis of Varroa-free Australian honey bees (*Apis mellifera*) shows a diverse Picornavirales virome. *J Gen Virol*. 2018 Jun;99(6):818–26.
18. Casadevall A, Pirofski LA. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. *Infect Immun*. 1999 Aug;67(8):3703–13.

19. Lezcano OM, Fuhrmann L, Ramakrishnan G, Beerenwinkel N, Huynen MA, van Rij RP. Parallel evolution and enhanced virulence upon in vivo passage of an RNA virus in *Drosophila melanogaster*. *Virus Evol.* 2023 Dec 15;9(2):vead074.
20. Kerr CH, Wang QS, Keatings K, Khong A, Allan D, Yip CK, et al. The 5' untranslated region of a novel infectious molecular clone of the dicistrovirus cricket paralysis virus modulates infection. *J Virol.* 2015 Jun;89(11):5919–34.
21. Valles SM, Chen Y, Firth AE, Guérin DMA, Hashimoto Y, Herrero S, de Miranda JR, Ryabov E, ICTV Report Consortium. ICTV virus taxonomy profile: dicistroviridae. *J Gen Virol.* 2017 Mar 98:355–356.
22. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997 Sep 1;25(17):3389–402.
23. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005 Jul 1;33(Web Server issue):W244–8.
24. Käll L, Krogh A, Sonnhammer ELL. Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res.* 2007 Jul;35(Web Server issue):W429–32.
25. de Jong AS, Wessels E, Dijkman HBPM, Galama JMD, Melchers WJG, Willems PHGM, et al. Determinants for membrane association and permeabilization of the coxsackievirus 2B protein and the identification of the Golgi complex as the target organelle. *J Biol Chem.* 2003 Jan 10;278(2):1012–21.
26. Martínez-Gil L, Bañó-Polo M, Redondo N, Sánchez-Martínez S, Nieva JL, Carrasco L, et al. Membrane integration of poliovirus 2B viroporin. *J Virol.* 2011 Nov;85(21):11315–24.
27. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature.* 2024 Jun;630(8016):493–500.
28. McKnight KL, Swanson KV, Austgen K, Richards C, Mitchell JK, McGivern DR, Fritch E, Johnson J, Remlinger K, Magid-Slav M, Kapustina M, You S, Lemon SM. Stimulator of interferon genes (STING) is an essential proviral host factor for human rhinovirus species A and C. *Proc Natl Acad Sci USA.* 2020 Nov;117:27598–27607.
29. Longdon B, Day JP, Alves JM, Smith SCL, Houslay TM, McGonigle JE, et al. Host shifts result in parallel genetic changes when viruses evolve in closely related species. *PLoS Pathog.* 2018 Apr 12;14(4):e1006951.
30. Remold SK, Rambaut A, Turner PE. Evolutionary genomics of host adaptation in vesicular stomatitis virus. *Mol Biol Evol.* 2008 Jun 1;25(6):1138–47.
31. Sanjuán R, Domingo-Calap P. Genetic diversity and evolution of viral populations. *Encyclopedia of Virology.* 2021:53–61.
32. Wichman HA, Millstein J, Bull JJ. Adaptive molecular evolution for 13,000 phage generations: a possible arms race. *Genetics.* 2005 May;170(1):19–31.
33. Mattenberger F, Vila-Nistal M, Geller R. Increased RNA virus population diversity improves adaptability. *Sci Rep.* 2021 Mar 25;11(1):6824.
34. Riemersma KK, Jaeger AS, Crooks CM, Braun KM, Weger-Lucarelli J, Ebel GD, et al. Rapid evolution of enhanced Zika virus virulence during direct vertebrate transmission chains. *J Virol.* 2021 Mar 25;95(8):e02218–20.
35. Bons E, Leemann C, Metzner KJ, Regoes RR. Long-term experimental evolution of HIV-1 reveals effects of environment and mutational history. *PLoS Biol.* 2020 Dec 28;18(12):e3001010.

36. Kaczorowska J, Timmerman AL, Deijs M, Kinsella CM, Bakker M, van der Hoek L. Anellovirus evolution during long-term chronic infection. *Virus Evol.* 2023 Jan 5;9(1):vead001.
37. Wahl LM, Tanaka MM. Hazardous Loss of Genetic Diversity through Selective Sweeps in Asexual Populations. *Am Nat.* 2022 Mar;199(3):313–29.
38. Pennings PS, Kryazhimskiy S, Wakeley J. Loss and recovery of genetic diversity in adapting populations of HIV. *PLoS Genet.* 2014 Jan 23;10(1):e1004000.
39. Rouzine IM, Rodrigo A, Coffin JM. Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiol Mol Biol Rev.* 2001 Mar;65(1):151–85.
40. Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. *Nat Rev Micro.* 2004 Apr;2(4):279–88.
41. Jónás Á, Taus T, Kosiol C, Schlötterer C, Futschik A. Estimating the Effective Population Size from Temporal Allele Frequency Changes in Experimental Evolution. *Genetics.* 2016 Oct;204(2):723–35.
42. Nei M, Kumar S. *Molecular evolution and Phylogenetics* (New York, NY, 2000; online edn, Oxford Academic, 31 Oct. 2023).
43. Nelson CW, Hughes AL. Within-host nucleotide diversity of virus populations: insights from next-generation sequencing. *Infect Genet Evol.* 2015 Mar;30:1–7.
44. Li Z, Zou Z, Jiang Z, Huang X, Liu Q. Biological function and application of picornaviral 2B protein: A new target for antiviral drug development. *Viruses.* 2019 Jun 4;11(6):510.
45. Nieva JL, Madan V, Carrasco L. Viroporins: structure and biological functions. *Nat Rev Micro.* 2012 Jul 2;10(8):563–74.
46. Martínez-Gil L, Mingarro I. Viroporins, Examples of the Two-Stage Membrane Protein Folding Model. *Viruses.* 2015 Jun 26;7(7):3462–82.
47. Cuconati A, Xiang W, Lahser F, Pfister T, Wimmer E. A protein linkage map of the P2 nonstructural proteins of poliovirus. *J Virol.* 1998 Feb;72(2):1297–307.
48. Agirre A, Barco A, Carrasco L, Nieva JL. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *J Biol Chem.* 2002 Oct 25;277(43):40434–41.
49. de Jong AS, Schrama IJW, Willems PHGM, Galama JMD, Melchers WJG, van Kuppeveld FJM. Multimerization reactions of coxsackievirus proteins 2B, 2C and 2BC: a mammalian two-hybrid analysis. *J Gen Virol.* 2002 Apr;83(Pt 4):783–93.
50. Patargias G, Barke T, Watts A, Fischer WB. Model generation of viral channel forming 2B protein bundles from polio and coxsackie viruses. *Mol Membr Biol.* 2009 Aug 21;26(5):309–20.
51. Ao D, Sun S-Q, Guo H-C. Topology and biological function of enterovirus non-structural protein 2B as a member of the viroporin family. *Vet Res.* 2014 Aug 28;45:87.
52. Baggen J, Thibaut HJ, Strating JRPM, van Kuppeveld FJM. The life cycle of non-polio enteroviruses and how to target it. *Nat Rev Micro.* 2018 Jun;16(6):368–81.
53. Triantafilou M, Ramanjulu J, Booty LM, Jimenez-Duran G, Keles H, Saunders K, et al. Human rhinovirus promotes STING trafficking to replication organelles to promote viral replication. *Nat Commun.* 2022 Mar 17;13(1):1406.
54. Merklings SH, van Rij RP. Analysis of resistance and tolerance to virus infection in *Drosophila*. *Nat Protoc.* 2015 Jul;10(7):1084–97.

55. Fuhrmann L, Jablonski KP, Topolsky I, Batavia AA, Borgsmüller N, Baykal PI, et al. V-pipe 3.0: a sustainable pipeline for within-sample viral genetic diversity estimation. *Gigascience*. 2024 Jan 2;13:giae065.
56. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009 Jul 15;25(14):1754–60.
57. Posada-Céspedes S, Seifert D, Topolsky I, Jablonski KP, Metzner KJ, Beerenwinkel N. V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data. *Bioinformatics*. 2021 Jul 19;37(12):1673–1680.
58. Fuhrmann L, Langer B, Topolsky I, Beerenwinkel N. VILOCA: sequencing quality-aware viral haplotype reconstruction and mutation calling for short-read and long-read data. *NAR Genom Bioinform*. 2024 Dec;6(4):lqae152.
59. Nelson CW, Moncla LH, Hughes AL. SNPGenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. *Bioinformatics*. 2015 Nov 15;31(22):3709–11.
60. Kuznetsova A, Brockhoff PB, Christensen RHB. Package “lmerTest.” R package version. 2015 2(0):734.
61. Lüdtke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *JOSS*. 2021 Apr 21;6(60):3139.
62. Kassambara A. Pipe-Friendly framework for basic statistical tests’. R package version 0.7.0. 2020.
63. Searle SR, Speed FM, Milliken GA. Population marginal means in the linear model: an alternative to least squares means. *Am Stat*. 1980 Nov;34(4):216–21.
64. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1–48.
65. Kerr CH, Dalwadi U, Scott NE, Yip CK, Foster LJ, Jan E. Transmission of Cricket paralysis virus via exosome-like vesicles during infection of *Drosophila* cells. *Sci Rep*. 2018 Nov 26;8(1):17353.
66. Marée AF, Keulen W, Boucher CA, De Boer RJ. Estimating relative fitness in viral competition experiments. *J Virol*. 2000 Dec;74(23):11067–72.
67. Chevin LM. On measuring selection in experimental evolution. *Biol Lett*. 2011 Apr 23;7(2):210–3.
68. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009 Dec 15;10:421.
69. Papadopoulos JS, Agarwala R. COBAL: constraint-based alignment tool for multiple protein sequences. *Bioinformatics*. 2007 May 1;23(9):1073–9.
70. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*. 2004 Jun;14(6):1188–90.
71. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012 Jun 28;9(7):676–82.
72. Bolte S, Cordelières FP. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc*. 2006 Dec;224(Pt 3):213–32.

Supplementary figures

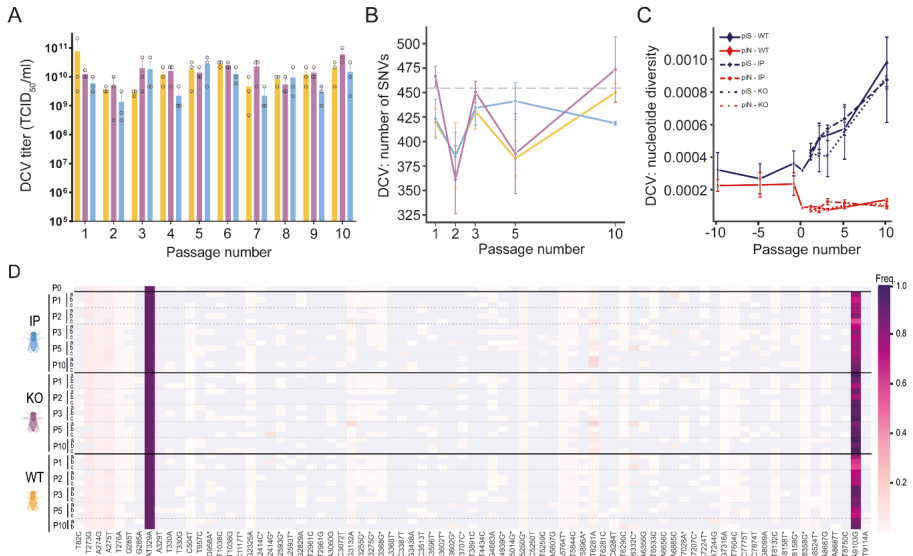


Figure S1. No evidence for adaptive evolution of DCV in its natural host, *Drosophila melanogaster*, with differing cGAS-STING response. **(A)** DCV titers in lysates of flies after each passage in the indicated host background. Color coding corresponds to Fig 1; yellow, WT; purple, *Sting* KO; blue, immune primed (IP). Bars represent the means and SD of the three replicate viral lineages; circles indicate individual data points. **(B)** The number of single-nucleotide variants (SNVs) in DCV populations from WT, *Sting* KO, and IP flies across passages. Data are shown as means and SEM across the three replicate lineages for each condition. The dashed line marks the number of SNVs in the parental stock. **(C)** Nucleotide diversity per synonymous (piS) and non-synonymous (piN) site of DCV populations across passages. Diversity values from DCV evolved in WT flies from our previous study (19) are integrated and denoted as passages -10, -5, and -1. Passage 0 is the parental stock, and passages 1 to 10 are the evolved DCV populations of the current study. Data are shown as means and standard errors across the three replicate lineages for each genotype. **(D)** Heatmap of SNV frequencies of all SNVs for DCV populations from WT, *Sting* KO and IP flies. Non-variable sites are indicated in gray. The x-axis represents genome positions relative to the DCV EB reference strain (NC_001834.1). Only SNVs that occurred in at least one sample with an observed frequency ≥ 0.1 are shown. Non-synonymous mutations are marked with an asterisk.

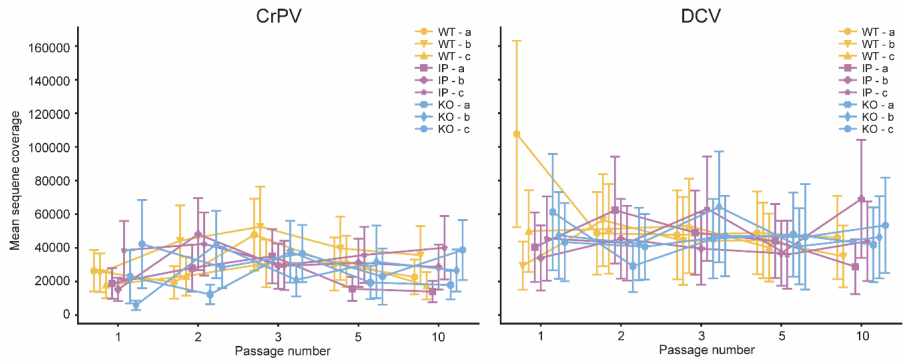


Figure S2. Sequence coverage across genome positions for CrPV and DCV samples. For each sample, the plotted value represents the average sequencing coverage across all positions in the viral genome. Error bars denote the standard deviation of coverage across these positions, illustrating the variability of read depth within each genome. Lowercase letters (a, b, c) indicate the three independent evolutionary lineages.

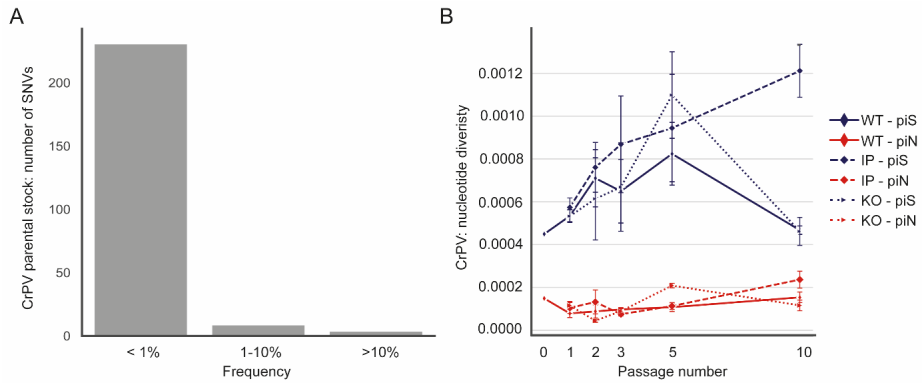


Figure S3. Single-nucleotide variants (SNVs) in the CrPV parental stock and nucleotide diversity in the evolutionary lineages. **(A)** Frequency distribution of SNVs in the indicated bins for the parental CrPV stock. **(B)** CrPV population nucleotide diversity per synonymous (piS) and non-synonymous (piN) site across passages. Data are shown as means and standard error across the three replicate lineages for each condition.

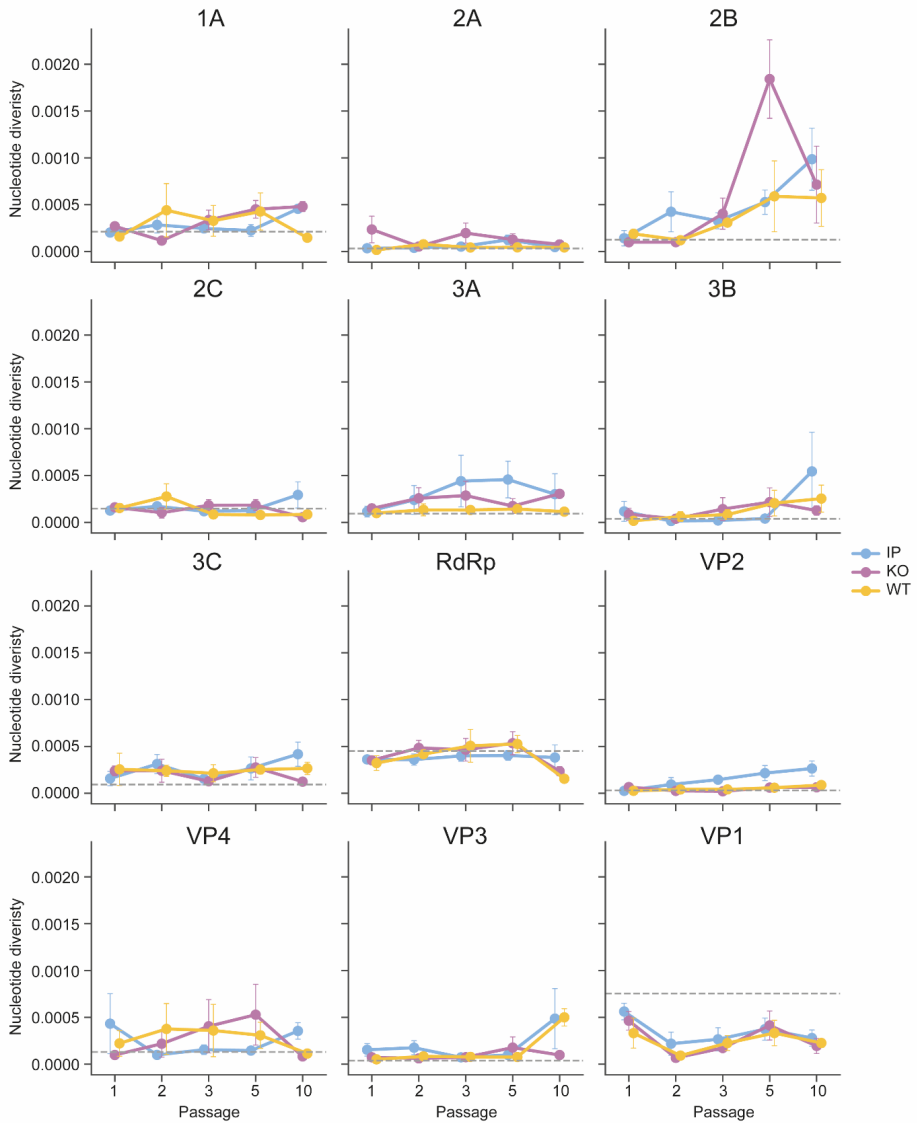


Figure S4. Temporal dynamics of mean nucleotide diversity per site of each CrPV gene. Data are shown as means and standard errors across the three replicate lineages for each host background.



C

Q ss_pred		H	
Q Q_4254442	123	D	123 (138)
Q Consensus	123	e	123 (138)
		.	
T Consensus	1096	~	1096 (2193)
T T_4254442	1096	K	1096 (2193)
T ss_pred		H	

< Figure S5. Structural predictions and homology of the CrPV 2B protein. **(A)** Transmembrane region prediction plot for CrPV 2B, created with PolyPhobius (1) using default parameters. **(B, C)** Alignment of the CrPV 2B protein (top, "Q", uniprot id: Q9IJX4) with the complete polyprotein of **(B)** poliovirus (bottom, "T", uniprot id: Q1PHW2) and **(C)** the coxsackievirus A10 (CVA10) (bottom, "T", uniprot id: A0A6M2Z865). In order to find support for the homology of CrPV 2B with enterovirus 2B, we ran HHpred in the pairwise alignment mode with default parameters. Among the polyproteins of poliovirus and CVA10 polyproteins, the 2B proteins (position 491–589 in poliovirus and position 1013–1111 in CVA10) are the most likely homologous region with CrPV 2B, even though the E-value is not significant. The top and bottom red letters indicate the predicted secondary structure per sequence (H, alpha helix; C, coil). Sequences of the viruses are indicated in multicolor font, and the consensus sequences are shown with symbols indicating moderately conserved ("+") and strongly conserved ("I") positions. The lack of a consensus sequence for the polio and coxsackievirus 2B proteins is caused by the low level of conservation of the 2B protein in both enteroviruses relative to the other enterovirus proteins.

Supplementary tables

Table S1. Summary of fixed and random effects from the linear mixed model predicting the log-transformed DCV nucleotide diversity based on host background (condition) and passage number as fixed effects and replicate lineage as a random effect.

Effect type	Effect	Estimate	Std. Error	t-value	p-value
Fixed Effects	Intercept	-3.6350	0.02096	-173.3920	< 0.0001
	Passage	0.0156	0.0029	5.2617	< 0.0001
	Condition IP	-0.0026	0.0276	-0.0963	0.9264
	Condition KO	-0.0563	0.0276	-2.0377	0.0877
		Std. Dev.			
Random Effect	Replicate (Intercept)	0.0201			
	Residual	0.0308			

Table S2. Summary of fixed and random effects from the linear mixed model predicting the log-transformed CrPV nucleotide diversity based on host background (condition), passage number, passage 10 as a separate binary factor, and an interaction term with fly condition and passage 10 as fixed effects and replicate as a random effect.

Effect type	Effect	Estimate	Std. Error	t-value	p-value
Fixed Effects	Intercept	-3.77	0.048	-78.180	< 0.001
	ConditionIP	-0.003	0.05	-0.06	0.9544
	ConditionKO	-0.001	0.05	-0.023	0.982
	Passage	0.046	0.012	3.920	< 0.001
	Passage10 (binary indicator)	-0.391	0.107	-3.645	< 0.001
	Condition IP: Passage10	0.245	0.095	2.592	0.014
	Condition KO: Passage10	-0.028	0.095	-0.294	0.771
		Std. Dev			
Random Effect	Replicate (Intercept)	0.027			
	Residual	0.102			

Table S3. Results of post hoc pairwise comparisons of mean diversity between viral genes across passages. Pairwise contrasts were performed using estimated marginal means with Bonferroni adjustment for multiple comparisons. Only significant contrasts are shown.

	Estimate	SE	df	t-ratio	p-value
2A - 2B	-4.085e-4	3.989e-05	22	-10.244	<.0001
2B - VP2	4.07e-4	3.989e-05	22	10.203	<.0001
2B - 3B	3.580e-4	3.989e-05	22	8.977	<.0001
2B - 3B	3.580e-4	3.989e-05	22	8.977	<.0001
2B - 2C	3.425e-4	3.989e-05	22	8.586	<.0001
2B - VP3	3.389e-4	3.989e-05	22	8.497	<.0001
2A - RdRp	-3.128e-4	3.989e-05	22	-7.842	<.0001
RdRp - VP2	3.111e-4	3.989e-05	22	7.802	<.0001
2B - 3A	2.662e-4	3.989e-05	22	6.673	0.0001
3B - RdRp	-2.623e-4	3.989e-05	22	-6.575	0.0001
2B - 3C	2.532e-4	3.989e-05	22	6.347	0.0001
2C - RdRp	-2.467e-4	3.989e-05	22	-6.185	0.0002
RdRp - VP3	2.431e-4	3.989e-05	22	6.096	0.0002
2B - VP4	2.298e-4	3.989e-05	22	5.761	0.0004
1A - 2A	2.238e-4	3.989e-05	22	5.612	0.0006
1A - VP2	2.222e-4	3.989e-05	22	5.571	0.0006
2B - VP1	2.084e-4	3.989e-05	22	5.224	0.0014
2A - VP1	-2.002e-4	3.989e-05	22	-5.020	0.0023
VP2 - VP1	-1.986e-4	3.989e-05	22	-4.979	0.0025
1A - 2B	-1.848e-4	3.989e-05	22	-4.632	0.0056
2A - VP4	-1.788e-4	3.989e-05	22	-4.483	0.0078
VP2 - VP4	-1.772e-4	3.989e-05	22	-4.443	0.0086
1A - 3B	1.733e-4	3.989e-05	22	4.344	0.0107
3A - RdRp	-1.704e-4	3.989e-05	22	-4.271	0.0126
1A - 2C	1.577e-4	3.989e-05	22	3.954	0.0254
3C - RdRp	-1.574e-4	3.989e-05	22	-3.945	0.0259
2A - 3C	-1.554e-4	3.989e-05	22	-3.897	0.0288
1A - VP3	1.542e-4	3.989e-05	22	3.865	0.0308
3C - VP2	1.538e-4	3.989e-05	22	3.856	0.0314
3B - VP1	-1.497e-4	3.989e-05	22	-3.753	0.0392

Table S4. Summary of the linear mixed model predicting the log-transformed CrPV viral loads of CrPV wildtype and D29N mutant.

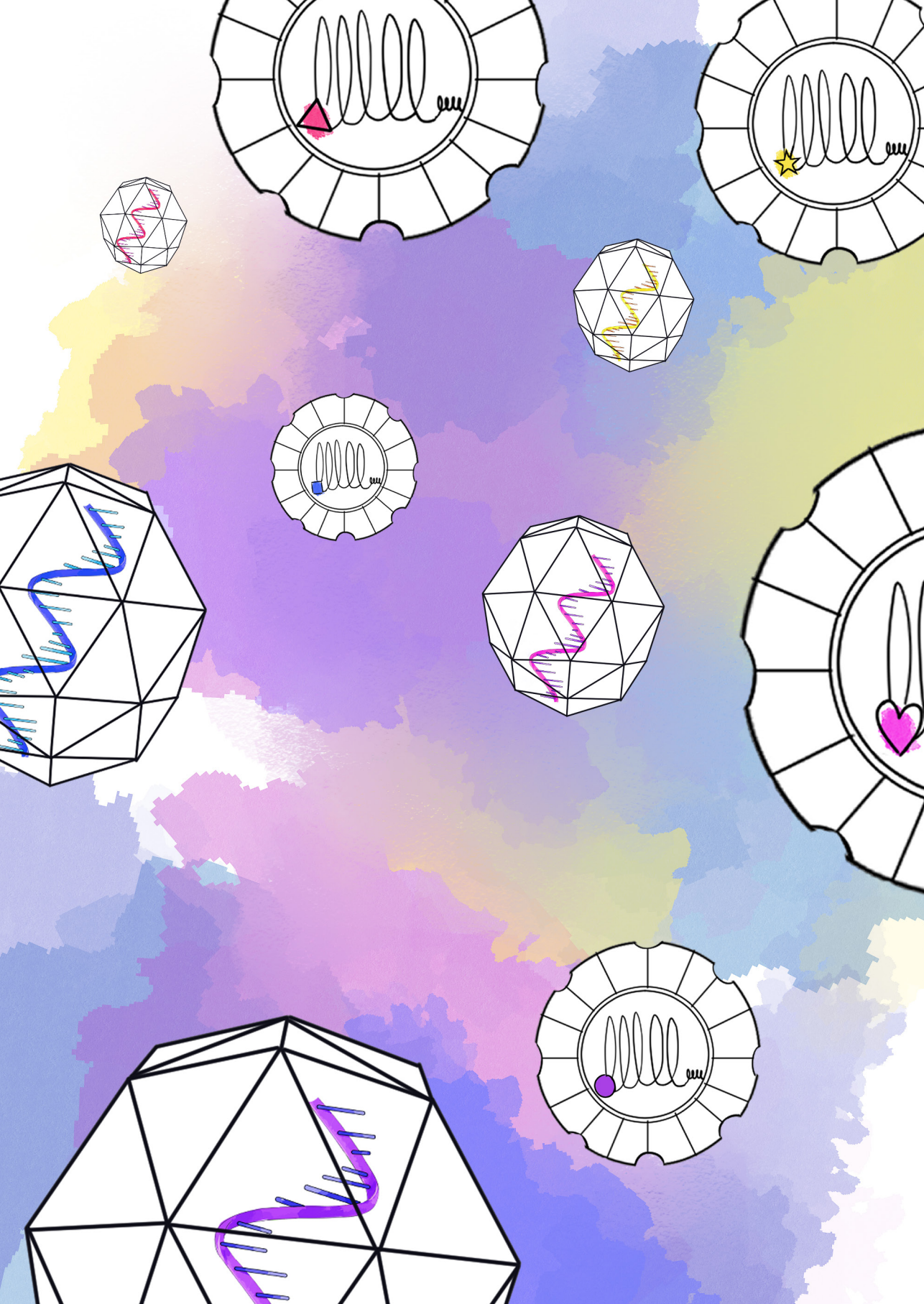
Effect	Estimate	Std. Error	t-value	p-value
Intercept	-8.768	0.457	-19.175	< 0.0001
Time post infection	0.14	0.006	22.925	< 0.0001
Variant D29N	2.632	0.382	6.886	< 0.0001
Condition IP	-1.856	0.468	-3.964	0.0014
Condition KO	0.000	0.468	0.000	0.999

Table S5. Estimates of selection coefficients based on mutation frequency changes (s_{freq}) and CrPV RNA levels (s_{RNA}).

Estimation based on mutation frequency changes		
Evolved lineage	s_{freq}	mean
KOa	0.249	0.23
WTa	0.180	
WTc	0.26	
Estimation based on CrPV RNA levels		
Sample	s_{RNA}	mean
IP-1	0.161	0.328
IP-2	0.568	
IP-3	0.256	
KO-1	0.506	0.375
KO-2	0.296	
KO-3	0.322	
WT-1	0.23	0.097
WT-2	0.107	
WT-3	-0.047	

Supplementary references

1. Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* 2004 May 14;338(5):1027–36.



Chapter 4

Exploring the role of RNA-modifying enzymes in insect RNA virus infection

Oscar M. Lezcano¹, Haitao Wang¹, María Paz Menafrá², Emina Fodolović², Erik Storkebaum², Martijn A. Huynen³, Ronald P. van Rij¹

¹ Department of Medical Microbiology, Radboud University Medical Center, Nijmegen, 6500 HB, Nijmegen, the Netherlands

² Molecular Neurobiology Laboratory, Donders Center for Neuroscience, Donders Institute for Brain, Cognition and Behaviour and Faculty of Science, Radboud University, 6525 EN, Nijmegen, the Netherlands

³ Department of Medical BioSciences, Radboud University Medical Center, 6500 HB, Nijmegen, the Netherlands

Corresponding author: ronald.vanrij@radboudumc.nl

Abstract

Covalent chemical marks added to cellular RNA by RNA-modifying enzymes help regulate the antiviral immune response. It has been suggested that RNA viruses have adapted to utilize these enzymes to maximize their gene expression, modulate host responses, and evade immunity, sometimes by modifying their own genome. However, results from different modification mapping techniques are conflicting and require orthogonal validation. Therefore, we aimed to functionally characterize the role of RNA-modifying enzymes in virus replication using a genetic knockdown approach. We used transgenic *Drosophila* RNA interference (RNAi) lines to knockdown 18 genes involved in RNA editing and recognition, using two RNAi lines as biological replicates when possible. The constitutive knockdown of 8 of those genes resulted in developmental lethality, therefore a conditional knockdown with the temperature inducible *Gal80^{ts}; Gal4* system was used instead. Knockdown flies were infected with *Drosophila* C virus (DCV) and viral RNA levels were analyzed. Despite the lack of detectable reduction in mRNA levels for most of the target genes, knockdown of the RNA methyltransferase *Trmt61* and the pseudouridine synthase *Nop60B* reduced DCV RNA levels in two independent RNAi lines. Additionally, the depletion of *Trmt61* in persistently infected *Drosophila* cells reduced viral RNA levels of DCV, *Drosophila* B birnavirus, Flock House virus, and *Drosophila* totivirus. Given the technical limitations of our approach, we are cautious about drawing strong conclusions and therefore present the two genes as potential proviral factors that require further confirmation.

Introduction

RNA from living organisms can undergo covalent modification with more than 170 chemical marks (1) in a dynamic process that affects gene expression, providing a post-transcriptional mechanism of gene regulation (2, 3). RNA modifications were discovered more than 60 years ago (4-8), and soon after, it was shown that viral RNA genomes were also subject to modification (9-11). However, the inability to precisely map RNA modifications and the lack of knowledge about the cellular machinery involved hampered understanding of their functions. Over the past 15 years, advancements in high-throughput methods for mapping modifications at single-nucleotide resolution have reignited interest in this field (12), now termed epitranscriptomics (13).

Epitranscriptomic changes profoundly impact RNA metabolism, affecting RNA stability (14), structure (15, 16), splicing (17), nuclear export (18), and translation efficiency (19), among others. Most RNA modifications occur in non-coding RNA (1), especially in transfer (tRNA) and ribosomal RNA (rRNA), but modifications also occur in protein-coding messenger RNA (mRNA) (20). The most abundant and well-studied modified RNA nucleotides are *N*⁶-methyladenosine (m⁶A), pseudouridine (Ψ), 5-methylcytidine (m⁵C) and *N*¹-methyladenosine (m¹A) (21). Typically, the modifications are reversibly deposited on RNA by “writer” enzymes, removed by “eraser” enzymes, and recognized by “reader” enzymes, although Ψ is an exception as it is thought to be an irreversible mark (22).

Diverse viruses have been reported to be RNA modified (23-26), for instance the genome of the human immunodeficiency virus 1 (HIV-1) is extensively modified by marks such as m⁶A (27-29), m⁵C (30, 31), pseudouridine, inosine (27), 2'-O-methylation (32) and N4-acetylcytidine (ac4C) (33). Although the functional relevance of some of these modifications is unknown, HIV-1 translation is enhanced by the NSUN2-mediated m⁵C modification of viral RNA (31). Moreover, m⁵C modifications of HIV-1 RNA are involved in promoting viral latency and hence important for establishing virus reservoirs (30). Additionally, several viruses from the *Flaviviridae* family and chikungunya virus (*Togaviridae* family) have been reported to contain m⁶A modifications (34-36). However, these findings were based on antibody-dependent techniques that have high rates of false positives (37), and orthogonal validation by nanopore direct RNA sequencing and a single-base elongation- and ligation-based qPCR amplification method (38) failed to detect m⁶A in dengue virus or chikungunya

virus RNA (39). As many studies used antibody-based methods to profile viral RNA modifications, further functional validation of epitranscriptomic control of viral infection is required.

To study the role of RNA modifications in virus infection, we performed an RNA interference (RNAi) screen using *Drosophila melanogaster* as a model. We selected modifying enzymes of chemical marks suggested to be involved in immunity (40, 41), knocked them down using transgenic RNAi, infected flies with *Drosophila C virus* (DCV), and quantified the amount of viral RNA. For *in vivo* knockdown, we used *Drosophila* RNAi lines from the Vienna *Drosophila* Research Center (VDRC), which carry an inverted repeat gene fragment preceded by the Upstream Activation Sequence (UAS) promoter (42). Crossing the RNAi lines with a driver line expressing the Gal4 transcription factor leads to the specific expression of the long double-stranded RNA hairpin (hpRNA), which is processed into small interfering RNAs to mediate specific gene silencing. Leveraging this genetic tool, we found two enzymes that might act as proviral factors, as their knockdown reduced viral RNA levels.

Results and discussion

Exploring the effect of RNA-modifying enzymes on *Drosophila C virus* replication

To study the function of RNA modifications in viral infection, we selected transgenic fly RNAi lines to knockdown enzymes responsible for producing m⁶A, m¹A, m⁵C, pseudouridine, ac⁴C, 3-methylcytidine (m³C), inosine, 2'-O-methylation, 5-methoxycarbonylmethyl-2-thiouridine (mcm⁵s²U) and dihydrouridine, modifications that have been linked with immunity (40, 41). We focused on writer enzymes as the readers and erasers are unknown for many RNA modifications. However, for the extensively studied m⁶A and m⁵C modifications, we included the reader proteins Ythdf and Ref, respectively. If available, we used flies from two independent RNAi libraries, called GD and KK, as biological replicates. Both encode long hpRNA, but transgenes are randomly integrated in the GD library, whereas transgenes are inserted in the same genomic position in the KK library, reducing the risk of disrupting other genes during transgenesis.

We crossed all RNAi lines (UAS-hpRNA) with a driver constitutively expressing *Gal4* under control of the *tubulin* promoter (Tub-Gal4), selected the offspring

expressing the hpRNA (UAS-hpRNA/Gal4), inoculated them with DCV, and analyzed viral RNA levels with respect to *w*¹¹¹⁸ or KK control flies (Figure 1, see M&M for the KK control stock description). Unexpectedly, knockdown efficiency was limited: at least 50% knockdown was observed in only 4 out of 10 RNAi lines, 60-70% in 2 lines, and no detectable knockdown in 4 lines (Figure S1A). In contrast, eight crosses produced no offspring expressing both the Gal4 driver and the UAS-hpRNA transgene, likely a sign of knockdown-induced lethality during early developmental stages (Figure 1). The lethality upon constitutive gene silencing suggests a functional knockdown.

Seeking to improve the knockdown, we tried enhancing the RNAi pathway activity by crossing RNAi lines with a driver expressing *Gal4* under control of the *actin* promoter in combination with a UAS-*Dicer-2* transgene to overexpress *Dicer-2* (UAS-*Dicer-2*; Act-Gal4). Unfortunately, knockdown efficiency was even lower than with the Tub-Gal4 driver (24% knockdown in only 1 out of 11 lines tested, data not shown), suggesting that overexpression of *Dicer-2* does not increase knockdown efficiency. However, it is difficult to directly compare the results from these experiments as expression of the hpRNA was controlled by different promoters (*Tubulin* and *Actin*).

The lethal phenotype prevented the screen of many of the selected RNAi lines; hence, we switched to a conditional knockdown system using Gal80^{ts}, a thermosensitive inhibitor of Gal4 that is active at 18°C but heat-inactivated at 29°C (Tub-Gal80^{ts}; Tub-Gal4). This system allowed us to rear flies at a low temperature and induce the knockdown only during adulthood by transferring flies to 29°C. Unfortunately, crosses with the Tub-Gal80^{ts}; Tub-Gal4 driver produced also poor gene silencing: 25% knockdown was found in 2 out of 14 lines tested, whereas no significant knockdown was found in the other lines (Figure S1B).

The inefficient silencing of our target genes contrasts with the extensive use of these RNAi libraries (e.g., [42-46]). However, we have solely measured knockdown efficiency at the mRNA level, and it has been reported that mild mRNA reductions can translate into high knockdown efficiency at the protein level (47). It will therefore be necessary to analyze protein levels to draw conclusive results about the knockdown efficiency of our target genes.

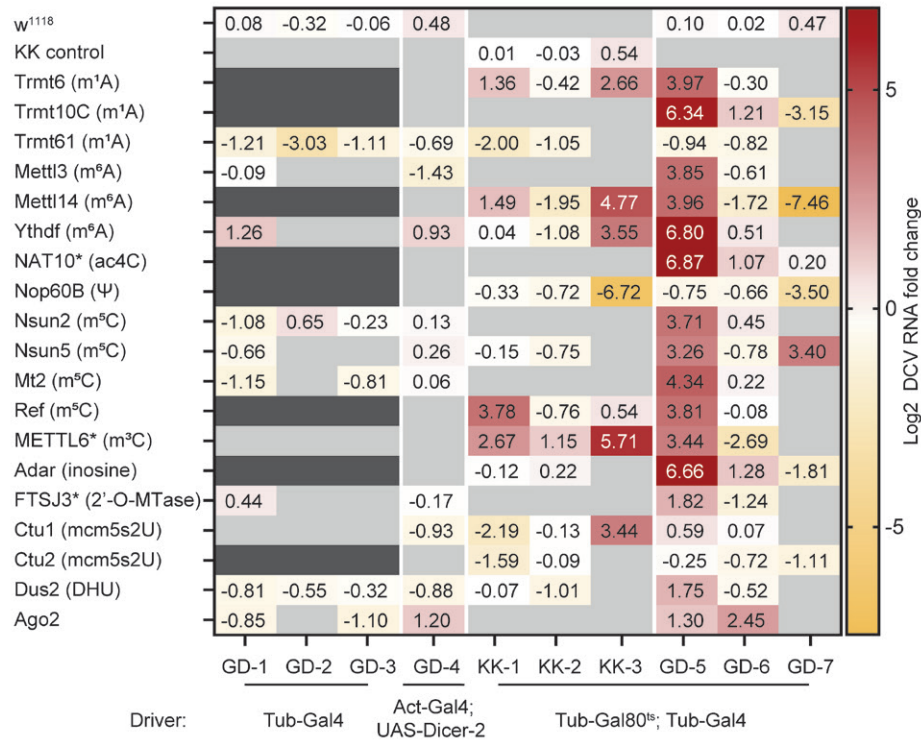


Figure 1. RNAi knockdown screen of RNA-modifying enzymes upon *Drosophila C* virus infection. Male RNAi flies (UAS-hpRNA) were crossed with female virgins of the indicated Gal4 driver lines. Two to five-day-old F1 females expressing both the hpRNA and Gal4 transgenes (UAS-hpRNA/Gal4) were selected. UAS-hpRNA/Gal4 flies from crosses with Tub-Gal4 and UAS-Dicer-2; Act-Gal4 flies were directly infected, while those crossed with Tub-Gal80^{ts}; Tub-Gal4 flies were transferred from 18°C to 29°C for 4 days to induce target gene knockdown before infection, and were kept at 29°C until harvesting. The UAS-hpRNA/Gal4 flies were injected with 100 TCID₅₀ of DCV and three pools of five flies were harvested at 24 hours post-infection for viral RNA quantification by RT-qPCR. The presented data are the median DCV RNA levels of the three pools expressed as Log2 fold change compared to control flies (w¹¹¹⁸ for GD lines and KK control for KK lines), using the housekeeping gene *Rp49* for normalization. Lethal phenotype of the offspring are represented with dark gray shading. Light gray cells indicate conditions that were not taken along in a given experiment. Gene names follow FlyBase annotation, except for CG1994, CG34195 and CG8939 (*) for which no name has yet been assigned and the name of the human ortholog was used instead. GD and KK represent the two RNAi libraries used; suffixes indicate the experiment number.

Despite the inefficient silencing, we proceeded with assessing viral RNA levels in the knockdown flies. Unfortunately, the results of the screen were not very consistent. Not only were the differences in viral RNA levels between the UAS-hpRNA/Gal4 and control flies variable between experiments, but also

the phenotype changed between replicates for some genes (Figure 1). For instance, knockdown of *Argonaute-2*, encoding a core protein of the antiviral RNAi pathway and used as a positive antiviral control, resulted in increased viral RNA levels in three of the replicate experiments, but reduced viral RNA levels in two other replicates. Despite the discrepancies, the knockdown of the methyltransferase *Trmt61* (tRNA (adenine(58)-N1)-methyltransferase activity) and the pseudouridine synthase *Nop60B* (H/ACA ribonucleoprotein complex catalytic subunit) reduced DCV RNA levels in all experimental replicates (Figure 1, $n = 8$ for *Trmt61* and $n = 6$ for *Nop60B*). Thus, we decided to focus on these genes for follow-up experiments using the conditional knockdown system.

Reduced DCV replication upon *Trmt61* knockdown

To confirm the results from our screen, we determined the viral titers of *Trmt61* knockdown flies (GD and KK lines) upon infection with DCV or Cricket paralysis virus (CrPV), a dicistrovirus closely related to DCV. We evaluated the *Trmt61* knockdown efficiency in the RNAi lines (UAS-hpTrmt61) and found 40% knockdown in KK line but no knockdown in the GD line (Figure 2A). Accordingly, viral titers in the KK line were reduced approximately 40-fold on average for DCV, albeit not reaching statistical significance, and 20-fold for CrPV, but showed no difference between *Trmt61* knockdown and control for the GD line (Figure 2B and C, unpaired *t*-test). Next, we analyzed DCV growth kinetics in *Trmt61* KK flies and found that the total viral RNA accumulation over the time course was approximately 1.6-fold lower than in the control flies (estimated from the area under the curve, see Materials and Methods), showing a significant difference in DCV replication kinetics ($P < 0.001$, two-way ANOVA, Figure 2D).

To validate these results, we knocked down *Trmt61* in *Drosophila* S2-R+ cells, which reduced its expression by 80% knockdown compared to the control knockdown (Figure S2). These cells are persistently infected with *Drosophila* A virus (DAV), *Drosophila* B birnavirus (DBV), DCV, *Drosophila* X virus (DXV), Flock House virus (FHV) and *Drosophila* totivirus (DTV). Interestingly, we found that RNA levels of all six viruses were reduced upon *Trmt61* knockdown (14 to 36% relative to control knockdown; average reduction of 24.7%), although the difference was statistically significant only for DBV, DCV, FHV and DTV (Figure 3). The knockdown did not visibly affect cell viability. Together with the *in vivo* data, these results suggest that *Trmt61* is a proviral factor that is required for efficient RNA virus replication.

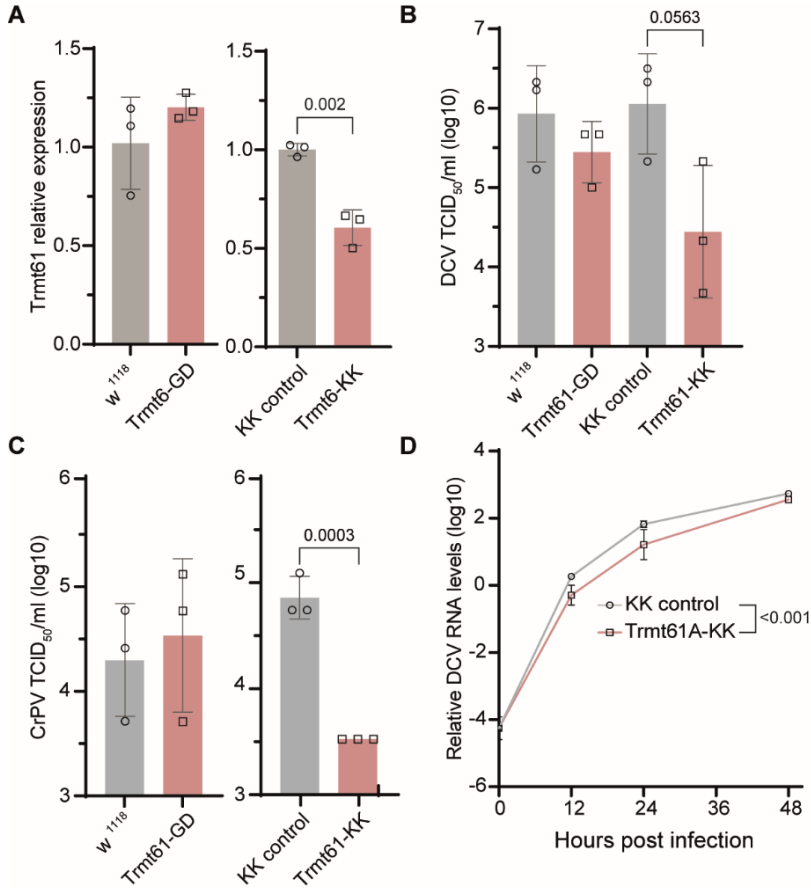


Figure 2. *Trmt61* knockdown reduces DCV and CrPV replication *in vivo*. UAS-hpTrmt61 males from the GD and the KK lines were crossed with Tub-Gal80^{ts}; Tub-Gal4 female virgins. UAS-hpTrmt61/Tub-Gal80^{ts}; Tub-Gal4 were incubated at 29°C for four days and then inoculated with either 100 TCID₅₀ of DCV or 1,000 TCID₅₀ of CrPV. **(A)** *Trmt61* knockdown efficiency relative to control flies, **(B)** DCV titers, and **(C)** CrPV titers were determined at 24 hpi from three pools of five flies. The differences were analyzed by unpaired *t*-tests and titers were determined by end-point dilution assay. Data in (A) and (C) are presented in two separate plots to indicate that the experiments were performed on different days. **(D)** DCV growth curve in *Trmt61* knockdown (KK line) and control flies. Flies were crossed as indicated above, inoculated with 100 TCID₅₀ of DCV, and three pools of five flies were collected at the indicated time points. DCV RNA levels were determined by RT-qPCR and expressed relative to *Rp49* housekeeping gene. Differences in DCV RNA levels between conditions were analyzed over time with a two-way ANOVA, and the interaction *P*-value is reported. Data are shown as mean and SD across the three replicates.

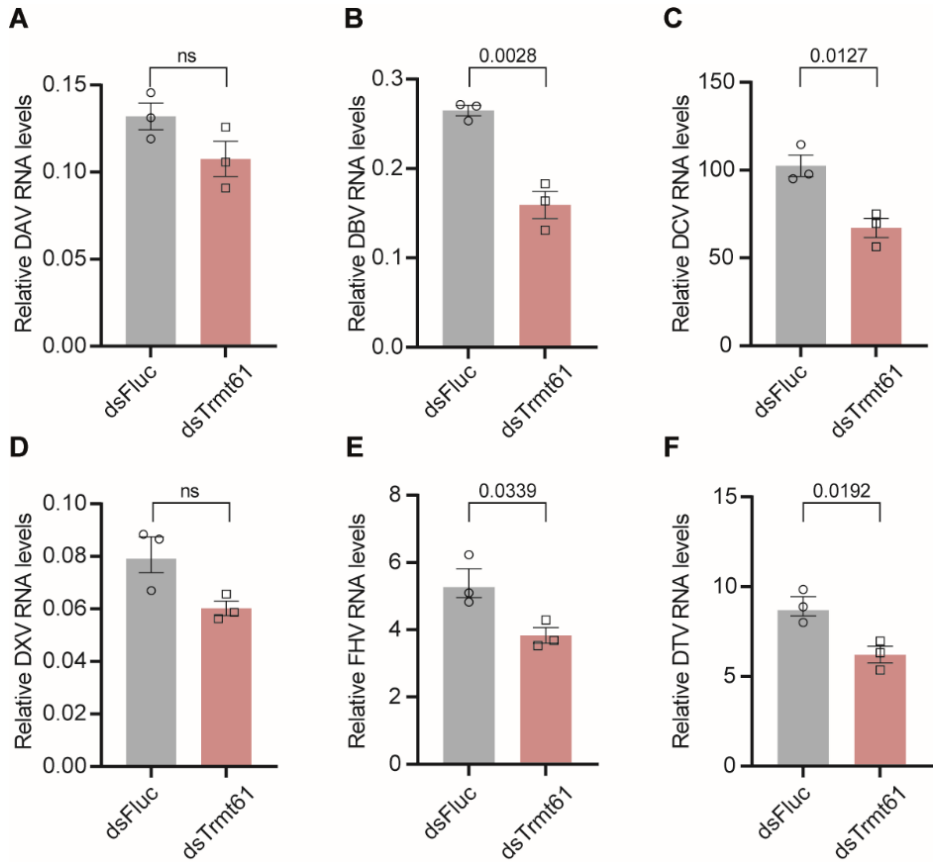


Figure 3. *Trmt61* depletion reduces virus RNA levels in persistently infected *Drosophila* cells. S2-R+ cells were treated with double-stranded RNA against *Trmt61* (dsTrmt61) or firefly luciferase (dsFluc) for 48 hours. Total RNA was then extracted from three biological replicates and viral RNA was determined by RT-qPCR for *Drosophila* A virus (DAV, **A**), *Drosophila* B birnavirus (DBV, **B**), *Drosophila* C virus (DCV, **C**), *Drosophila* X virus (DXV, **D**), Flock House virus (FHV, **E**) and *Drosophila* totivirus (DTV, **F**). Differences in viral RNA levels between *Trmt61* knockdown (dsTrmt61) and control knockdown (*firefly luciferase*, dsFluc) were analyzed by unpaired *t*-tests. Data are shown as mean and SD across the three replicates.

The apparent reduction in viral RNA replication upon *Trmt61* knockdown seems to conflict with the (inconsistent) results for *Trmt6* (Figure 1), considering that the human and yeast homologs of *Trmt61* and *Trmt6* dimerize in a complex that acts as a methyltransferase (48). *Trmt61* is the catalytic subunit that transfers a methyl group from the cofactor *S*-adenosyl-L-methionine to the RNA and *Trmt6* is the RNA-binding protein that helps refold the tRNA to make the substrate nucleotide more accessible (49). Together they predominantly methylate the N1 position of adenosine (m¹A) at nucleotide 58 of tRNAs, a modification that is

essential for tRNA structural stability (50), but they also deposit m¹A in mRNA, albeit to a lesser extent (51, 52). It remains elusive how *Trmt61* could have a proviral function independent of *Trmt6*, as, to the best of our knowledge, it has not been reported to act as a separate monomer. Interestingly, in a recent screen of human RNA-modifying enzymes, knockdown of a *Trmt61* homolog was found to increase viral RNA levels of respiratory syncytial virus, Semliki Forest virus and human metapneumovirus, while no effect of *Trmt6* knockdown was reported (53). Given these conflicting results among the human and the fly proteins, and between the monomers of the *Trmt61*-*Trmt6* complex, further research is required to elucidate the function of *Trmt61* in RNA virus infection.

Nop60B is possibly involved in RNA virus replication

Our RNAi screen further suggested that *Nop60B* depletion resulted in reduced DCV RNA levels (Figure 1). Unexpectedly, in none of these RNAi lines (UAS-*hpNop60B*) a detectable knockdown of target mRNA levels was observed (Figure 4A). Still, we analyzed the effect of *Nop60B* knockdown on infectious particle production and observed a significant 108-fold reduction of DCV titers in the GD, but not the KK line (Figure 4B). In contrast, CrPV titers were significantly reduced almost 9-fold upon *Nop60B* knockdown in the KK line (Figure 4C), and in an independent experiment, CrPV titers were similarly reduced although this difference did not reach statistical significance (Figure S3). Considering the lack of detectable knockdown, we can therefore not conclude that these phenotypes are directly caused by loss of *Nop60B*, and they might be due to a potential off-target effect instead.

Nop60B is the pseudouridine synthase of a ribonucleoprotein (RNP) complex consisting of a heterotetrameric protein core and RNAs containing the H/ACA motif (54). It plays essential functions in ribosomal RNA biogenesis in the nucleolus and defects in its human ortholog DKC-1 cause X-linked dyskeratosis congenita (55). Likewise, using CrPV IRES-mediated luciferase translation as a reporter, the Nop60B ortholog was shown to be required for efficient translation of internal ribosome entry site (IRES)-containing mRNAs in humans, mice and yeast (56, 57). Additionally, it has been proposed that H/ACA RNPs can localize in the cytoplasm and be recruited to stress granules (58), whose machinery is co-opted and exploited by some RNA viruses (59, 60). Recently, *Nop60B* was reported to have a proviral function in Sindbis virus (SINV) infection, based on an observed decrease in SINV RNA levels upon *Nop60B* knockdown in flies and an increase in SINV RNA in fly and human cells upon overexpression of *Nop60B* or its human ortholog *DKC-1*, respectively (61). In accordance, the authors

identified pseudouridine sites in the SINV genome that are complementary to the *Nop60B* RNA guide. Accordingly, introduction of a silent mutation in one of those loci slightly reduced SINV replication, suggesting that *Nop60B* directly pseudouridylates the viral genome to enhance viral replication, but further validation and the characterization of the underlying molecular mechanism are required.

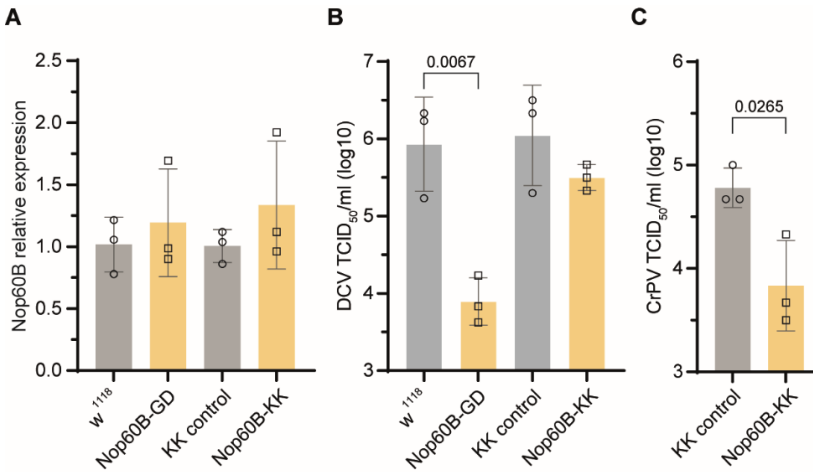


Figure 4. Effect of *Nop60B* knockdown on DCV and CrPV titers. UAS-hpNop60B males from the GD and the KK lines were crossed with Tub-Gal80^{ts}; Tub-Gal4 female virgins. hpNop60B/Tub-Gal80^{ts}; Tub-Gal4 were incubated at 29°C for four days, then inoculated with either 100 TCID₅₀ of DCV or 1,000 TCID₅₀ of CrPV. **(A)** *Nop60B* knockdown efficiency relative to control flies, **(B)** DCV titers, and **(C)** CrPV titers were determined at 24 hpi from three pools of five flies. For the latter, only the KK line was used due to the low number of GD flies available for this experiment. Titers were determined by end-point dilution assay. Differences were analyzed by unpaired *t*-tests. Expression was analyzed by RT-qPCR, normalized to *Rp49* housekeeping gene and expressed as fold change relative to the control flies. Data are shown as mean and SD across the three replicates.

The putative proviral function of *Trmt61* and *Nop60B* is independent of protein translation defects

Trmt61 and *Nop60B* are involved in protein synthesis as they predominantly modify tRNA and rRNA, respectively. We therefore explored if reduced protein translation efficiencies in the *Trmt61* and *Nop60B* knockdown lines could explain the observed reduction in virus replication. To this end, we profiled newly synthesized proteins in *Trmt61* and *Nop60B* depleted flies inoculated with CrPV or PBS as a control using THRONCAT (62). This technique uses the bioorthogonal threonine analog β -ethynylserine to label nascent proteins and azide-alkyne cycloaddition to visualize them. However, we found that the

protein synthesis was not altered by *Trmt61* or *Nop60B* depletion, nor in CrPV infected flies (Figure S4).

Conclusion

We have tried to silence expression of RNA-modifying enzymes to study their involvement in RNA virus replication in *Drosophila*. We found two candidates, *Trmt61* and *Nop60B*, whose knockdown reduced viral RNA levels. Their knockdown also appears to reduce the production of infectious viral particles. Our findings are constrained by technical limitations in the RNAi knockdown. Despite routine use in genome-wide studies, transgenic RNAi did not consistently reduce RNA expression of the target enzymes in our experiments.

To further characterize the putative proviral function of *Trmt61* and *Nop60B*, it would be necessary to determine whether the viral RNA is directly modified, or whether the phenotype is mediated by effects on a host protein required for replication or indirect effects on cell physiology. To determine the presence of modifications in the virus genome, viral RNA could be purified and analyzed by mass spectrometry. If the viral RNA were modified, the modified sites could subsequently be determined by modification mapping techniques, such as MeRIP-seq (Methylated RNA Immunoprecipitation Sequencing) or Nanopore direct RNA sequencing, and abrogated by introducing silent mutations in the modified loci to study their relevance for virus replication.

Materials and methods

Fly strains and husbandry

Flies were maintained at 25°C in standard fly food. Eggs were bleached and treated with tetracycline as described before (63), and the absence of Nora virus, DCV, DXV, CrPV and *Wolbachia* was verified by (RT-)PCR as described earlier (64).

The following RNAi stocks from Vienna Drosophila Research Center (42) were used:

RNA mod.	Gene name (<i>D. mel</i>)	Gene name (human)	CG gene number	GD lines ID	KK lines ID	Enzyme name
m1A	<i>Trmt6</i>	<i>TRMT6</i>	CG9596	27025	106324	tRNA (adenine (58)-N(1))-methyltransferase non-catalytic subunit
m1A	<i>Trmt61</i>	<i>TRMT61A</i>	CG14544	32279	107994	tRNA (adenine (58)-N(1))-methyltransferase SAM-dependent
m1A	<i>roswell</i>	<i>TRMT10C</i>	CG5190	28282		tRNA (adenine (9)-N(1))-methyltransferase
m6A	<i>Mettl3</i>	<i>METTL3</i>	CG5933	20968		N6-adenosine-methyltransferase
m6A	<i>METTL14</i>	<i>METTL14</i>	CG7818	48560	105434	N6-adenosine-methyltransferase non-catalytic subunit
m6A	<i>Ythdf</i>	<i>YTHDF1-3</i>	CG6422	27614	106251	YTH N6-methyladenosine RNA binding protein
m5C	<i>Nsun2</i>	<i>NSUN2</i>	CG6133	37601		RNA cytosine-C(5)-methyltransferase
m5C	<i>Nsun5</i>	<i>NSUN5</i>	CG42358	46814	106765	RNA cytosine-C(5)-methyltransferase SAM-dependent
m5C	<i>Mt2</i>	<i>DNMT2/ TRDMT1</i>	CG10692	37815		DNA cytosine-C(5)-methyltransferase
m5C	<i>Ref1</i>	<i>ALYREF</i>	CG1101	12031	104471	RNA-binding domain superfamily and export factor
m3C		<i>METTL6</i>	CG34195	34311	105449	tRNA N(3)-methylcytidine methyltransferase SAM-dependent
ac4C	<i>lethal(1) G0020</i>	<i>NAT10</i>	CG1994	13479		RNA cytidine acetyltransferase acetyl-CoA dependent
ψ	<i>Nop60B</i>	<i>DKC1</i>	CG3333	34597	109616	pseudouridine synthase - catalytic domain
Inosine	<i>Adar</i>	<i>ADARB1</i>	CG12598	7763	105612	double-stranded RNA adenine deaminase
(2'-O-)-MTase		<i>FTSJ3</i>	CG8939	40726		rRNA (guanine and uridine-2'-O-)-methyltransferase
mcm5s2U	<i>Ctu1</i>	<i>CTU1</i>	CG8078	11107	104409	Cytosolic thiouridylase subunit 1
mcm5s2U	<i>Ctu2</i>	<i>CTU2</i>	CG10189	52087	100704	Cytosolic thiouridylase subunit 2
DHU	<i>Dus2</i>	<i>DUS2</i>	CG1434	30994	104876	tRNA-dihydrouridine(20) synthase [NAD(P) (+)]
	<i>AGO2</i>	<i>AGO1-4</i>	CG7439	49473		Argonaute 2
	KK control				60100	

Three Gal4-expressing driver lines were used: *Tubulin-Gal4/TM3 Sb* (Tub-gal4), *w¹¹¹⁸*, *UAS-Dicer-2; Actin-Gal4/CyO* (*UAS-Dicer-2; Act-Gal4*) and *Tubulin-Gal80^{ts}; Tubulin-Gal4/TM6, Hu, Tb* (*Tub-gal80^{ts}; Tub-Gal4*). These lines carry balancer chromosomes (TM3, CyO, and TM6, respectively) that contain dominant phenotypic markers, allowing easy identification of flies carrying the balancer, as well as recessive lethal or sterile alleles, which ensures that the balancer cannot exist in homozygosis.

Male *UAS-hpRNA* and control flies were crossed with virgin female flies from the driver lines, after which flies lacking the *Sb*, *CyO* or *Hu* markers were selected from the F1 offspring and two to five-day-old female flies were used for the virus infection. As controls, *w¹¹¹⁸* were used for the GD lines and *w¹¹¹⁸* flies with a target insertion in the KK landing site (position chr2L: 22019296, cytological band 40D3) was used as the control for the KK lines (referred to KK control). Of note, KK lines have also been shown to have a previously non-annotated landing site at position chr2L: 9437482, cytological band 30B3 (65).

Lethality caused by constitutive knockdown was only assessed once, as all *hpRNA/Tub-gal4* replicates (GD1, GD2 and GD3 in Figure 1) were produced from a single fly cross.

The *Tub-gal80^{ts}; Tub-Gal4* driver was used in part of the screen (Figure 1) and in all follow-up experiments. Here, fly crosses were kept at 18°C, after which flies lacking the *Hu* marker were selected from the F1 offspring and two to five-day-old female flies were transferred to a 29°C incubator for 4 days, injected with the specified virus inoculum and incubated for 24h at 29°C before harvesting.

Cells and viruses

The parental DCV and CrPV stocks were produced in *Drosophila* S2 cells (Invitrogen), which were maintained in Schneider's medium supplemented with 10% fetal calf serum (Sigma) and 50 U/ml penicillin and 50 µg/ml streptomycin (Gibco) at 27°C. CrPV stock was produced from an infectious clone as described in Chapter 3 (66). DCV previously passaged in WT flies for 10 generations in Chapter 2 (64) was bottlenecked by serial dilution in a 24-well plate, selecting the highest dilution that still showed cytopathic effects to inoculate a T25 flask of confluent S2 cells containing 7 ml of medium. After 72 h, the supernatant was collected, centrifuged at 300 *g* for 5 min, and stored in aliquots at -80°C. Titers were determined by end-point dilution in 96-

well plates, as described before (63), and expressed as median tissue culture infectious dose (TCID₅₀), calculated with the Reed–Muench method.

Females flies of each genotype were intrathoracically inoculated with 100 TCID₅₀ of DCV or 1,000 TCID₅₀ of CrPV as specified in 50 nl of PBS pH 7.3, using a Nanoject II microinjector (Drummond Scientific).

Templates for *in vitro* transcription of *Trmt61* double-stranded RNA (dsRNA) were generated by PCR on complementary DNA from *Drosophila* S2 cells (Invitrogen), using primers Trmt61For (taatacgactcactatagggaga CCCCACCGCACACAGATTAT) and Trmt61Rev (taatacgactcactatagggaga TTGCAACGTCCCGGTGATA). Templates for control dsRNA (*firefly luciferase*) were produced by PCR on a plasmid containing *firefly luciferase* sequence, using primers fLucFor (taatacgactcactatagggagaTATGAAGAGATACGCCCTGGTT) and fLucRev (taatacgactcactatagggagaTAAAACCGGGAGGTAGATGAGA). The T7 promoter sequence was introduced at both forward and reverse primers, corresponding to the lower case portion of each primer. The resulting PCR products were used as templates for *in vitro* transcription with T7 RNA polymerase for 4 h at 37 °C. RNA annealing was performed by denaturing the transcripts at 95 °C followed by gradual cooling to room temperature. The dsRNAs were purified using RNA Clean & Concentrator Kits (cat. no. R1014, Zymo Research) according to the manufacturer's instructions.

Drosophila S2-R+ cells persistently infected with DAV, DBV, DCV, DXV, FHV and DTV were plated in 24-wells plates in complete medium, and 1 µg of dsRNA against *Trmt61* or *firefly luciferase* was added to the culture medium (soaking). After an incubation for 48 hours, total RNA was extracted and viral RNA quantified by RT-qPCR.

RT-qPCR

RT-qPCR was performed as previously described (64), using primers CrPVFor (ACGAGGAAGCAACTCAAGG) and CrPVRev (GAGCCCGCTGAGATGTAAAG) for CrPV, and DCVFor (TCAAGAAAAGTTGCGTGGGT) and DCVRev (CAGAGCGTCTTGGAGAGTG) for DCV. Target gene expression was quantified using the following primers:

Gene	Orientation	Sequence (5' to 3')
<i>Trmt61</i>	For	CAAGGCGTTGAAGCTGTCCG
	Rev	TCTTCACCACGTTCTCCTGC
<i>Mettl3</i>	For	CCAACCAGGGTGCAATTGTG
	Rev	TCAACATCGTCGACTTCCAAC
<i>Ythdf</i>	For	CTGCCTGTCGATTCCCACT
	Rev	GACAAAGAATCTGGCCGATCC
<i>Nsun2</i>	For	GCAAACGTCAAAAGCGGGAG
	Rev	AACTGCGTCCACTCCTCATC
<i>Nsun5</i>	For	GATGAGGACGACGAGCCG
	Rev	GCCTCCGCCAAGCTGTATAG
<i>Mt2</i>	For	TGGAACACGTGCCACTCA
	Rev	CCAATTGGGCATAATTAAAGGCA
<i>FTSJ3</i>	For	CCGCAACTAAAGACAGCAGC
	Rev	CCTTCTTGGCATTGTACCTG
<i>Ctu2</i>	For	CTTCGCGCAACACAGAACA
	Rev	CTGCAGAGCTTCCAAGTCCA
<i>Dus2</i>	For	GAGCTGGTGGACATCAAGC
	Rev	TTTGTAGTACCAGCCGTGAG
<i>Ago2</i>	For	CCGTAACGCCTATCCTGATCA
	Rev	CACCACAATCACGCAGCAAA
<i>Trmt6</i>	For	TCAAATAGTGGTAGAGAAGGAGCA
	Rev	TTTCACAAGAGGCAGCAGGG
<i>Trmt10C</i>	For	GATGAGCCCTTCGATCTGCA
	Rev	CAGTGGGGCGTTAGGTAGAC
<i>Mettl14</i>	For	CGGTCTTTCAACGCACCAAG
	Rev	GCTGCCAAACTCCTCTTCCT
<i>NAT10</i>	For	GAGCCAGTGAATCCCGCAG
	Rev	GCTTCAGTTGCTTGTCGACC
<i>Nop60B</i>	For	CCAAAGTCACGGGTTGCCT
	Rev	TTTACGGCGGAGATGAGTGG
<i>Ref</i>	For	TCAGAAGGCCAAGTTCCAC
	Rev	CCAGGTTACCGACGATGAGG
<i>Mettl6</i>	For	GCTCTCCGCCATACATCCAC
	Rev	TCTGCTATCTTGTTGCCCGG
<i>Adar</i>	For	TGCCACAGAAGAACGTACCC
	Rev	CCCGTCGAAACTGAAATAACTTT
<i>Ctu1</i>	For	AGCGCACGTATTAATACTGCT
	Rev	GCATCTGGTAGTCGTACGG

Nascent protein labelling

THRONCAT was used in *Drosophila melanogaster* to label newly synthesized proteins (62). In short, two to five-day-old female flies were incubated at 29°C for 4 days, inoculated with 1,000 TCID₅₀ of CrPV or PBS as mock infection and transferred to 4mM β-ethynylserine (βES)-containing food for 24 h at 29°C. Flies were harvested, snap frozen, and stored at -80°C. Pools of five flies were homogenized with a pestle and proteins were extracted by cell lysis in 50 mM HEPES, pH 7.4, 150 mM NaCl, 0.5% Nonidet P-40, 1% Sodium deoxycholate, 1% SDS, 0.5% Triton-X-100, 5% Glycerol, 1 x EDTA-free protease inhibitor cocktail. Lysates were cleared by centrifugation at 13,000 g for 30 min at 4°C, and the supernatants were transferred to new tubes. Total protein content was normalized based on BCA measurements (Thermo Fisher). Newly synthesized proteins, metabolically labelled with βES, were conjugated with Cy5-azide via Copper(I)-catalyzed Azide-Alkyne Cycloaddition (CuAAC). The CuAAC was assembled in a sequential reaction by mixing 50 µg of protein lysate in dilution buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 1% Triton-X-100, 1 x EDTA-free protease inhibitor cocktail) and adding CuSO₄ (10 mM), THPTA (20 mM), Cy5-azide (10 µM) and sodium ascorbate (100 mM). After each addition, the solution was vortexed for 10 sec and at the end incubated at 37°C for 30 min in an orbital shaker at 550 rpm. The proteins were precipitated by adding 400 µl of ice-cold acetone and centrifuging at 5,000 g 20 min at 4°C. Samples were resuspended in Laemmli buffer (Bio-Rad), incubated at 70°C for 10 min, and analysed on an SDS-PAGE (12% polyacrylamide gel). The gel was fixed in 50% methanol and 7% acetic acid and total proteins were stained with SYPRO Ruby. In-gel fluorescence was measured with an Odyssey XF scanner using the 600 nm and 700 nm lasers for SYPRO® and Cy5, respectively.

Statistical analyses

Viral titers between knockdown and control flies, relative expression of target genes, as well as relative viral RNA levels of endogenous viruses infecting S2-R+ cells, were compared using unpaired *t*-tests.

To assess differences in DCV replication kinetics over time between *Trmt61* knockdown and KK control flies, relative viral RNA levels were analyzed using a two-way analysis of variance (ANOVA) with replication. The two factors included time post-infection (0, 12, 24, and 48 hpi) and genotype. Three independent biological replicates were analyzed for each genotype at each time point. The analysis evaluated the main effects of time and genotype, as well as their interaction.

To evaluate the overall magnitude of viral RNA levels across the infection time course, the area under the curve (AUC) was calculated for each biological replicate using the trapezoidal rule:

$AUC_i = (t_{i+1} - t_i) \times (y_i + y_{i+1})/2$ where t represents time post-infection and y represents relative DCV RNA levels. AUC values were obtained by calculating viral RNA levels over time between consecutive time points (0–12, 12–24, and 24–48 hpi), and summing these values to yield a total AUC for each replicate. Mean AUC values were then calculated for each genotype, and the relative difference in viral RNA levels per genotype was expressed as the ratio of the mean AUC of KK control flies to that of *Trmt61* knockdown flies.

Acknowledgements

We thank members of the laboratory for discussion, Natalia Mora for kindly providing the driver lines UAS-Dicer-2; Act-Gal4 and Tub-gal80^{ts}; Tub-Gal4, as well as the Vienna Drosophila Resource Center (VDRC, Vienna, Austria) for providing the UAS-hpRNA flies.

Funding

This study was financially supported by the European Union's Horizon 2020 research and innovation program, under the Marie Skłodowska-Curie Actions Innovative Training Networks grant agreement no. 955974 (VIROINF) to RpvR and the China Scholarship Council program (Grant No. 202208320155) to HW. The funders had no role in study design, data collection and analysis, or preparation of the manuscript.

References

1. Cappannini A, Ray A, Purta E, Mukherjee S, Boccaletto P, Moafinejad SN, et al. MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Research*. 2024;52(D1):D239-D44.
2. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation. *Cell*. 2017;169(7):1187-200.
3. Zhao BS, Roundtree IA, He C. Post-transcriptional gene regulation by mRNA modifications. *Nature Reviews Molecular Cell Biology*. 2017;18(1):31-42.
4. Cohn WE. Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *Journal of Biological Chemistry*. 1960;235(5):1488-98.
5. Davis FF, Allen FW. Ribonucleic acids from yeast which contain a fifth nucleotide. *Journal of Biological Chemistry*. 1957;227(2):907-15.
6. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences*. 1974;71(10):3971-5.
7. Dubin DT, Taylor RH. The methylation state of poly A-containing-messenger RNA from cultured hamster cells. *Nucleic Acids Research*. 1975;2(10):1653-68.
8. Holley RW, Everett GA, Madison JT, Zamir A. Nucleotide sequences in the yeast alanine transfer ribonucleic acid. *Journal of Biological Chemistry*. 1965;240(5):2122-8.
9. Canaani D, Kahana C, Lavi S, Groner Y. Identification and mapping of N 6-methyladenosine containing sequences in simian virus 40 RNA. *Nucleic Acids Research*. 1979;6(8):2879-99.
10. Dimock K, Stoltzfus CM. Sequence specificity of internal methylation in B77 avian sarcoma virus RNA subunits. *Biochemistry*. 1977;16(3):471-8.
11. Krug RM, Morgan MA, Shatkin AJ. Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *Journal of Virology*. 1976;20(1):45-53.
12. Helm M, Motorin Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics*. 2017;18(5):275-91.
13. Saletore Y, Meyer K, Korfach J, Vilfan ID, Jaffrey S, Mason CE. The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome Biology*. 2012;13(10):175.
14. Boo SH, Kim YK. The emerging role of RNA modifications in the regulation of mRNA stability. *Experimental and Molecular Medicine*. 2020;52(3):400-8.
15. Roost C, Lynch SR, Batista PJ, Qu K, Chang HY, Kool ET. Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *Journal of the American Chemical Society*. 2015;137(5):2107-15.
16. Harcourt EM, Kietrys AM, Kool ET. Chemical and structural effects of base modifications in messenger RNA. *Nature*. 2017;541(7637):339-46.
17. Zhao X, Yang Y, Sun B-F, Shi Y, Yang X, Xiao W, et al. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Research*. 2014;24(12):1403-19.

18. Yang X, Yang Y, Sun B-F, Chen Y-S, Xu J-W, Lai W-Y, et al. 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Research*. 2017;27(5):606–25.
19. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N6-methyladenosine modulates messenger RNA translation efficiency. *Cell*. 2015;161(6):1388–99.
20. Gilbert WV, Bell TA, Schaening C. Messenger RNA modifications: form, distribution, and function. *Science*. 2016;352(6292):1408–12.
21. Nachtergaele S, He C. Chemical modifications in the life of an mRNA transcript. *Annual Review of Genetics*. 2018;52(1):349–72.
22. Xue C, Chu Q, Zheng Q, Jiang S, Bao Z, Su Y, et al. Role of main RNA modifications in cancer: N6-methyladenosine, 5-methylcytosine, and pseudouridine. *Signal Transduction and Targeted Therapy*. 2022;7(1):142.
23. Price A, Hayer K, McIntyre A, Gokhale N, Abebe J, Della Fera A, et al. Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing. *Nature Communication*. 2020;11(1):6016.
24. Saikia B, Riquelme-Barrios S, Carell T, Brameyer S, Jung K. Depletion of m6A-RNA in *Escherichia coli* reduces the infectious potential of T5 bacteriophage. *Microbiology Spectrum*. 2024;12(12):e01124–24.
25. Kim D, Lee J-Y, Yang J-S, Kim JW, Kim VN, Chang H. The architecture of SARS-CoV-2 transcriptome. *Cell*. 2020;181(4):914–21. e10.
26. Ge L, Pan F, Jia M, Pott DM, He H, Shan H, et al. RNA modifications in plant biotic interactions. *Plant Communications*. 2025;6(2):101232.
27. Bosmeny MS, Pater AA, Zhang L, Sha BE, Lyu Z, Larkai L, et al. An HIV-1 Reference Epitranscriptome. *bioRxiv [Preprint]*. 2025:2025.01.30.635805.
28. Baek A, Lee G-E, Golconda S, Rayhan A, Manganaris AA, Chen S, et al. Single-molecule epitranscriptomic analysis of full-length HIV-1 RNAs reveals functional roles of site-specific m6As. *Nature Microbiology*. 2024;9(5):1340–55.
29. Kennedy EM, Bogerd HP, Kornepati AV, Kang D, Ghoshal D, Marshall JB, et al. Posttranscriptional m6A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host and Microbe*. 2016;19(5):675–85.
30. Kong W, Biswas A, Zhou D, Fiches G, Fujinaga K, Santoso N, et al. Nucleolar protein NOP2/NSUN1 suppresses HIV-1 transcription and promotes viral latency by competing with Tat for TAR binding and methylation. *PLoS Pathogens*. 2020;16(3):e1008430.
31. Courtney DG, Tsai K, Bogerd HP, Kennedy EM, Law BA, Emery A, et al. Epitranscriptomic addition of m5C to HIV-1 transcripts regulates viral gene expression. *Cell Host and Microbe*. 2019;26(2):217–27. e6.
32. Ringgaard M, Marchand V, Decroly E, Motorin Y, Bennasser Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature*. 2019;565(7740):500–4.
33. Tsai K, Vasudevan AAJ, Campos CM, Emery A, Swanstrom R, Cullen BR. Acetylation of cytidine residues boosts HIV-1 gene expression by increasing viral RNA stability. *Cell Host and Microbe*. 2020;28(2):306–12. e6.
34. Gokhale NS, McIntyre AB, McFadden MJ, Roder AE, Kennedy EM, Gandara JA, et al. N6-methyladenosine in flaviviridae viral RNA genomes regulates infection. *Cell Host and Microbe*. 2016;20(5):654–65.

35. Lichinchi G, Zhao BS, Wu Y, Lu Z, Qin Y, He C, et al. Dynamics of human and viral RNA methylation during Zika virus infection. *Cell Host and Microbe*. 2016;20(5):666-73.
36. Kim B, Arcos S, Rothamel K, Jian J, Rose KL, McDonald WH, et al. Discovery of widespread host protein interactions with the pre-replicated genome of CHIKV using VIR-CLASP. *Molecular Cell*. 2020;78(4):624-40. e7.
37. McIntyre AB, Gokhale NS, Cerchietti L, Jaffrey SR, Horner SM, Mason CE. Limits in the detection of m6A changes using MeRIP/m6A-seq. *Scientific Reports*. 2020;10(1):6590.
38. Xiao Y, Wang Y, Tang Q, Wei L, Zhang X, Jia G. An elongation-and ligation-based qPCR amplification method for the radiolabeling-free detection of locus-specific N6-methyladenosine modification. *Angewandte Chemie International Edition*. 2018;57(49):15995-6000.
39. Baquero-Pérez B, Yonchev ID, Delgado-Tejedor A, Medina R, Puig-Torrents M, Sudbery I, et al. N 6-methyladenosine modification is not a general trait of viral RNA genomes. *Nature Communications*. 2024;15(1):1964.
40. Thompson MG, Sacco MT, Horner SM. How RNA modifications regulate the antiviral response. *Immunological Reviews*. 2021;304(1):169-80.
41. Esteve-Puig R, Bueno-Costa A, Esteller M. Writers, readers and erasers of RNA modifications in cancer. *Cancer letters*. 2020;474:127-37.
42. Dietzl G, Chen D, Schnorrer F, Su K-C, Barinova Y, Fellner M, et al. A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*. 2007;448(7150):151-6.
43. Schnorrer F, Schonbauer C, Langer CC, Dietzl G, Novatchkova M, Schernhuber K, et al. Systematic genetic analysis of muscle morphogenesis and function in *Drosophila*. *Nature*. 2010;464(7286):287-91.
44. Yapici N, Kim YJ, Ribeiro C, Dickson BJ. A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour. *Nature*. 2008;451(7174):33-7.
45. Neely GG, Hess A, Costigan M, Keene AC, Goulas S, Langeslag M, et al. A genome-wide *Drosophila* screen for heat nociception identifies alpha2delta3 as an evolutionarily conserved pain gene. *Cell*. 2010;143(4):628-38.
46. Neely GG, Kuba K, Cammarato A, Isobe K, Amann S, Zhang L, et al. A global in vivo *Drosophila* RNAi screen identifies NOT3 as a conserved regulator of heart function. *Cell*. 2010;141(1):142-53.
47. Kaya-Copur A, Schnorrer F. A Guide to genome-wide in vivo RNAi applications in *Drosophila*. *Methods in Molecular Biology*. 2016;1478:117-43.
48. Ozanick S, Krecic A, Andersland J, Anderson JT. The bipartite structure of the tRNA m1A58 methyltransferase from *S. cerevisiae* is conserved in humans. *Rna*. 2005;11(8):1281-90.
49. Finer-Moore J, Czudnochowski N, O'Connell III JD, Wang AL, Stroud RM. Crystal structure of the human tRNA m1A58 methyltransferase-tRNA3Lys complex: refolding of substrate tRNA allows access to the methylation target. *Journal of Molecular Biology*. 2015;427(24):3862-76.
50. Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes and Development*. 2004;18(11):1227-40.
51. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, et al. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature*. 2017;551(7679):251-5.

52. Li X, Xiong X, Zhang M, Wang K, Chen Y, Zhou J, et al. Base-resolution mapping reveals distinct m1A methylome in nuclear-and mitochondrial-encoded transcripts. *Molecular Cell*. 2017;68(5):993-1005. e9.
53. Jarmsom IG. Investigating the role of RNA modifying enzymes in RNA virus infection: Doctoral dissertation, University of Glasgow; 2025.
54. Garus A, Autexier C. Dyskerin: an essential pseudouridine synthase with multifaceted roles in ribosome biogenesis, splicing, and telomere maintenance. *Rna*. 2021;27(12):1441-58.
55. Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ, et al. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature Genetics*. 1998;19(1):32-8.
56. Yoon A, Peng G, Brandenburg Y, Zollo O, Xu W, Rego E, et al. Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science*. 2006;312(5775):902-6.
57. Jack K, Bellodi C, Landry DM, Niederer RO, Meskauskas A, Musalgaonkar S, et al. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Molecular Cell*. 2011;44(4):660-6.
58. Belli V, Matrone N, Sagliocchi S, Incarnato R, Conte A, Pizzo E, et al. A dynamic link between H/ACA snoRNP components and cytoplasmic stress granules. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. 2019;1866(12):118529.
59. Lloyd RE. How do viruses interact with stress-associated RNA granules? *PLoS Pathogens*. 2012;8(6):e1002741.
60. Lloyd RE. Regulation of stress granules and P-bodies during RNA virus infection. *Wiley Interdisciplinary Reviews: RNA*. 2013;4(3):317-31.
61. Stark N, Podicheti R, Garcia L, Krenz A, Rusch DB, Newton IL, et al. Pseudouridine synthases are proviral factors for Sindbis virus in insect and mammalian cells. *mBio*. 2025:e01329-25.
62. Ignacio BJ, Dijkstra J, Mora N, Slot EF, van Weijsten MJ, Storkebaum E, et al. THRONCAT: metabolic labeling of newly synthesized proteins using a bioorthogonal threonine analog. *Nature Communications*. 2023;14(1):3367.
63. Merklings SH, Van Rij RP. Analysis of resistance and tolerance to virus infection in *Drosophila*. *Nature Protocols*. 2015;10(7):1084-97.
64. Lezcano OM, Fuhrmann L, Ramakrishnan G, Beerenwinkel N, Huynen MA, van Rij RP. Parallel evolution and enhanced virulence upon *in vivo* passage of an RNA virus in *Drosophila melanogaster*. *Virus Evolution*. 2023;9(2):vead074.
65. Green EW, Fedele G, Giorgini F, Kyriacou CP. A *Drosophila* RNAi collection is subject to dominant phenotypic effects. *Nature Methods*. 2014;11(3):222-3.
66. Lezcano OM, Fuhrmann L, Bos RT, Wang H, Stevens M, Beerenwinkel N, Huynen MA, van Rij RP. A recurrent adaptive mutation in the transmembrane 2B protein of an insect picornaviralike virus in a nonnative host. *J Virol*. 2026;100(1):e0123925.

Supplementary figures

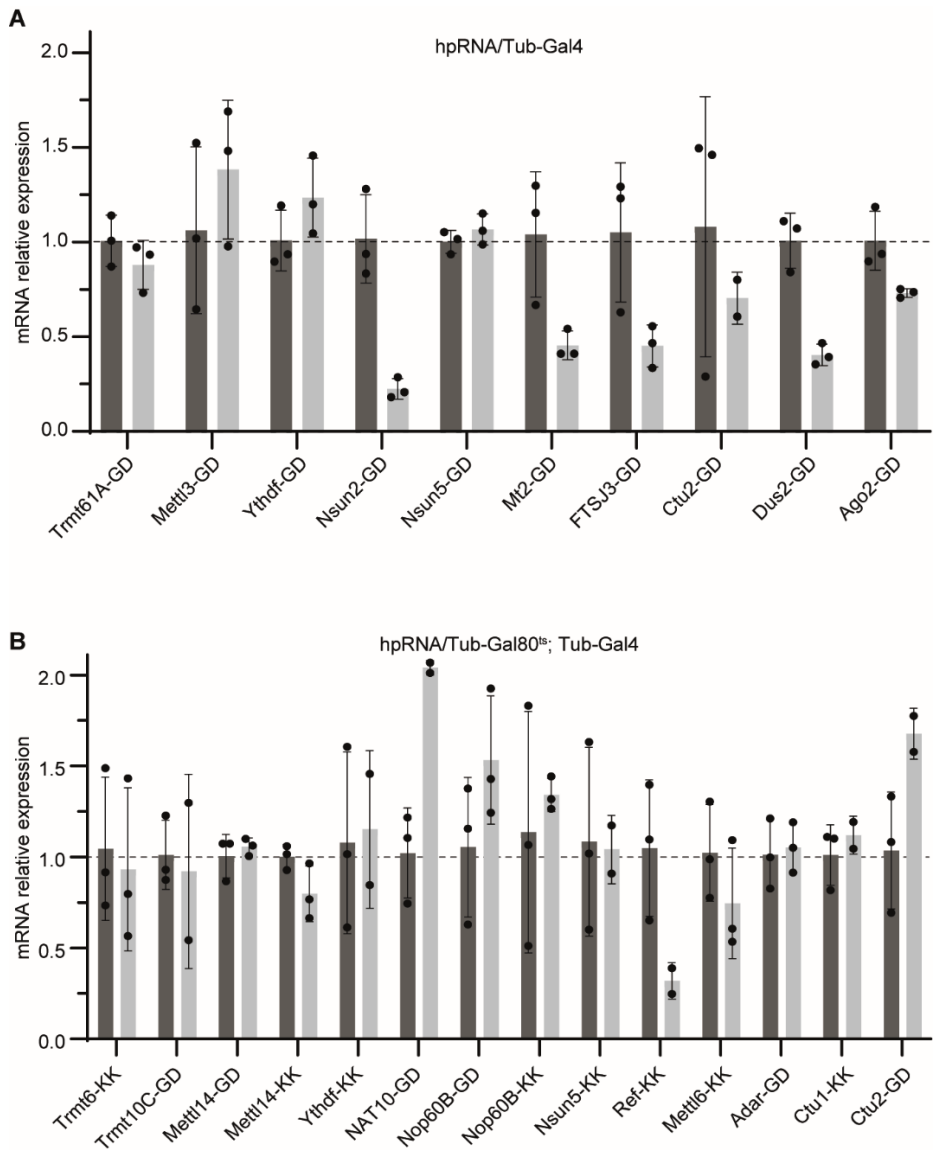


Figure S1. Knockdown efficiency of RNAi lines. Target gene expression in UAS-hpRNA flies (light gray bars) relative to control flies (dark gray bars). Gene expression was assessed by RT-qPCR from samples from crosses with the Tub-Gal4 driver line used in Figure 1 (experiment GD-1; **A**) and samples from crosses with the Tub-Gal80^{ts}; Tub-Gal4 driver from Figure 1 (experiment KK-3 and GD-7; **B**), using gene-specific primers.

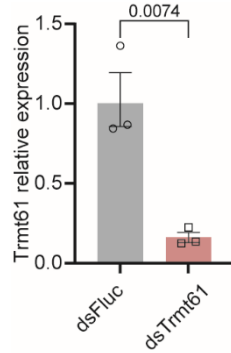


Figure S2. *Trmt61* knockdown in *Drosophila* cells. S2-R+ cells were soaked with double-stranded RNA against *Trmt61* (dsTrmt61) or *firefly luciferase* (dsFluc) for 48 hours and *Trmt61* expression relative to the control was measured by RT-qPCR. Data were expressed as mean and SD of three replicates and significant differences were analyzed by unpaired *t*-test.

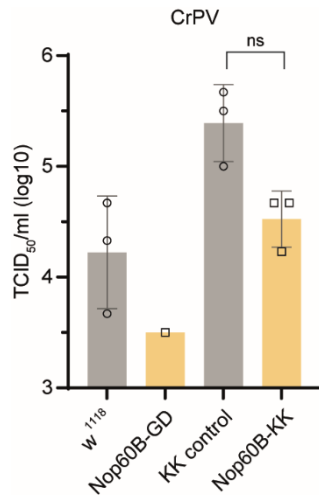


Figure S3. Biological replicate of *Nop60B* knockdown effect on CrPV titers. UAS-hpNop60B males from the GD and the KK lines were crossed with Tub-Gal80^{ts}; Tub-Gal4 female virgins. hpNop60B/Tub-Gal80^{ts}; Tub-Gal4 were incubated at 29°C for four days, then inoculated with 1,000 TCID₅₀ of CrPV. Titters were determined at 24 hpi by end-point dilution assay from three pools of five flies (KK line) and from only one pool for the GD line due to the low number of available flies for this experiment. Differences were analyzed by unpaired *t*-tests (KK line). Data are shown as mean and SD.

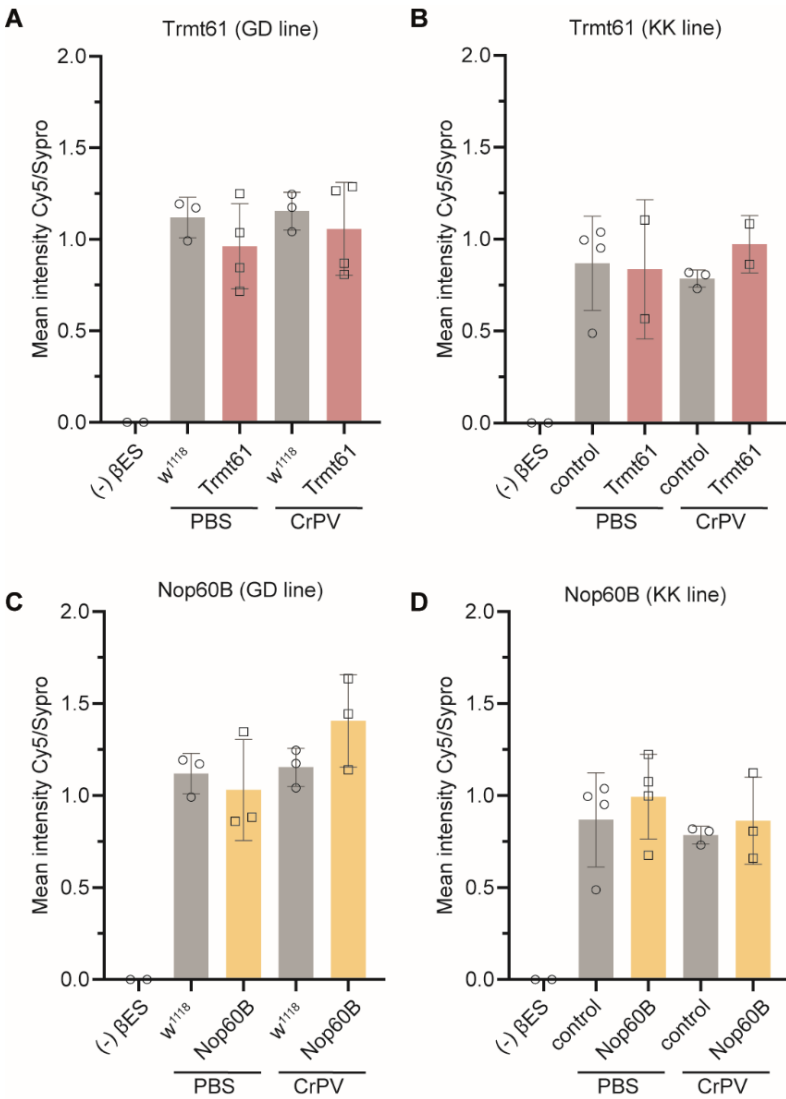
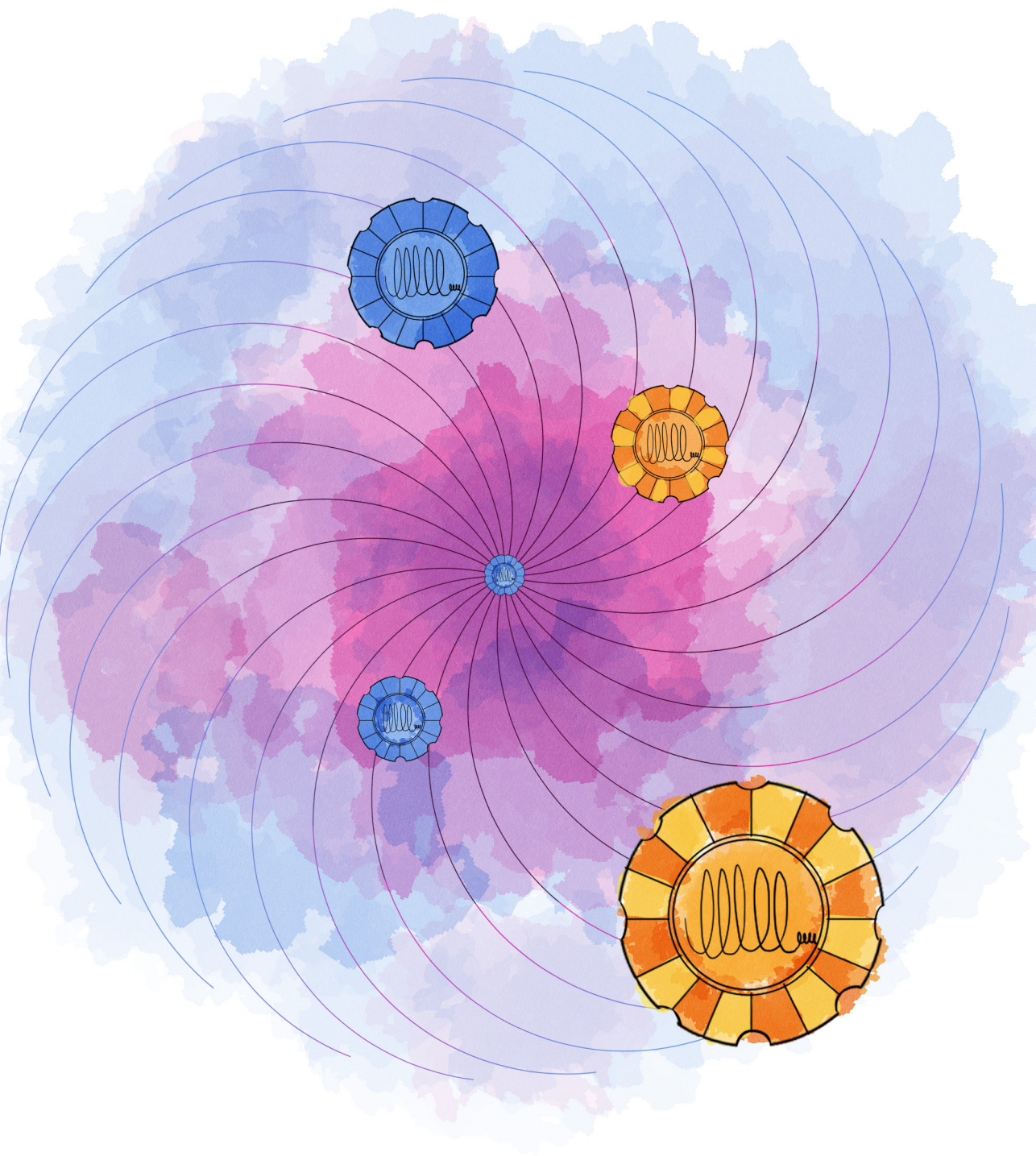


Figure S4. Total protein synthesis is not affected by *Trmt61* and *Nop60B* knockdown nor by cricket paralysis virus infection. UAS-hpNop60B and UAS-hpTrmt61 males from the GD and the KK lines were crossed with Tub-Gal80^{ts}; Tub-Gal4 female virgins. *Trmt61* or *Nop60B* knockdown and control flies were selected and incubated at 29°C for four days, inoculated with 1000 TCID₅₀ of CrPV and transferred to food containing 4 mM β -ethynylserine (β ES) for 24 hours. Proteins from pools of five flies were extracted and newly incorporated β ES was labeled with Cy5-azide. Proteins were separated by SDS-PAGE and visualized by in-gel fluorescence, with total protein stained using SYPRO. The amount of nascent protein in *Trmt61* (A, B) and *Nop60B* (C, D) knockdown flies is shown as the ratio Cy5/SYPRO relative to the controls in the infected and non-infected conditions. Data are shown as mean and SD across the replicates. GD or KK lines were used as indicated.



Chapter 5

General discussion

Understanding how host innate immunity shapes virus evolution remains an open question in virology. During my PhD, I developed an experimental model to directly test how modulation of key antiviral pathways influences viral adaptation and pathogenesis. By experimentally abrogating or stimulating innate immune responses and serially passaging viruses in fruit flies, I aimed to determine whether changes in immune pressure resulted in viral adaptation through inoculation of well over 15,000 flies. Using the native fly pathogen *Drosophila C virus* (DCV), I passaged virus populations in *Drosophila melanogaster* in which the activity of the RNA interference pathway (**Chapter 2**) and the cGAS-Sting pathway (**Chapter 3**) was experimentally modulated. In the latter case, I also studied cross-species transmission by serial passage of Cricket paralysis virus (CrPV), which was originally isolated from crickets (*Teleogryllus oceanicus* and *Teleogryllus commodus*) and has not previously been detected in wild fly populations. Finally, I investigated the role of RNA-modifying enzymes in RNA virus replication (**Chapter 4**), which have recently attracted attention as potential regulators of infection.

Here, I discuss our findings and open questions in light of the literature on virus evolution, ecology and phylodynamics. I present viruses as integral components of ecosystems, rather than merely disease agents, and I elaborate on the importance of environmental factors and host switching as central catalysts for virus evolution. I review our knowledge on immune driven evolution, and finally I discuss the pros and cons of our experimental model.

Why study virus evolution?

Viruses have a major impact on human societies, driving substantial morbidity and mortality, and imposing a severe economic burden worldwide. The Coronavirus Disease 2019 (COVID-19) pandemic caused by SARS-Coronavirus 2 (SARS-CoV-2) resulted in more than 14 million excess deaths, according to estimates of the World Health Organization (1), though the exact figure remains debated (2, 3). The pandemic caused a 6.5% decrease in economic activity in the Eurozone in 2020 (4), and associated healthcare costs in the United States exceeded \$200 billion in the same year (5). Other epidemic viruses are also leaving devastating death tolls: 630,000 people died of acquired immunodeficiency syndrome (AIDS) in 2024 (6), and influenza virus was responsible for 27,000 to 130,000 deaths in the 2024-2025 season (7).

It should therefore be a social and political priority to control these pathogenic agents. Translational efforts aiming to improve patients' quality of life are

important but insufficient when dealing with emerging viruses. Ergo, the role of basic research in this matter is to unravel virus-host interactions at the molecular level and identify vulnerabilities in the virus lifecycle that inform targeted strategies for virus control and vaccine development. A landmark example is the discovery of reverse transcriptase enzymes (8, 9), which enabled the development of nucleoside analog reverse transcriptase inhibitors such as Zidovudine, the first Food and Drug Administration-approved antiretroviral drug for the treatment of AIDS (10). It was also observed that highly pathogenic avian influenza strain A/H5N1 remained sensitive to the antiviral drug oseltamivir after acquisition of airborne transmission in ferrets through fixation of nine amino acid substitutions (11). A deep understanding of the process of virus-host adaptation, and in particular the interplay between the virus and the host immune system, could provide a conceptual framework for predicting how viruses acquire drug resistance and evade immune responses.

Beyond their relevance to public health, viruses provide a powerful lens to examine the principles of molecular biology and genome evolution. The use of viruses as relevant experimental models is assisted by their ubiquity (12), relatively simple lifecycle, small genomes, and short generation times. Historically, researching viruses, and particularly bacteriophages, has led to seminal discoveries that shaped the understanding of life. For instance, the fact that genetic material is composed of DNA instead of proteins (13), and that messenger RNA (mRNA) is a transient carrier of genetic information between genes and ribosomes (14), were discoveries made by studying bacteriophage infection in bacteria. Therefore, by dissecting the molecular architecture of viruses, we gain relevant insights into life itself.

Virus-host co-evolution is often described as an arms race, in which two species continuously have to adapt to each other to retain their fitness—a dynamic also referred to as the Red Queen hypothesis (15, 16). Under this concept, viruses evading the immune response have a fitness advantage and are thus positively selected from the population. In turn, the host immune machinery adapts to better fight the evolved pathogens, subjecting immune genes to strong selective pressures that accelerates protein evolution and causes the diversification of immune components of closely related species (17, 18). Signature immune pathways under strong selection are the RNA interference (RNAi) in insects (19–22) and the interferon pathway in mammals (23, 24).

Virus ecology: the tip of the iceberg

Viruses are in the spotlight because they are a threat to public health and the economy. However, viruses are also major components of our ecosystems and play important homeostatic functions. Viruses are extremely abundant in the oceans, reaching more than a million viral particles in just one milliliter of North Atlantic seawater (25), and they are responsible for killing around 20% of the total oceanic microbial biomass per day (26). Viruses are probably the richest source of genetic diversity, as it is estimated that each bacterium hosts at least 10 different bacteriophages, and phage genes are so diverse that the majority of distinct genes in the biosphere belong to viral genomes (27-29).

Viruses affect ecosystems at all ecological scales (30). At the individual level, many viruses induce behavioral changes in their hosts, although it is generally hard to differentiate between virus manipulation and protective host responses (31-34). Viruses can drastically reduce the host population size too. For example, outbreaks of rabbit hemorrhagic disease virus destroyed populations of the European rabbit (*Oryctolagus cuniculus*) in Spain (35, 36), and this, in turn, initiated a trophic cascade that reduced the rabbit's natural predators, the Iberian Lynx (*Lynx pardinus*) and the Spanish Imperial eagle (*Aquila adalberti*) (37), thus altering the whole ecosystem.

We are still unaware of most viruses that exist in nature (38). We have only sampled a tiny portion of the virosphere because virus discovery has been anthropocentrically skewed towards disease-causing agents in humans, pets, and plants and animals of economic value. This is changing thanks to the unbiased virus discovery using metagenomic techniques (39, 40), which sequence the genetic material of entire communities of organisms and therefore provide essential information to study host-pathogen population dynamics in less studied organisms. To illustrate this, long-term surveillance of microbial structures in a freshwater lake demonstrated a positive virus-host pair correlation that repeated seasonally (41), suggesting that bacteria-phage dynamics shape the lake's microbial community. Considering that viruses are the most abundant Earth inhabitants and regulate our environment (27), we should study them more comprehensively as critical components of ecological communities that play a central role in maintaining ecosystem balance.

The ecosystem within the ecosystem: the replicative environment as a dominant driver of virus evolution

Virus emergence via cross-species transmission (also called host jumps) lies at the intersection of virus ecology, virus evolution, and public health. Three critical steps mediate host jumps: the virus gains the ability to infect cells from the new host species, it reaches efficient replication that ensures between-host transmission, and it spreads epidemically to establish itself as an endemic pathogen of the new host (42). Cross-species transmission therefore entails a substantial challenge that requires acquiring one or several mutations that grant adaptation to the new host. Thus, many host jumps remain transient spillover infections (43).

From the virus perspective, each host can be considered as its own ecosystem, where within-host virus evolution is affected by abiotic factors, like temperature, and biotic factors, such as immunity. In our experimental setting, we independently manipulated the host RNAi and cGAS-Sting pathways to determine how each shapes viral evolution. Interestingly, the changes in the local ecosystem (which I will refer to as the replicative environment), governed viral evolution in a manner largely independent of immune modulation. In **Chapter 2**, we detected a preexisting mutation in the DCV parental stock (Y95 of the VP3 capsid gene). This mutation exerted antagonistic pleiotropy, as it allowed DCV to replicate faster in *Drosophila* S2 cells, but almost abolished its replication in flies. As soon as the replicative environment changed to the *Drosophila* host *in vivo*, Y95 was selected against and replaced by almost all possible single nucleotide variants (SNVs) that could originate from codon 95, one of which was the Y to S reversion present in the wild-type virus. In **Chapter 3**, we passaged a DCV variant pre-adapted to wild-type flies, as well as CrPV in flies differing in their cGAS-Sting response. For DCV, we did not observe the fixation of any SNV, but CrPV adapted to the new ecosystem, becoming more virulent in flies by the fixation of a mutation in the transmembrane protein 2B (D29N) in multiple parallel lines. Below I will speculate on the apparent lack of selection by innate immunity in our experimental model.

The antigenic wave: how do viruses escape from the adaptive immune response?

Immunity can be grossly divided into two main branches with diffuse borders: innate and adaptive immunity. Innate immunity provides a generic response that is broadly shared, although highly diversified, across the tree of life; indeed, some components of the innate immune response are derived from ancient prokaryotic cells (44). Adaptive immunity is exclusive to jawed

vertebrates (45) and relies on somatic recombination of antigen receptors to randomly generate an immense repertoire of highly specific receptors.

The adaptive immune system elicits pathogen-specific responses by recognizing unique viral epitopes and preserving immune memory against reinfection (46). At the population level, immune memory drives viral diversification in antigenic space, a process modeled as an “antigenic wave” or a moving quasispecies pursued by immune cross-reactivity (47). Among the key forces driving this wave and favoring the emergence of escape variants is the antibody response, exemplified using SARS-CoV-2 as a case study. The Delta lineage rapidly displaced Alpha as the dominant lineage following vaccination campaigns, aided by escape mutations reducing sensitivity to vaccine-induced humoral responses and by higher replication efficiency (48-51). Later, the Omicron lineage emerged with unprecedented antigenic novelty, carrying over 30 mutations in the immunogenic spike protein targeted by vaccines (51, 52). Omicron exhibited markedly higher resistance to both vaccine-elicited and infection-derived antibodies compared to earlier variants (53-59). Different Omicron subvariants acquired convergent mutations in the spike receptor-binding domain, enhancing antibody escape (57, 60). These patterns highlight how antibody-mediated immunity strongly shapes RNA virus evolution, driving antigenic diversification and the emergence of escape variants in a process that is known as antigenic drift.

The presentation of viral antigens to cytotoxic CD8⁺ T lymphocytes (CTLs) represents another important factor of viral evolution. Infected cells process peptides derived from viral proteins, load them onto major histocompatibility complex class I (MHC-I) molecules, and present them on the cell surface for surveillance by CTLs. If these antigenic peptides are recognized by the T cell receptor (TCR), the infected cell is targeted for elimination (61). Consequently, viral peptides that are highly visible to T cells are under strong negative selection. A sheer abundance of escape strategies has been reported, such as the down-regulation of MHC-I surface expression and mutation of viral epitopes to impair processing, MHC-I binding, or TCR recognition (62). Well-documented examples of human viruses escaping CTLs responses are SARS-CoV-2 (63-65), human immunodeficiency virus type 1 (HIV-1) (66-68), and influenza A virus (69-72).

Does immunodeficiency affect virus evolution?

Next, I explore the mechanisms underlying the emergence of SARS-CoV-2 pathogenic lineages, given their direct relevance to the research questions

addressed in this thesis. In phylogenetic trees, long branches separate the Omicron lineage from earlier lineages and intermediate sequences are lacking (51, 53, 73, 74). In fact, the Omicron lineage is so different from its predecessors that it was proposed to represent a different viral serotype (75, 76). Three hypotheses have been formulated to explain the large phylogenetic gap between Omicron and other lineages (74): (i) the new Omicron lineage emerged in undersampled geographic regions; (ii) the virus evolved in an unknown animal reservoir and invaded the human population via zoonotic transmission; (iii) the Omicron lineage arose through within-host viral evolution in chronically infected, immunocompromised individuals. Although it is possible that all three hypotheses play a role in the emergence of Omicron, the last one has received the strongest experimental support, with several studies reporting escape mutations arising under such conditions (65, 77-81). Within-host virus evolution is favored by a high population size that promotes deterministic selection and supports the high multiplicity of infection needed for viral recombination (82), while preventing transmission bottlenecks that commonly occur between hosts (83). The rate of mutation events (defined as the gain or loss of a specific mutation) during SARS-CoV-2 evolution in chronically infected individuals has been estimated from eight longitudinal studies, some involving immunocompromised patients. On average, SARS-CoV-2 long-term within-host evolution produced four mutations per week, compared to roughly 0.5 mutations per week during between-host transmission (73). These findings suggest that chronic infections, particularly in immunocompromised hosts, are the most plausible origin of the Omicron lineage.

Along the same line, there are two observations from **Chapter 2** that are worth mentioning. Firstly, we detected a significant increase in polymorphisms in virus populations evolving in RNAi-deficient flies, an immunocompromised host, compared to wild-type and *Dicer-2*-overexpressing flies. This increase correlated with enhanced viral RNA synthesis, further supporting that viral replication within immunocompromised hosts is a source of genetic diversity and a driver of virus evolution. Indeed, a model that combines antigenic evolution with epidemiological data showed that the presence of immunocompromised hosts in the host population accelerates antigenic evolution (84). Secondly, we observed that the only evolved DCV population that retained the parental Y95 residue in the capsid gene *VP3* while recovering replicative capacity in flies was a lineage evolved in RNAi-deficient hosts. This lineage acquired a mutation in the vicinity of Y95 (P92T), in tight association with an additional substitution in *VP1* (N114S). Together, these mutations likely compensated for the *in vivo*

fitness cost attributable to Y95; in fact, the resulting double mutant was the fittest among all evolved variants. Although this occurred in only one out of five replicate lineages in the RNAi-deficient flies and in none of the ten lineages in immunocompetent hosts, the observation suggests that reduced immune pressure allows sequences with negative fitness effects to persist long enough in the population to serve as precursors of new mutational combinations that would otherwise be eliminated by purifying selection.

Evading the initial barriers: viral antagonism of innate immunity

Innate immunity is the first line of defense against invading pathogens. Innate antiviral immunity in mammals relies on cytosolic sensors such as RIG-I-like receptors and cGAS that detect molecules associated with infection, especially viral nucleic acids (85). Upon activation, these receptors trigger a signaling cascade leading to the production of pro-inflammatory cytokines known as interferons (IFNs). IFNs subsequently act in an autocrine and paracrine manner to upregulate the expression of interferon-stimulated genes that set up an antiviral state.

The capacity to antagonize the interferon pathway is widespread among different viruses (86), which do so through myriad mechanisms (87) that can be differentiated into three routes of action: inhibition of IFN production, inhibition of IFN signaling, and inhibition of the effector function of interferon stimulated genes (88). Numerous viruses encode dedicated IFN antagonists that are often, but not exclusively (89-91), non-structural proteins that have many different functions in virus biology (92). For instance, non-structural proteins of several picornaviruses (93), flaviviruses (94-96), and respiratory viruses (97), like influenza virus (98) and SARS-CoV-2 (99, 100), inhibit the IFN response.

As the first barrier to infection, the evasion of innate immunity is a key step mediating host switching and determining the pandemic potential of viruses. For instance, simian immunodeficiency virus has jumped to humans at least 13 independent times, but only a single HIV-1 lineage (group M, for “major”) established durably in the human population and caused a pandemic. The success of HIV-1(M) over the non-pandemic strains, in particular HIV-1 group O (for “outlier”) and HIV-2, has been linked to its ability to evade the activation of the cGAS sensor and the retrovirus restriction factor TRIM5α (101). Interestingly, the reduced activation of these pathways by HIV-1(M) seems to be mediated by mutations at two positions in the capsid protein. Confining viral nucleic acids within the capsid to limit recognition by cytosolic sensors

appears to be a particularly important immune evasion strategy of retroviruses. HIV-1 reverse transcription occurs inside the core, which allows the entrance of nucleotides through capsid pores and remains mostly intact during nuclear import, until the virus uncoats near the integration site (102-104).

SARS-CoV-2 also improved its ability to suppress the human innate response during the pandemic. The Alpha lineage increased the expression of known immune antagonists (the nucleocapsid (N), ORF6, and ORF9b) at both subgenomic RNA and protein levels, and regulated ORF9b activity by post-translational phosphorylation (105). Likewise, the Omicron lineage convergently evolved innate immune suppression by enhanced expression of N and ORF6 in the BA.4 and BA.5 subvariants compared to earlier Omicron subvariants (106). Overall, SARS-CoV-2 appears to rewire host innate immune responses by increasing the expression of viral immune antagonists, whose activity can be fine-tuned through phosphorylation (107). These processes are expected to be predominantly governed by mutations outside the *S* gene. More generally, adaptive mutations—regardless of the protein in which they occur—tend to be enriched at surface-exposed residues, where structural constraints are lower and protein-protein interactions are concentrated (107, 108). Thus, adaptive evolution of viral proteins involved in innate immune antagonism—often outside canonical entry factors—represents a key determinant of pandemic potential.

Evolutionary constraints of innate antagonists

In light of the preceding discussion, one must ask: why was immune-mediated adaptation limited in our experimental models in **Chapters 2** and **3**? Below, I speculate about possible reasons.

Innate immunity recognizes generic traits that are associated with basic functions of pathogens (109) and, as such, they are highly conserved structures. This may limit the virus's ability to escape innate immunity as opposed to adaptive immunity, which can be avoided by altering targeted sequences that are typically pathogen-specific. These features (e.g., the capsid-specific structure) can escape adaptive immune recognition by simply acquiring one or few point mutations, as seen in rapidly evolving HIV-1 epitopes exposed to CTL responses (110). Alternatively, molecules recognized by innate immunity, like double-stranded RNA, are essential for all RNA viruses, and their recognition cannot be prevented by directly mutating the viral genome. Instead, RNA viruses have evolved viral proteins with IFN antagonist activity, as introduced earlier (111), or viral suppressors of RNAi (112, 113) to combat innate immunity.

Considering the limited size and compact nature of RNA virus genomes, viruses have developed sophisticated mechanisms to express IFN antagonists. For instance, viruses from the *Paramyxoviridae* family lack a dedicated antagonist in their genome. Instead, they evolved an mRNA editing mechanism in which the viral polymerase stutters in homopolymeric regions and inserts one or two extra guanosines in their *P* gene, changing the mRNA reading frame (114, 115). Paramyxoviruses generate in this way several new proteins that suppress the IFN response with different mechanisms (116, 117).

Now, imagine a mutation arising in one of these IFN antagonists that confers improved immune evasion. Such a mutation is linked through pleiotropic interactions with the *P* gene, as it is in the same genomic region encoding separate viral proteins. The selection regime and potential fixation of this mutation depend on the compound fitness effect of the individual proteins originating from the *P* gene. The *P* protein is essential for virus replication, as it is a cofactor of the viral polymerase (118). It is likely that, irrespective of this mutation being beneficial for interferon antagonism, it worsens virus replication and creates a negative net effect that would purify this mutation from the viral population. This illustrates how proteins with a wider variety of biochemical functions are more restricted to change.

The complexity of viral genomes hampers the fixation of adaptive mutations, particularly through antagonistic epistatic and pleiotropic effects that are dominant in RNA viruses (119). Pleiotropy occurs when changes in one gene affect several phenotypic traits. Pleiotropism and epistasis—the dependence of a mutation’s effect on the presence of other mutations—create rugged fitness landscapes that can constrain evolutionary trajectories (42). Both phenomena have been modelled to inversely correlate with genome robustness, i.e. the tolerance to mutagenesis (120). It has been proposed that antagonistic pleiotropism reduced positive selection in *D. melanogaster* (121) and hindered the fixation of mutations in the influenza A virus host receptor-binding *haemagglutinin* gene (122). In addition, pleiotropism increased parallel selection in flies and reduced sequence divergence of pleiotropic genes (123). Lastly, studying the fitness effect of combining beneficial mutations in RNA viruses showed that the addition of a second mutation produced either detrimental effects or diminishing-returns epistasis, the reduced fitness gain of a beneficial mutation as an organism becomes fitter (124, 125). Diminishing-returns epistasis has been proposed to slow adaptation rates (126).

Taken these observations together, it is likely that the evolution of multifunctional proteins of RNA viruses is limited by their complex genetic architecture.

The immune-mediated evolutionary signatures detected in our experimental models were limited to two observations: an increase in the number of polymorphisms in DCV evolved in *Dicer-2* knockout flies (**Chapter 2**), and the incapacity of the beneficial 2B D29N allele to reach fixation in CrPV evolved in cGAS-Sting stimulated flies (**Chapter 3**). Both findings remain intriguing, and the underlying mechanisms can at present only be speculated upon. I believe the first result is explained by a reduction in purifying selection of viruses replicating in the absence of the RNAi pathway, which allowed virus population to further explore the sequence space. As for the latter result, I argue that the reduction in viral population size due to the activation of the cGAS-Sting pathway caused a less deterministic evolution restricting the effect of positive selection.

Beyond these two observations, we obtained limited evidence of innate immune-mediated evolution in **Chapters 2 and 3**. What could be the reason? One can hypothesize that the ten passages in our fly model, corresponding to a 30-day adaptation period, together with the capped number of virus particles produced by the relatively small host population size (30-40 flies per passage), may not have generated sufficient genetic diversity to promote adaptive changes in innate immune antagonists whose evolution is constrained by pleiotropy and epistasis. It is also possible that a less stringent selective pressure could be masked by stronger forces shaping viral evolution, such as adaptation to the replicative environment (adaptation to *in vivo* replication in **Chapter 2** and cross-species transmission in **Chapter 3**).

Adaptation in structural and non-structural functions

During viral adaptation in our experimental model, we observed evolution of both structural (**Chapter 2**) and non-structural proteins (**Chapter 3**). Below, I elaborate on the role of these different proteins in the context of host switching and adaptation to novel environments. Structural proteins are pivotal determinants of viral success during early phases of infection. Mutations in the structural proteins that bestow higher binding affinity to host receptors are key facilitators of cell entry in new hosts (127). For example, mutations accumulated in the receptor-binding glycoproteins of SARS-CoV-2 (128-130), avian influenza A virus (131), and Ebola virus (132-134) as part of their adaptation to humans. In **Chapter 2**, we observed DCV alleles at a specific position of the capsid protein VP3 mediating adaptation to either

cell culture or *in vivo* replication. However, it seems unlikely that enhanced cell entry was responsible for the observed adaptation, as the parental DCV variant replicated very efficiently in S2 cells, where the mechanism of cell entry is likely the same as *in vivo*. Therefore, we conjectured that the mechanism driving the selection of the new variants upon *in vivo* replication could be recognition by some immune receptor participating in the systemic immune response.

Despite structural proteins being crucial during the first steps of infection, non-structural proteins execute a vast array of molecular functions in infected cells that are essential for virus biology and host adaptation. For instance, the influenza A virus non-structural protein NS1 shows species-specific adaptive mutations that are allegedly involved in host switching (135, 136). Likewise, a deletion in the SARS-CoV-2 NSP6 protein, a protein that is involved in membrane remodeling and replication complex formation, was selected on multiple independent occasions during the evolution of the Omicron lineage (137). Lastly, mutations in SARS-CoV-2 NSP4, NSP8, NSP9 and the spike protein were observed during viral adaptation to mice, of which the mutations in the non-structural proteins were required for efficient mouse infection, whereas the spike mutation alone produced a mild increase in virulence (138). These examples align with our results from **Chapter 3**, where CrPV adaptation to a novel host led to the recurrent fixation of the D29N mutation in the non-structural transmembrane protein 2B, conferring markedly higher virulence in flies. However, the mechanism by which 2B mediates host switching, and the specific role of the D29N allele, remain to be elucidated. Taken together, these cases highlight how non-structural proteins are decisive for viral adaptation. They shape host interactions, fine-tune immune evasion, and reorganize the cellular machinery to create an environment where the virus can efficiently replicate and spread. Their versatility makes them key determinants of host range and, ultimately, the epidemic potential of viruses.

Measuring the fitness effect of adaptive mutations

Evaluating the impact of adaptive changes requires the quantification of their contribution to viral fitness, thereby elucidating the dynamics of alleles under selection. The fitness advantage (or disadvantage) granted by an adaptive mutation relative to a wild-type genotype can be expressed as a selection coefficient, s (139). It is a key coefficient in population genetics which is used especially in experimental evolution studies (140). We estimated that s of the CrPV D29N mutant was on average larger than 0.2, calculated using two independent approaches: from the increase in allele frequency over time and the difference

in RNA replication kinetics of mutant versus wild-type virus (**Chapter 3**). An s larger than 0.2 indicates a > 20% increase in reproductive success per virus generation, which increases the mutant population size quickly over the wild-type with a generation time of around 10 hours. This led me to assess the circumstances under which s values greater than 0.2 occur, and how commonly they are observed. In general, viral mutations are suggested to have a higher s than those from other groups of microorganisms (141). Still, the s we reported is above the s distribution that was inferred from phylogenetic data for HIV-1 *env* and influenza A *haemagglutinin* genes (142), which are known to be under strong selective pressure. Such an s is not unprecedented in viral populations though, as illustrated by a mutation in the receptor-binding domain of the SARS-CoV-2 spike protein that increased its binding affinity for the cell entry receptor (143). Strikingly, the s value of this mutation was dependent on the vaccination status of the host population, as it was estimated to be 0.26 in an unvaccinated population and 0.41 in a completely vaccinated population (144). Similarly, HIV-1 escape mutants of CTLs had an estimated s varying between 0.2 to 0.4 during acute infection, but only between 0.0024 to 0.15 during chronic infection (110). It is likely that the combination of large population sizes—such as those produced during acute infection—and strong selective pressures, including adaptive immunity or abrupt changes in the replicative environment, drives a high s .

What is the right experimental model?

An experimental model should accurately reflect the responses to the trait of interest that would be observed in the corresponding natural system, while minimizing removable sources of variation in this trait. However, we could wonder if, after such simplification, a model would still mimic natural conditions and remain relevant. In Manfred Eigen's words:

"A theory has only the alternative of being right or wrong. A model has a third possibility: it may be right, but irrelevant" (145).

Drosophila is a powerful experimental model for studying virus-host interactions in insects. Its genetic tractability, well-characterized innate immune pathways, and easy rearing enable tightly controlled experiments. In **Chapters 2** and **3**, we used this model to reduce transmission bottlenecks by inoculating high viral loads, to minimize host population fluctuations caused by predation, migration, or resource availability, and to limit the effects of genetic background, temperature variation, persistent infections, and

pathogen interactions arising from coinfections with multiple microorganisms. In this simplified environment, we hypothesized that variations in the host immune pressure would translate into viral adaptation to either the relaxation of selective constraints on the immunocompromised hosts (e.g., viral suppressors of immunity being able to mutate due to the lack of selection) or to the overactivated immune response (e.g., the selection of escape mutants, higher activity of immune suppressors).

In nature, viruses infecting flies face a mosaic of pressures that have a combinatorial effect on virus evolution. Almost one third of wild-caught flies have wounds (146), which in turn may affect susceptibility to infection and their immune response. Moreover, more than 30% of wild flies are infected with RNA viruses (147), with some of the persistent infections impacting basic life traits (148, 149). Exposure to predators can change fly behavior too, as seen in the reduction of oviposition as a defense to parasitoid wasps (150), which likely impacts host population size. However, the way these factors affect virus evolution remains to be established.

The simplicity that makes our model strong for isolating the influence of innate immunity on viral evolution inevitably limits its resemblance to the dynamics of natural evolution. Although our results do not demonstrate innate immune-driven evolution, this absence is itself informative, as it highlights how repeatably viruses adapt to controlled environmental cues. Furthermore, the findings from **Chapter 3** reveal a previously unrecognized molecular player involved in cross-species transmission. Successful host jumps may require the virus to adapt to the cellular environment of the new species for the effective formation of replication organelles, a process that could be targeted for antiviral development. Taken together, these insights underscore that while simplified models cannot fully capture the complexity of natural evolution, they remain invaluable for uncovering mechanistic constraints that shape RNA virus evolution.

Using the fly genetic toolbox to identify host factors for infection

Our experimental model allow careful evaluation of host genetic determinants of virus infection. In **Chapter 4**, we exploited the *Drosophila* genetic toolbox to silence RNA-modifying enzymes to assess their role in RNA virus replication. Unfortunately, knockdown efficiency obtained with the RNAi libraries was limited for our set of genes, and the screen did not yield the expected sensitivity and consistency. Despite RNAi screens being widely and successfully

used (151-154), constitutive gene silencing is generally inefficient, achieving an average knockdown efficiency of 48.6% and 38.1% compared to controls in 75 RNAi lines from the same libraries that we used (KK and GD, respectively) (155). This modest reduction in expression levels masks subtle phenotypes and may explain the almost 40% false negative rate reported for the GD library (151). This is instructive though, as it signals the difficulty of achieving robust gene knockdown *in vivo* and underlines that models must be validated and benchmarked thoroughly.

While RNAi screens remain popular due to their scalability and ease of use, they often suffer from incomplete silencing and off-target effects, leading to inconsistent phenotypes. An alternative to RNAi for large-scale screens is the use of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9 based approaches, which offer high specificity and can induce a complete loss-of-function (156, 157). The CRISPR-Cas9 knockout system uses a Cas9 endonuclease guided by a single-guide RNA to introduce targeted double-strand breaks in DNA, resulting in gene disruption through error-prone repair pathways (158). To screen for essential genes that cannot be knocked out, the transcriptionally repressive CRISPR interference can be used instead, where Cas9 has been catalytically inactivated and induces knockdown by sterically blocking the polymerase or by direct coupling to a transcriptional repressor (159). Overall, CRISPR-Cas9 based systems provide powerful alternatives for conducting screens that could outperform RNAi (160).

Concluding remarks

Our experimental evolution work produced unexpected results that were largely unrelated to our original research question concerning the influence of innate immunity on virus adaptability. However, we obtained clear evidence of a directed and highly repeatable evolutionary path that led to remarkable fitness gains within just a few generations, representing an adaptive viral response to novel environments. These observations suggest that natural selection defines the sequence space available for adaptation. Thus, parallel mutations may reflect the reduced number of molecular solutions accessible to the architecturally constrained genomes of RNA viruses, or perhaps simply indicate that the fittest variants will always win when deterministic evolution is favored. Regardless, my research shows that viral evolution is not random but follows a predictable trajectory shaped by biological constraints. Importantly, we identified functional hotspots for virus adaptation through repeatable evolutionary outcomes. These are not restricted to classical candidates

(e.g., entry proteins), but emphasize the role of viral non-structural proteins as determinants of cross-species transmission. Our findings highlight how limited our understanding of virus-host interactions remains, stress how challenging it is to anticipate all variables affecting evolution, and illustrate that even well-founded hypotheses can yield unforeseen results in controlled experimental systems.

References

1. WHO. 14.9 million excess deaths were associated with the COVID-19 pandemic in 2020 and 2021. World Health Organization. 2022. Available from: 14.9 million excess deaths associated with the COVID-19 pandemic in 2020 and 2021.
2. Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. *Nature*. 2023;613(7942):130-7, doi: 10.1038/s41586-022-05522-2.
3. Adam D. The pandemic's true death toll: millions more than official counts. *Nature*. 2022;601(7893):312-5, doi: 10.1038/d41586-022-00104-8.
4. Muggenthaler P, Schroth J, Sun Y. The heterogeneous economic impact of the pandemic across euro area countries. European Central Bank. 2021. Available from: The heterogeneous economic impact of the pandemic across euro area countries.
5. Kaye AD, Okeagu CN, Pham AD, Silva RA, Hurley JJ, Arron BL, et al. Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives. *Best Pract Res Clin Anaesthesiol*. 2021;35(3):293-306, doi: 10.1016/j.bpa.2020.11.009.
6. UNAIDS. Global HIV Factsheet – Latest global and regional statistics on the status of the AIDS epidemic. 2025. Available from: Fact sheet - Latest global and regional statistics on the status of the AIDS epidemic.
7. Preliminary Estimated Flu Disease Burden 2024-2025 Flu Season [press release]. Centers for Disease Control and Prevention. 2025. Available from: Preliminary Estimated Flu Disease Burden 2024-2025 Flu Season | Flu Burden | CDC
8. Temin HM, Mizutani S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature*. 1970;226(5252):1211-3, doi: 10.1038/2261211a0.
9. Baltimore D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature*. 1970;226(5252):1209-11, doi: 10.1038/2261209a0.
10. Yarchoan R, Mitsuya H, Myers CE, Broder S. Clinical pharmacology of 3'-azido-2',3'-dideoxythymidine (zidovudine) and related dideoxynucleosides. *N Engl J Med*. 1989;321(11):726-38, doi: 10.1056/NEJM198909143211106.
11. Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, et al. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science*. 2012;336(6088):1534-41, doi: 10.1126/science.1213362.
12. Iranzo J, Puigbo P, Lobkovsky AE, Wolf YI, Koonin EV. Inevitability of genetic parasites. *Genome Biol Evol*. 2016;8(9):2856-69, doi: 10.1093/gbe/evw193.
13. Hershey AD, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol*. 1952;36(1):39-56, doi: 10.1085/jgp.36.1.39.
14. Brenner S, Jacob F, Meselson M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature*. 1961;190:576-81, doi: 10.1038/190576a0.
15. Tenthorey JL, Emerman M, Malik HS. Evolutionary landscapes of host-virus arms races. *Annu Rev Immunol*. 2022;40:271-94, doi: 10.1146/annurev-immunol-072621-084422.
16. Van Valen L. A new evolutionary law. *Evol Theory* 1973;1:1-30.
17. Palmer WJ, Jiggins FM. Comparative genomics reveals the origins and diversity of arthropod immune systems. *Mol Biol Evol*. 2015;32(8):2111-29, doi: 10.1093/molbev/msv093.

18. van der Lee R, Wiel L, van Dam TJP, Huynen MA. Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. *Nucleic Acids Res.* 2017;45(18):10634-48, doi: 10.1093/nar/gkx704.
19. Obbard DJ, Gordon KH, Buck AH, Jiggins FM. The evolution of RNAi as a defence against viruses and transposable elements. *Philos Trans R Soc Lond B Biol Sci.* 2009;364(1513): 99-115, doi: 10.1098/rstb.2008.0168.
20. Obbard DJ, Jiggins FM, Bradshaw NJ, Little TJ. Recent and recurrent selective sweeps of the antiviral RNAi gene *Argonaute-2* in three species of *Drosophila*. *Mol Biol Evol.* 2011;28(2):1043-56, doi: 10.1093/molbev/msq280.
21. Obbard DJ, Jiggins FM, Halligan DL, Little TJ. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol.* 2006;16(6):580-5, doi: 10.1016/j.cub.2006.01.065.
22. Palmer WH, Hadfield JD, Obbard DJ. RNA-interference pathways display high rates of adaptive protein evolution in multiple invertebrates. *Genetics.* 2018;208(4):1585-99, doi: 10.1534/genetics.117.300567.
23. Manry J, Laval G, Patin E, Fornarino S, Itan Y, Fumagalli M, et al. Evolutionary genetic dissection of human interferons. *J Exp Med.* 2011;208(13):2747-59, doi: 10.1084/jem.20111680.
24. Xu L, Yang L, Liu W. Distinct evolution process among type I interferon in mammals. *Protein Cell.* 2013;4(5):383-92, doi: 10.1007/s13238-013-3021-1.
25. Parada V, Sintes E, van Aken HM, Weinbauer MG, Herndl GJ. Viral abundance, decay, and diversity in the meso- and bathypelagic waters of the north atlantic. *Appl Environ Microbiol.* 2007;73(14):4429-38, doi: 10.1128/AEM.00029-07.
26. Suttle CA. Marine viruses--major players in the global ecosystem. *Nat Rev Microbiol.* 2007;5(10):801-12, doi: 10.1038/nrmicro1750.
27. Koonin EV, Dolja VV. A virocentric perspective on the evolution of life. *Curr Opin Virol.* 2013;3(5):546-57, doi: 10.1016/j.coviro.2013.06.008.
28. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, et al. Origins of highly mosaic mycobacteriophage genomes. *Cell.* 2003;113(2):171-82, doi: 10.1016/s0092-8674(03)00233-2.
29. Rohwer F. Global phage diversity. *Cell.* 2003;113(2):141, doi: 10.1016/s0092-8674(03)00276-9.
30. French RK, Holmes EC. An ecosystems perspective on virus evolution and emergence. *Trends Microbiol.* 2020;28(3):165-75, doi: 10.1016/j.tim.2019.10.010.
31. Varaldi J, Petit S, Bouletreau M, Fleury F. The virus infecting the parasitoid *Leptopilina boulardi* exerts a specific action on superparasitism behaviour. *Parasitology.* 2006;132(Pt 6):747-56, doi: 10.1017/S0031182006009930.
32. Stroeymeyt N, Grasse AV, Crespi A, Mersch DP, Cremer S, Keller L. Social network plasticity decreases disease transmission in a eusocial insect. *Science.* 2018;362(6417):941-5, doi: 10.1126/science.aat4793.
33. Tallon AK, Lorenzo MG, Moreira LA, Martinez Villegas LE, Hill SR, Ignell R. Dengue infection modulates locomotion and host seeking in *Aedes aegypti*. *PLoS Negl Trop Dis.* 2020;14(9):e0008531, doi: 10.1371/journal.pntd.0008531.
34. Wei Xiang BW, Saron WAA, Stewart JC, Hain A, Walvekar V, Misse D, et al. Dengue virus infection modifies mosquito blood-feeding behavior to increase transmission to the host. *Proc Natl Acad Sci U S A.* 2022;119(3):e2117589119, doi: 10.1073/pnas.2117589119.

35. Villafuerte R, Calvete C, Blanco J, Lucientes J. Incidence of viral hemorrhagic disease in wild rabbit populations in Spain. *Mammalia*. 1995;59(4):651-60, doi: 10.1136/vr.150.25.776.
36. Moleon M, Almaraz P, Sanchez-Zapata JA. An emerging infectious disease triggering large-scale hyperpredation. *PLoS One*. 2008;3(6):e2307, doi: 10.1371/journal.pone.0002307.
37. Monterroso P, Garrote G, Serronha A, Santos E, Delibes-Mateos M, Abrantes J, et al. Disease-mediated bottom-up regulation: An emergent virus affects a keystone prey, and alters the dynamics of trophic webs. *Sci Rep*. 2016;6:36072, doi: 10.1038/srep36072.
38. Zhang YZ, Chen YM, Wang W, Qin XC, Holmes EC. Expanding the RNA virosphere by unbiased metagenomics. *Annu Rev Virol*. 2019;6(1):119-39, doi: 10.1146/annurev-virology-092818-015851.
39. Zhang YZ, Shi M, Holmes EC. Using metagenomics to characterize an expanding virosphere. *Cell*. 2018;172(6):1168-72, doi: 10.1016/j.cell.2018.02.043.
40. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, et al. Redefining the invertebrate RNA virosphere. *Nature*. 2016;540(7634):539-43, doi: 10.1038/nature20167.
41. Zhou Z, Tran PQ, Martin C, Rohwer RR, Baker BJ, McMahon KD, et al. Unravelling viral ecology and evolution over 20 years in a freshwater lake. *Nat Microbiol*. 2025;10(1):231-45, doi: 10.1038/s41564-024-01876-7.
42. Dennehy JJ. Evolutionary ecology of virus emergence. *Ann N Y Acad Sci*. 2017;1389(1):124-46, doi: 10.1111/nyas.13304.
43. Geoghegan JL, Holmes EC. Predicting virus emergence amid evolutionary noise. *Open Biol*. 2017;7(10):170189, doi: 10.1098/rsob.170189.
44. Aravind L, Nicastro GG, Iyer LM, Burroughs AM. The prokaryotic roots of eukaryotic immune systems. *Annu Rev Genet*. 2024;58(1):365-89, doi: 10.1146/annurev-genet-111523-102448.
45. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet*. 2010;11(1):47-59, doi: 10.1038/nrg2703.
46. Lam N, Lee Y, Farber DL. A guide to adaptive immune memory. *Nat Rev Immunol*. 2024;24(11):810-29, doi: 10.1038/s41577-024-01040-6.
47. Marchi J, Lassig M, Walczak AM, Mora T. Antigenic waves of virus-immune coevolution. *Proc Natl Acad Sci U S A*. 2021;118(27):e2103398118, doi: 10.1073/pnas.2103398118.
48. Davis C, Logan N, Tyson G, Orton R, Harvey WT, Perkins JS, et al. Reduced neutralisation of the Delta (B.1.617.2) SARS-CoV-2 variant of concern following vaccination. *PLoS Pathog*. 2021;17(12):e1010022, doi: 10.1371/journal.ppat.1010022.
49. Garcia-Beltran WF, Lam EC, St Denis K, Nitido AD, Garcia ZH, Hauser BM, et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell*. 2021;184(9):2523, doi: 10.1016/j.cell.2021.04.006.
50. Mlcochova P, Kemp SA, Dhar MS, Papa G, Meng B, Ferreira I, et al. SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature*. 2021;599(7883):114-9, doi: 10.1038/s41586-021-03944-y.
51. Carabelli AM, Peacock TP, Thorne LG, Harvey WT, Hughes J, Consortium C-GU, et al. SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat Rev Microbiol*. 2023;21(3):162-77, doi: 10.1038/s41579-022-00841-7.
52. Syed AM, Ciling A, Taha TY, Chen IP, Khalid MM, Sreekumar B, et al. Omicron mutations enhance infectivity and reduce antibody neutralization of SARS-CoV-2 virus-like particles. *Proc Natl Acad Sci U S A*. 2022;119(31):e2200592119, doi: 10.1073/pnas.2200592119.

53. Tegally H, Moir M, Everatt J, Giovanetti M, Scheepers C, Wilkinson E, et al. Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat Med.* 2022;28(9):1785-90, doi: 10.1038/s41591-022-01911-2.
54. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature.* 2022;603(7902):679-86, doi: 10.1038/s41586-022-04411-y.
55. Martin DP, Lytras S, Lucaci AG, Maier W, Gruning B, Shank SD, et al. Selection analysis identifies clusters of unusual mutational changes in Omicron lineage BA.1 that likely impact spike function. *Mol Biol Evol.* 2022;39(4):msac06, doi: 10.1093/molbev/msac061.
56. Andrews N, Stowe J, Kirsebom F, Toffa S, Rickeard T, Gallagher E, et al. Covid-19 vaccine effectiveness against the Omicron (B.1.1.529) Variant. *N Engl J Med.* 2022;386(16):1532-46, doi: 10.1056/NEJMoa2119451.
57. Cao Y, Jian F, Wang J, Yu Y, Song W, Yisimayi A, et al. Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature.* 2023;614(7948):521-9, doi: 10.1038/s41586-022-05644-7.
58. Cele S, Jackson L, Khoury DS, Khan K, Moyo-Gwete T, Tegally H, et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature.* 2022;602(7898):654-6, doi: 10.1038/s41586-021-04387-1.
59. Newman J, Thakur N, Peacock TP, Bialy D, Elrefaey AME, Bogaardt C, et al. Neutralizing antibody activity against 21 SARS-CoV-2 variants in older adults vaccinated with BNT162b2. *Nat Microbiol.* 2022;7(8):1180-8, doi: 10.1038/s41564-022-01163-3.
60. Ito J, Suzuki R, Uriu K, Itakura Y, Zahradnik J, Kimura KT, et al. Convergent evolution of SARS-CoV-2 Omicron subvariants leading to the emergence of BQ.1.1 variant. *Nat Commun.* 2023;14(1):2671, doi: 10.1038/s41467-023-38188-z.
61. Harari A, Dutoit V, Cellerai C, Bart PA, Du Pasquier RA, Pantaleo G. Functional signatures of protective antiviral T-cell immunity in human virus infections. *Immunol Rev.* 2006;211:236-54, doi: 10.1111/j.0105-2896.2006.00395.x.
62. Wu Y, Sun Z, Xia L, Tian P, Jiao L, Li Y, et al. MHC-I pathway disruption by viruses: insights into immune evasion and vaccine design for animals. *Front Immunol.* 2025;16:1540159, doi: 10.3389/fimmu.2025.1540159.
63. Zhang Y, Chen Y, Li Y, Huang F, Luo B, Yuan Y, et al. The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-Iota. *Proc Natl Acad Sci U S A.* 2021;118(23):e2024202118, doi: 10.1073/pnas.2024202118.
64. Moriyama M, Lucas C, Monteiro VS, Yale S-C-GSI, Iwasaki A. Enhanced inhibition of MHC-I expression by SARS-CoV-2 Omicron subvariants. *Proc Natl Acad Sci U S A.* 2023;120(16):e2221652120, doi: 10.1073/pnas.2221652120.
65. Stanevich OV, Alekseeva EI, Sergeeva M, Fadeev AV, Komissarova KS, Ivanova AA, et al. SARS-CoV-2 escape from cytotoxic T cells during long-term COVID-19. *Nat Commun.* 2023;14(1):149, doi: 10.1038/s41467-022-34033-x.
66. Goulder PJ, Walker BD. HIV and HLA class I: an evolving relationship. *Immunity.* 2012;37(3):426-40, doi: 10.1016/j.immuni.2012.09.005.
67. Neuner-Jehle N, Zeeb M, Thorball CW, Fellay J, Metzner KJ, Frischknecht P, et al. Using viral diversity to identify HIV-1 variants under HLA-dependent selection in a systematic viral genome-wide screen. *PLoS Pathog.* 2024;20(8):e1012385, doi: 10.1371/journal.ppat.1012385.
68. Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, Gaschen B, et al. HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS Pathog.* 2009;5(5):e1000414, doi: 10.1371/journal.ppat.1000414.

69. Boon AC, de Mutsert G, Graus YM, Fouchier RA, Sintnicolaas K, Osterhaus AD, et al. Sequence variation in a newly identified HLA-B35-restricted epitope in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. *J Virol.* 2002;76(5):2567-72, doi: 10.1128/jvi.76.5.2567-2572.2002.
70. Koutsakos M, McWilliam HEG, Aktepe TE, Fritzlar S, Illing PT, Mifsud NA, et al. Downregulation of MHC class I expression by influenza A and B viruses. *Front Immunol.* 2019;10:1158, doi: 10.3389/fimmu.2019.01158.
71. Valkenburg SA, Quinones-Parra S, Gras S, Komadina N, McVernon J, Wang Z, et al. Acute emergence and reversion of influenza A virus quasispecies within CD8+ T cell antigenic peptides. *Nat Commun.* 2013;4:2663, doi: 10.1038/ncomms3663.
72. van de Sandt CE, Kreijtz JH, Rimmelzwaan GF. Evasion of influenza A viruses from innate and adaptive immune responses. *Viruses.* 2012;4(9):1438-76, doi: 10.3390/v4091438.
73. Hill V, Du Plessis L, Peacock TP, Aggarwal D, Colquhoun R, Carabelli AM, et al. The origins and molecular evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. *Virus Evol.* 2022;8(2):veac080, doi: 10.1093/ve/veac080.
74. Mallapaty S. Where did Omicron come from? Three key theories. *Nature.* 2022;602(7895):26-8, doi: 10.1038/d41586-022-00215-2.
75. van der Straten K, Guerra D, van Gils MJ, Bontjer I, Caniels TG, van Willigen HDG, et al. Antigenic cartography using sera from sequence-confirmed SARS-CoV-2 variants of concern infections reveals antigenic divergence of Omicron. *Immunity.* 2022;55(9):1725-31 e4, doi: 10.1016/j.immuni.2022.07.018.
76. Simon-Loriere E, Schwartz O. Towards SARS-CoV-2 serotypes? *Nat Rev Microbiol.* 2022;20(4):187-8, doi: 10.1038/s41579-022-00708-x.
77. Weigang S, Fuchs J, Zimmer G, Schnepf D, Kern L, Beer J, et al. Within-host evolution of SARS-CoV-2 in an immunosuppressed COVID-19 patient as a source of immune escape variants. *Nat Commun.* 2021;12(1):6405, doi: 10.1038/s41467-021-26602-3.
78. Lustig G, Ganga Y, Rodel HE, Tegally H, Khairallah A, Jackson L, et al. SARS-CoV-2 infection in immunosuppression evolves sub-lineages which independently accumulate neutralization escape mutations. *Virus Evol.* 2024;10(1):vead075, doi: 10.1093/ve/vead075.
79. Halfmann PJ, Minor NR, Haddock Iii LA, Maddox R, Moreno GK, Braun KM, et al. Evolution of a globally unique SARS-CoV-2 Spike E484T monoclonal antibody escape mutation in a persistently infected, immunocompromised individual. *Virus Evol.* 2023;9(2):veac104, doi: 10.1093/ve/veac104.
80. Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, et al. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *N Engl J Med.* 2020;383(23):2291-3, doi: 10.1056/NEJMc2031364.
81. Clark SA, Clark LE, Pan J, Coscia A, McKay LGA, Shankar S, et al. SARS-CoV-2 evolution in an immunocompromised host reveals shared neutralization escape mechanisms. *Cell.* 2021;184(10):2605-17 e18, doi: 10.1016/j.cell.2021.03.027.
82. Jackson B, Boni MF, Bull MJ, Collieran A, Colquhoun RM, Darby AC, et al. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. *Cell.* 2021;184(20):5179-88 e8, doi: 10.1016/j.cell.2021.08.014.
83. Lythgoe KA, Hall M, Ferretti L, de Cesare M, MacIntyre-Cockett G, Trebes A, et al. SARS-CoV-2 within-host diversity and transmission. *Science.* 2021;372(6539), doi: 10.1126/science.abg0821.
84. Kumata R, Sasaki A. Antigenic escape is accelerated by the presence of immunocompromised hosts. *Proc Biol Sci.* 2022;289(1986):20221437, doi: 10.1098/rspb.2022.1437.

85. Wu J, Chen ZJ. Innate immune sensing and signaling of cytosolic nucleic acids. *Annu Rev Immunol.* 2014;32:461-88, doi: 10.1146/annurev-immunol-032713-120156.
86. Versteeg GA, Garcia-Sastre A. Viral tricks to grid-lock the type I interferon system. *Curr Opin Microbiol.* 2010;13(4):508-16, doi: 10.1016/j.mib.2010.05.009.
87. Garcia-Sastre A. Ten Strategies of Interferon Evasion by Viruses. *Cell Host Microbe.* 2017;22(2):176-84, doi: 10.1016/j.chom.2017.07.012.
88. Hoffmann HH, Schneider WM, Rice CM. Interferons and viruses: an evolutionary arms race of molecular interactions. *Trends Immunol.* 2015;36(3):124-38, doi: 10.1016/j.it.2015.01.004.
89. Weber-Gerlach M, Weber F. To conquer the host, influenza virus is packing it in: interferon-antagonistic strategies beyond NS1. *J Virol.* 2016;90(19):8389-94, doi: 10.1128/JVI.00041-16.
90. Oh SJ, Shin OS. SARS-CoV-2 nucleocapsid protein targets RIG-I-like receptor pathways to inhibit the induction of the interferon response. *Cells.* 2021;10(3):530, doi: 10.3390/cells10030530.
91. Lei X, Dong X, Ma R, Wang W, Xiao X, Tian Z, et al. Activation and evasion of type I interferon responses by SARS-CoV-2. *Nat Commun.* 2020;11(1):3810, doi: 10.1038/s41467-020-17665-9.
92. Murray CL, Jones CT, Rice CM. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat Rev Microbiol.* 2008;6(9):699-708, doi: 10.1038/nrmicro1928.
93. Wang YN, Ma LN, Stipkovits L, Szathmary S, Li XR, Liu YS. The strategy of picornavirus evading host antiviral responses: non-structural proteins suppress the production of IFNs. *Frontiers in Microbiology.* 2018;9:2943, doi: 10.3389/fmicb.2018.02943.
94. Thurmond S, Wang B, Song J, Hai R. Suppression of type I interferon signaling by flavivirus NS5. *Viruses.* 2018;10(12):712, doi: 10.3390/v10120712.
95. Hee JR, Cheng D, Chen YH, Wang SH, Chao CH, Huang SW, et al. A non-structural protein 1 substitution of dengue virus enhances viral replication by interfering with the antiviral signaling pathway. *J Biomed Sci.* 2025;32(1):25, doi: 10.1186/s12929-024-01116-4.
96. Liu WJ, Wang XJ, Clark DC, Lobigs M, Hall RA, Khromykh AA. A single amino acid substitution in the West Nile virus nonstructural protein NS2A disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. *J Virol.* 2006;80(5):2396-404, doi: 10.1128/JVI.80.5.2396-2404.2006.
97. Piret J, Boivin G. Viral interference between respiratory viruses. *Emerg Infect Dis.* 2022;28(2):273-81, doi: 10.3201/eid2802.211727.
98. Marc D. Influenza virus non-structural protein NS1: interferon antagonism and beyond. *J Gen Virol.* 2014;95(Pt 12):2594-611, doi: 10.1099/vir.0.069542-0.
99. Kumar A, Ishida R, Strilets T, Cole J, Lopez-Orozco J, Fayad N, et al. SARS-CoV-2 Nonstructural Protein 1 Inhibits the Interferon Response by Causing Depletion of Key Host Signaling Factors. *J Virol.* 2021;95(13):e0026621, doi: 10.1128/JVI.00266-21.
100. Lee JH, Koepke L, Kirchhoff F, Sparrer KMJ. Interferon antagonists encoded by SARS-CoV-2 at a glance. *Med Microbiol Immunol.* 2023;212(2):125-31, doi: 10.1007/s00430-022-00734-9.
101. Zuliani-Alvarez L, Govasli ML, Rasaiyaah J, Monit C, Perry SO, Sumner RP, et al. Evasion of cGAS and TRIM5 defines pandemic HIV. *Nat Microbiol.* 2022;7(11):1762-76, doi: 10.1038/s41564-022-01247-0.

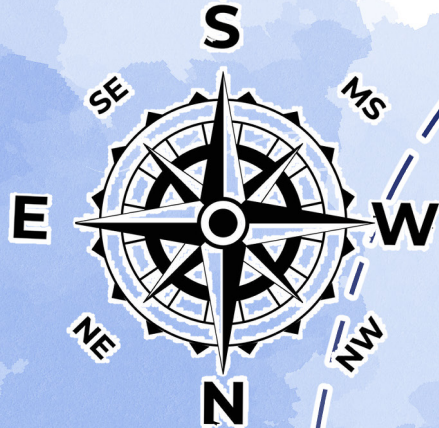
102. Jacques DA, McEwan WA, Hilditch L, Price AJ, Towers GJ, James LC. HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature*. 2016;536(7616):349-53, doi: 10.1038/nature19098.
103. Burdick RC, Li C, Munshi M, Rawson JMO, Nagashima K, Hu WS, et al. HIV-1 uncoats in the nucleus near sites of integration. *Proc Natl Acad Sci U S A*. 2020;117(10):5486-93, doi: 10.1073/pnas.1920631117.
104. Li C, Burdick RC, Nagashima K, Hu WS, Pathak VK. HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proc Natl Acad Sci U S A*. 2021;118(10):e2019467118, doi: 10.1073/pnas.2019467118.
105. Thorne LG, Bouhaddou M, Reuschl AK, Zuliani-Alvarez L, Polacco B, Pelin A, et al. Evolution of enhanced innate immune evasion by SARS-CoV-2. *Nature*. 2022;602(7897):487-95, doi: 10.1038/s41586-021-04352-y.
106. Reuschl AK, Thorne LG, Whelan MVX, Ragazzini R, Furnon W, Cowton VM, et al. Evolution of enhanced innate immune suppression by SARS-CoV-2 Omicron subvariants. *Nat Microbiol*. 2024;9(2):451-63, doi: 10.1038/s41564-023-01588-4.
107. Bouhaddou M, Reuschl AK, Polacco BJ, Thorne LG, Ummadi MR, Ye C, et al. SARS-CoV-2 variants evolve convergent strategies to remodel the host response. *Cell*. 2023;186(21):4597-614 e26, doi: 10.1016/j.cell.2023.08.026.
108. Arya R, Tripathi P, Nayak K, Ganesh J, Bihani SC, Ghosh B, et al. Insights into the evolution of mutations in SARS-CoV-2 non-spike proteins. *Microb Pathog*. 2023;185:106460, doi: 10.1016/j.micpath.2023.106460.
109. Medzhitov R, Janeway CA, Jr. Innate immunity: the virtues of a nonclonal system of recognition. *Cell*. 1997;91(3):295-8, doi: 10.1016/s0092-8674(00)80412-2.
110. Liu Y, McNevin J, Cao J, Zhao H, Genowati I, Wong K, et al. Selection on the human immunodeficiency virus type 1 proteome following primary infection. *J Virol*. 2006;80(19):9519-29, doi: 10.1128/JVI.00575-06.
111. Weber F, Kochs G, Haller O. Inverse interference: how viruses fight the interferon system. *Viral Immunol*. 2004;17(4):498-515, doi: 10.1089/vim.2004.17.498.
112. Jin Y, Zhao JH, Guo HS. Recent advances in understanding plant antiviral RNAi and viral suppressors of RNAi. *Curr Opin Virol*. 2021;46:65-72, doi: 10.1016/j.coviro.2020.12.001.
113. van Mierlo JT, Overheul GJ, Obadia B, van Cleef KW, Webster CL, Saleh MC, et al. Novel Drosophila viruses encode host-specific suppressors of RNAi. *PLoS Pathog*. 2014;10(7):e1004256, doi: 10.1371/journal.ppat.1004256.
114. Kulkarni S, Volchkova V, Basler CF, Palese P, Volchkov VE, Shaw ML. Nipah virus edits its P gene at high frequency to express the V and W proteins. *J Virol*. 2009;83(8):3982-7, doi: 10.1128/JVI.02599-08.
115. Vidal S, Curran J, Kolakofsky D. A stuttering model for paramyxovirus P mRNA editing. *EMBO J*. 1990;9(6):2017-22, doi: 10.1002/j.1460-2075.1990.tb08330.x.
116. Pisanelli G, Pagnini U, Iovane G, Garcia-Sastre A. Type I and type II interferon antagonism strategies used by *Paramyxoviridae*: previous and new discoveries, in comparison. *Viruses*. 2022;14(5):1107, doi: 10.3390/v14051107.
117. Audsley MD, Moseley GW. Paramyxovirus evasion of innate immunity: diverse strategies for common targets. *World J Virol*. 2013;2(2):57-70, doi: 10.5501/wjv.v2.i2.57.
118. Cox R, Plemper RK. The paramyxovirus polymerase complex as a target for next-generation anti-paramyxovirus therapeutics. *Front Microbiol*. 2015;6:459, doi: 10.3389/fmicb.2015.00459.

119. Elena SF, Sole RV, Sardanyes J. Simple genomes, complex interactions: epistasis in RNA virus. *Chaos*. 2010;20(2):026106, doi: 10.1063/1.3449300.
120. Mehra P, Hintze A. Reducing epistasis and pleiotropy can avoid the survival of the flattest tragedy. *Biology*. 2024;13(3):193, doi: 10.3390/biology13030193.
121. Fraisse C, Puixeu Sala G, Vicoso B. Pleiotropy Modulates the Efficacy of Selection in *Drosophila melanogaster*. *Mol Biol Evol*. 2019;36(3):500-15, doi: 10.1093/molbev/msy246.
122. Yu TC, Kikawa C, Dadonaite B, Loes AN, Englund JA, Bloom JD. Pleiotropic mutational effects on function and stability constrain the antigenic evolution of influenza haemagglutinin. *Nat Ecol Evol*. 2025, doi: 10.1038/s41559-025-02895-1.
123. Lai WY, Hsu SK, Futschik A, Schlotterer C. Pleiotropy increases parallel selection signatures during adaptation from standing genetic variation. *Elife*. 2025;13:RP102321, doi: 10.7554/eLife.102321.
124. Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA. Epistasis between beneficial mutations and the phenotype-to-fitness Map for a ssDNA virus. *PLoS Genet*. 2011;7(6):e1002075, doi: 10.1371/journal.pgen.1002075.
125. Sanjuan R, Moya A, Elena SF. The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc Natl Acad Sci U S A*. 2004;101(43):15376-9, doi: 10.1073/pnas.0404125101.
126. Chou HH, Chiu HC, Delaney NF, Segre D, Marx CJ. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*. 2011;332(6034):1190-2, doi: 10.1126/science.1203799.
127. Rochman ND, Wolf YI, Koonin EV. Molecular adaptations during viral epidemics. *EMBO Rep*. 2022;23(8):e55393, doi: 10.15252/embr.202255393.
128. Barton MI, MacGowan SA, Kutuzov MA, Dushek O, Barton GJ, van der Merwe PA. Effects of common mutations in the SARS-CoV-2 Spike RBD and its ligand, the human ACE2 receptor on binding affinity and kinetics. *Elife*. 2021;10:e70658, doi: 10.7554/eLife.70658.
129. Naresh G, Guruprasad L. Mutations in the receptor-binding domain of human SARS CoV-2 spike protein increases its affinity to bind human ACE-2 receptor. *J Biomol Struct Dyn*. 2023;41(6):2368-81, doi: 10.1080/07391102.2022.2032354.
130. Liu Y, Liu J, Plante KS, Plante JA, Xie X, Zhang X, et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature*. 2022;602(7896):294-9, doi: 10.1038/s41586-021-04245-0.
131. Yamada S, Suzuki Y, Suzuki T, Le MQ, Nidom CA, Sakai-Tagawa Y, et al. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature*. 2006;444(7117):378-82, doi: 10.1038/nature05264.
132. Wang MK, Lim SY, Lee SM, Cunningham JM. Biochemical basis for increased activity of Ebola glycoprotein in the 2013-16 epidemic. *Cell Host Microbe*. 2017;21(3):367-75, doi: 10.1016/j.chom.2017.02.002.
133. Durham ND, Jain A, Howard A, Luban J, Munro JB. Molecular basis for the increased fusion activity of the Ebola virus glycoprotein epidemic variant A82V: Insights from simulations and experiments. *Cell Rep*. 2025;44(4):115521, doi: 10.1016/j.celrep.2025.115521.
134. Dietzel E, Schudt G, Krahling V, Matrosovich M, Becker S. Functional characterization of adaptive mutations during the West African Ebola virus outbreak. *J Virol*. 2017;91(2):e01913-16, doi: 10.1128/JVI.01913-16.
135. Selman M, Dankar SK, Forbes NE, Jia JJ, Brown EG. Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerg Microbes Infect*. 2012;1(11):e42, doi: 10.1038/emi.2012.38.

136. Evseev D, Magor KE. Molecular evolution of the influenza A virus non-structural protein 1 in interspecies transmission and adaptation. *Front Microbiol.* 2021;12:693204, doi: 10.3389/fmicb.2021.693204.
137. Feng S, O'Brien A, Chen DY, Saeed M, Baker SC. SARS-CoV-2 nonstructural protein 6 from Alpha to Omicron: evolution of a transmembrane protein. *mBio.* 2023;14(4):e0068823, doi: 10.1128/mbio.00688-23.
138. Pan R, Li P, Meyerholz DK, Perlman S. Mutations in nonstructural proteins essential for pathogenicity in SARS-CoV-2-infected mice. *J Virol.* 2024;98(7):e0058424, doi: 10.1128/jvi.00584-24.
139. Crow JF, Kimura M. An introduction to population genetics theory. New York: Harper & Row; 1970, doi: 10.1002/tera.1420050318.
140. Chevin LM. On measuring selection in experimental evolution. *Biol Lett.* 2011;7(2):210-3, doi: 10.1098/rsbl.2010.0580.
141. Sanjuan R, Domingo-Calap P. Genetic diversity and evolution of viral populations. *Encyclopedia of virology.* 2021:53-61, doi: 10.1016/B978-0-12-809633-8.20958-8.
142. Nielsen R, Yang Z. Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Mol Biol Evol.* 2003;20(8):1231-9, doi: 10.1093/molbev/msg147.
143. Tian F, Tong B, Sun L, Shi S, Zheng B, Wang Z, et al. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *Elife.* 2021;10:e69091, doi: 10.7554/eLife.69091.
144. Jankowiak M, Obermeyer FH, Lemieux JE. Inferring selection effects in SARS-CoV-2 with Bayesian Viral Allele Selection. *PLoS Genet.* 2022;18(12):e1010540, doi: 10.1371/journal.pgen.1010540.
145. Eigen M. The origin of biological information. The physicist's conception of nature: Springer; 1973. p. 594-632.
146. Subasi BS, Grabe V, Kaltenpoth M, Rolff J, Armitage SAO. How frequently are insects wounded in the wild? A case study using *Drosophila melanogaster*. *R Soc Open Sci.* 2024;11(6):240256, doi: 10.1098/rsos.240256.
147. Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, Quintana JF, et al. The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. *PLoS Biol.* 2015;13(7):e1002210, doi: 10.1371/journal.pbio.1002210.
148. Castello-Sanjuan M, Gonzalez R, Romoli O, Blanc H, Nigg JC, Saleh MC. Persistent viral infections impact key biological traits in *Drosophila melanogaster*. *PLoS Biol.* 2025;23(10):e3003437, doi: 10.1371/journal.pbio.3003437.
149. Bruner-Montero G, Luque CM, Cesar CS, Ding SD, Day JP, Jiggins FM. Hunting *Drosophila* viruses from wild populations: A novel isolation approach and characterisation of viruses. *PLoS Pathog.* 2023;19(3):e1010883, doi: 10.1371/journal.ppat.1010883.
150. Kacsoh BZ, Bozler J, Ramaswami M, Bosco G. Social communication of predator-induced changes in *Drosophila* behavior and germ line physiology. *Elife.* 2015;4:e07423, doi: 10.7554/eLife.07423.
151. Dietzl G, Chen D, Schnorrrer F, Su KC, Barinova Y, Fellner M, et al. A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature.* 2007;448(7150):151-6, doi: 10.1038/nature05954.

152. Zeng X, Han L, Singh SR, Liu H, Neumuller RA, Yan D, et al. Genome-wide RNAi screen identifies networks involved in intestinal stem cell regulation in *Drosophila*. *Cell Rep.* 2015;10(7):1226-38, doi: 10.1016/j.celrep.2015.01.051.
153. Chen S, Zhang YE, Long M. New genes in *Drosophila* quickly become essential. *Science.* 2010;330(6011):1682-5, doi: 10.1126/science.1196380.
154. Heigwer F, Port F, Boutros M. RNA Interference (RNAi) screening in *Drosophila*. *Genetics.* 2018;208(3):853-74, doi: 10.1534/genetics.117.300077.
155. Xia S, VanKuren NW, Chen C, Zhang L, Kemkemer C, Shao Y, et al. Genomic analyses of new genes and their phenotypic effects reveal rapid evolution of essential functions in *Drosophila* development. *PLoS Genet.* 2021;17(7):e1009654, doi: 10.1371/journal.pgen.1009654.
156. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. *Science.* 2014;343(6166):80-4, doi: 10.1126/science.1246981.
157. Shalem O, Sanjana NE, Zhang F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet.* 2015;16(5):299-311, doi: 10.1038/nrg3899.
158. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337(6096):816-21, doi: 10.1126/science.1225829.
159. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell.* 2013;154(2):442-51, doi: 10.1016/j.cell.2013.06.044.
160. Evers B, Jastrzebski K, Heijmans JP, Grerum W, Beijersbergen RL, Bernards R. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol.* 2016;34(6):631-3, doi: 10.1038/nbt.3536.

GRAN CANARIA



NETHERLANDS

Appendices

Summary

Samenvatting

Research data management

Acknowledgements

Curriculum vitae

List of publications

PhD portfolio

Summary

Once a virus successfully enters and replicates within its host, the infection can have diverse results ranging from asymptomatic disease to the host's death. The outcome of infection depends greatly on the immune response of the host, which must mount an effective defense against the pathogen and prevent damage to the host. In turn, viruses rapidly adapt to the environment and can evolve resistance mechanisms to evade or suppress the immune response. Because of their high evolvability and transmissibility, viruses pose a considerable challenge to public health. This challenge is amplified by viruses whose genetic information is composed of RNA, which are masters of adaptation and display high rates of cross-species transmission. Such events, in which viruses infect new hosts by adapting to their cellular machinery, are often responsible for emerging diseases that can grow into pandemics.

In this context, we asked how innate immune pathways contribute to shaping viral evolution. We studied this in the model organism *Drosophila melanogaster*, by infecting the fruit flies with two viruses—one native to the *Drosophila* host and another that naturally infects a different insect species. We experimentally manipulated the major antiviral defenses, producing both immunocompromised and highly immune-competent flies. Together with wild-type control flies, we infected the three conditions for ten serial passages with both viruses. After each passage, we collected viral samples and deep-sequenced a subset to couple genetic adaptive changes with viral phenotypes.

In **Chapter 2**, we modulated the activity of RNA interference (RNAi), an RNA-based defense mechanism that degrades viral genomes, and monitored the adaptation of Drosophila C virus (DCV) to the three different fly conditions. We found that viral populations evolving in RNAi-deficient flies acquired more polymorphisms than those adapted to RNAi-competent flies. Strikingly, we observed the parallel loss of the parental viral variant in 14 out of 15 evolving lineages, upon the shift from cell culture infection to *in vivo* replication. In these lineages, we observed the fixation of one of five amino acids that replaced the aromatic tyrosine residue at position 95 of the VP3 capsid protein. The evolved populations gained replicative fitness and virulence in flies. Therefore, our results suggest that the environment in which the virus replicates exerted a stronger selective pressure than the altered immune pressure.

In **Chapter 3**, we modulated the activity of the cGAS-STING pathway, an inducible immune mechanism that senses RNA virus infection in flies and stimulates the expression of antiviral genes. We allowed DCV and Cricket paralysis virus (CrPV) to evolve in flies differing in their cGAS-STING response. Across viral populations evolved under normal or abolished cGAS-STING activity, we observed the recurrent fixation of a mutation in the CrPV transmembrane protein 2B, whereas this mutation was not selected in populations evolved in flies in which the cGAS-STING signaling was exogenously stimulated. The parallel evolution of this CrPV variant in both wild-type flies and in immunocompromised flies dramatically increased viral fitness in the *Drosophila* host, suggesting that the 2B protein mediates adaptation to novel hosts.

We next investigated how the course of RNA virus infections can be influenced by host enzymes that covalently modify RNA molecules. RNA-modifying enzymes can act on viral genomes in addition to cellular RNA, but the functional relevance of these chemical marks in RNA virus genomes remains largely unresolved. Therefore, we performed an RNAi-based screen in which we individually silenced RNA-modifying enzymes in flies (**Chapter 4**). We infected them with DCV or CrPV to assess the role of these proteins in virus replication. Unfortunately, knockdown efficiency was low in many RNAi lines, which calls for cautious interpretation of our results. Despite this limitation, we identified *Trmt61* and *Nop60B* as putative proviral factors that require further validation.

My thesis concludes with the General Discussion (**Chapter 5**), where I examine our findings in the context of current literature on virus evolution. I discuss evidence showing immune-driven virus evolution, speculate about reasons why innate immunity seemed to have little effect on virus evolution in our research, and consider possible limitations of our experimental model.



Samenvatting

Zodra een virus met succes een gastheer binnendringt en zich daarin weet te vermenigvuldigen, kan de infectie uiteenlopende gevolgen hebben, variërend van een asymptomatisch verloop tot de dood van de gastheer. De uitkomst van de infectie hangt in sterke mate af van de immuunrespons van de gastheer, die een efficiënte afweer moet opwekken tegen binnendringende ziekteverwekkers en schade in de gastheer moet voorkomen. Virussen passen zich op hun beurt snel aan hun omgeving aan en kunnen mechanismen ontwikkelen waarmee zij de immuunrespons ontwijken of onderdrukken. Door hun hoge evoleerbaarheid en overdraagbaarheid zijn virussen een aanzienlijke uitdaging voor de volksgezondheid. Dit wordt nog verder gecompliceerd doordat RNA-virussen ware meesters zijn in adaptatie en regelmatig overspringen op een nieuwe gastheersoort. Dergelijke gebeurtenissen—waarin virussen nieuwe gastheersoorten binnendringen en zich daar vermeerderen—liggen vaak ten grondslag aan opkomende infectieziekten die kunnen uitgroeien tot pandemieën.

In dit kader vroegen wij ons af hoe de aangeboren immuunrespons de virale evolutie beïnvloedt. Dit onderzochten wij in een modelorganisme, de fruitvlieg *Drosophila melanogaster*, na infectie met twee virussen—een virus dat *Drosophila* van nature infecteert en een virus dat een andere insectensoort infecteert. We manipuleerden experimenteel de belangrijkste antivirale afweermechanismen om zowel immuungecompromitteerde als sterk immuuncompetente vliegen te maken. Deze infecteerden we, naast wildtype vliegen als controle, gedurende tien seriële passages met beide virussen. Na elke passage verzamelden we virale monsters, waarin we virale sequenties bepaalden met behulp van next-generation sequencing om genetische veranderingen te koppelen aan virale fenotypes.

In **Hoofdstuk 2** moduleerden we de activiteit van RNA-interferentie (RNAi), een RNA-gebaseerd afweermechanisme dat virale genomen afbreekt, en volgden we de adaptatie van het *Drosophila* C virus (DCV) onder verschillende vliegcondities. We ontdekten dat virale populaties die evolueerden in RNAi-deficiënte vliegen meer polymorfismen verwierven dan populaties uit RNAi-competente vliegen. Daarnaast zagen we dat in 14 van de 15 evolutielijnen de ouderlijke virale variant vervangen werd door varianten met vergelijkbare mutaties in één specifiek viraal gen tijdens de seriële passages *in vivo*. In deze lijnen werd het aromatische tyrosine-residu op positie 95 in het VP3 capsid-eiwit vervangen door één van vijf mogelijke andere aminozuren die door een

enkele puntmutatie gecodeerd zouden kunnen worden. Deze geëvolueerde virale populaties hadden een hogere fitness en virulentie in vliegen. Deze resultaten suggereren dat de omgeving waarin het virus replicateert een sterkere selectiedruk uitoefent dan het veranderde immuunsysteem.

In **Hoofdstuk 3** veranderden we de functie van de cGAS-STING-signaalroute, een mechanisme dat RNA-virusinfecties in vliegen detecteert en de expressie van antivirale genen activeert. We lieten vervolgens DCV en het Cricket paralysisvirus (CrPV) evolueren in vliegen met verschillende cGAS-STING-responsen. Eén specifieke mutatie in het CrPV-transmembraaneiwit 2B werd gefixeerd in alle virale populaties die evolueerden onder normale of uitgeschakelde cGAS-STING-activiteit, maar niet in populaties van vliegen waarin de cGAS-STING-route experimenteel werd geactiveerd. De parallelle evolutie van deze CrPV variant verhoogde de virale fitness in de *Drosophila*-gastheer aanzienlijk, wat suggereert dat het 2B eiwit een rol kan spelen bij de aanpassing aan een nieuwe gastheer.

Vervolgens onderzochten we hoe het verloop van RNA-virusinfecties beïnvloed kan worden door gastheerenzymen die RNA-moleculen covalent modifieren. RNA-modificerende enzymen kunnen naast cellulaire RNA's ook virale RNA genomen veranderen, maar het belang van deze veranderingen voor virale infectie is nog grotendeels onbekend. Daarom voerden we een op RNAi gebaseerde screen uit waarin we RNA-modificerende enzymen individueel uitschakelden in vliegen (**Hoofdstuk 4**). We infecteerden deze vliegen vervolgens met DCV en CrPV om de rol van deze eiwitten in virusreproductie te analyseren. Helaas was knockdown-efficiëntie laag voor veel van de geteste eiwitten, hetgeen om een voorzichtige interpretatie van de resultaten vraagt. Ondanks deze limitatie, identificeerden we *Trmt61* en *Nop60B* als mogelijke provirale factoren. Deze vereisen verdere validatie en kunnen als basis dienen voor verder vervolgonderzoek.

Mijn proefschrift sluit af met een algemene discussie (**Hoofdstuk 5**), waarin ik onze bevindingen bespreek in de context van de huidige literatuur over virusevolutie. Ik bespreek bewijs voor virale evolutie onder druk van het immuunsysteem, speculeer over redenen waarom het aangeboren immuunsysteem weinig effect op virusevolutie had in ons onderzoek, en sta stil bij de mogelijke beperkingen van ons experimentele model.

Research data management

Ethics

For this thesis, only non-human data derived from laboratory experiments were used.

Data collection and storage

All data and analyses obtained during my PhD are stored digitally on the departmental Teams environment, which is backed up daily. Raw unaltered data were stored alongside processed files. The raw data incorporated in this thesis have been made accessible in data files that are organized by chapter and figure. Experimental procedures and results have been documented on Labguru, an electronic lab journal accessible to all members of the laboratory. The scripts used in bioinformatics analysis are available on GitHub, at <https://github.com/cbg-ethz/DCV-Evolution-Data-Processing> for Chapter 2 and at <https://github.com/cbg-ethz/DCV-CrPV-cGAS-STING-pathway-data-analysis> for Chapter 3. The sequences and location of the plasmids used in this thesis are documented in the plasmid database of the laboratory, and plasmids are available upon request from the principal investigator.

Data sharing

All raw sequencing data used in Chapters 2 and 3 of this thesis have been deposited open access in the NCBI BioProject under the accession numbers PRJNA993483 and PRJNA1257840, respectively. Other data underlying the results presented in this thesis are available in the Data Sharing Collections (DSCs) in the Radboud Data Repository, where they will be archived for a period of ten years under CC-BY-4 licensing. The DSC for Chapter 2 is available at DOI: 10.34973/agz4-m140, the DSC for Chapter 3 at DOI: 10.34973/w0vs-q122, and the DSC for Chapter 4 at DOI: 10.34973/1agj-b696. The chapters 2 and 3 are published in open access journals.

Acknowledgements

I am incredibly grateful for the opportunity I have been given during these years in the Netherlands. It has been a turning point in my life, opening many new doors for me, giving me countless amazing experiences, and broadening my perspective. Thank you to everyone who has shared this journey with me. Below, I would like to dedicate a few words to some of you.

In the first place I send a big hug to **Ronald**. I have massively enjoyed having you as a supervisor, I have learnt a lot from you, and I had a lot of fun along the way. We worked very well together; sharing excitement (often excessive on my side) while helping me navigate my PhD, as you prevented me from rushing into too many side quests. Thanks for always having your door open for a quick question, for giving me the space to develop as a scientist and for your wisdom.

Martijn, after passing your (very) strong selection filter, you let me into your friendly circle while still forcing me to be sharp and to the point. No fooling around until Friday afternoon when you were always ready to divagate with a beer in hand. With the carrot and the stick you have shown me most of what I know about evolutionary biology. Thanks for your enthusiasm about my project, it really helped me keep going.

Lara, what would I have done without you! Tirelessly developing algorithms to study virus evolution, performing all sorts of statistical tricks and finding the best color palettes. What a team we made! Perfect example of the expertise exchange that makes science richer. Thanks also to **Niko** for letting me into his lab in Basel for an amazing secondment.

Milou, from a fun intern to an indispensable person in my life in the blink of an eye! Thank you so much for sharing your relaxed energy, efficient planning and stunning mind, and for having my back when I needed it most! We've shared, learnt, grown and enjoyed so much along the way. I'm so grateful for all of it! I can't wait to discover what's in store for us, while focusing on making the most out of our present.

Chiara piccola! My time in Nijmegen would have been way grayer without you around! The mix of kindness, energy and amusement that sparks your life is so contagious that it turns everything into more light and fun. Such a privilege to have you as a paranymp. You're the kind of person that cannot age, as your



inner kid is always screaming through your eyes! Cheers to biting into life together for many more years! **Luc**, you've been a great addition to the team, and thanks for taking such good care of her!

Pascal, another young soul that is so fun to have around! Apart from being always willing to help with scientific input, your air guitar will echo in my house for eternity! I've enjoyed sooooo much the parties in your garden, keep them up!

Benny, supervisor, friend, partner in crime, and now, photographer! Who would have thought we'd evolve this way six years ago, eh! Together we're like the good wine, only getting better with time. I hope I could teach you some time management during my internship, as you were loaded with bright ideas, but short on common sense (even less than me dude!). Can't wait to find out our next adventure together! **Johanna**, since the first day in your house, where you persistently managed to make sure I couldn't bike home in a straight line, you've become essential in our plots.

Gijs, the icy technician and soul of the lab! You act all tough on the outside, but when poked enough, you turn into a gentle and attentive friend. I hope parenthood is calming your Dutch efficiency to the point in which brushing your teeth and taking out the trash can be accomplished as separate tasks!

Rebecca, although I found you slightly intimidating at the beginning, soon enough I discovered an incredibly helpful friend always keen to think along. Thanks for sharing your sharp mind and acidic sense of humor.

Ezgi, my best Turkish friend. Humble, bright and with an excellent taste in food. You've turned my PhD journey more pleasant, always ready to listen to my problems and share good times. Thanks for guiding us through the amazing chaos of Istanbul.

Valerie, I miss my party companion. Nobody can collapse a handstand as gracefully as you did! You always eased the vibe during stressful times.

Michelle, **Samara** and **Qu**, always ready to sweeten up a day with some delicious treats. **Lisa**, my fictitious wife and best roomie ever, it was great to have you around! **Bodine**, thanks for keeping me entertained with your constant sparring with Benny and keeping me fit by racing after your fast walk.

Bumi, until today “follow Bumi” down any slope and under any weather condition is one of the funniest and most challenging games I’ve ever played (and considering my record, my friend, that is a huge compliment). I hope we can keep playing for many years!

Megan, thanks for convincing me to go to Roscoff, it was great! And also for the good times as roomies, when you found love with the Mr. who broke into our house.

Tom, the best piano player I have ever played music with, humbling most people in chess and still talented in many other sports. It has been so cool to have you as a friend along our PhD, we’ve shared all sorts of discussions and adventures, from Istanbul to Gran Canaria. What’s next? **Jeroen**, invaluable lead guitar from the well-known ML4 and father of the future famous bass player Thomas. **Koen** and the rest of the MMBand, I can’t be more grateful for all the music we shared together!

Femke, I will always be thankful for showing up at my house in hard times, for your critical point of view that has taught me tough lessons and for the infinite fun I’ve had on our ski trips. **Ruben**, always a gentleman. I hope you dare to ski another speed track after such a fall!

Kirsten, from the earliest times of my internship to date, one of my essential friends in Nimma. One of the few people that can turn a coffee into hours on non-stop conversation with a sharp mockery.

Haitao, lekkerding, thank you for your kindness and generosity! I hope you can come back soon!

Femke M., the craziest new addition to the lab (Sorry **Julie**, I did not meet you on time). It’s good to see you making your way through a challenging project without losing your hobbies in the process. Hope you can keep teaching me to windsurf for a long time! **Nynke**, since we didn’t have a picture together for my photo book, let’s solve it by having some fun dances at my PhD party! **Zahië**, it was nice sharing Sinterklaas with you, all the best!

Reinder and **Carles**, my microbiology students. I’ve learnt a lot from you and enjoyed supervising you. I hope the feeling is mutual.



Mery, thanks for the long discussions about society, science or philosophy while solving the world. For reminding me the importance of letting the inner child play safely and feast on life. While certainly we've had a bumpy ride, I've grown from it and I'm grateful for that.

Emina, thank you so much for your help with the flies, for your kindness, your good taste in literature, and your inexhaustible energy to cook and share excellent Croatian food. It is a pleasure to share time with beautiful people.

Maria, it was very nice to have you around, easing up the times with a fun talk in Spanish and picking your mind on my project to help me out.

Cas and **Julie**, my favorite Dutch couple ever. You're just cool people, always in for an adventure on the bike, sharing a festival or even visiting me on Gran Canaria. I hope that you're not sore anymore after our Kayak adventure.

Felix, my 4daagse buddy, I wouldn't have done it without you, and together with **Carolina**, you are the hosts of amazing predrinks and birthday parties. So happy that your family is growing in a healthy and vigorous way. Cheers to Alma!

Maartje, A.K.A. Blondie, it's a pity we didn't get to know each other better until later in my PhD, but still; everybody knows that good things come in small packages. I wish you all the best! **Ezra** and **Laura**, the achievers and goal-oriented couple that still keep it low-key and fun! **Silvia**, thanks for hosting pre-drinks at your place! **Freia**, it was cool to have you around! Thanks for giving me that mighty glass of water after walking the 4daagse and being fully unable to move! **Kjerstin**, it was amazing to organize all the MMB marathons with you, I think it was a great initiative that should outlive us in the department.

Thanks to all other Parasitology and MMB members for the nice interactions, chill vibe and safe working space that you all created!

David, thanks to the least Dutch Dutchman I know for your spontaneity, enthusiasm and fun energy. You gave me a much-needed escape from the lab and shared so many fun nights, Viking runs and all kinds of sports with me!

Andrea y Marcos, fue un placer compartir con ustedes! Son un buen equipo con el que desahogar las dificultades del doctorado y pasarlo bien!

Papá, sin ti nada de esto hubiera sido posible. Mi soporte incondicional. Quiero agradecerte todo lo que has hecho por mí: aprender a ser padre y madre a la vez, asegurarte de que tuviéramos unas vacaciones fantásticas compartiendo tiempo juntos y enseñándome a pescar, marcar límites, permitirme desarrollar mis aficiones, escucharme y empujarme a volar mirando siempre por mi bienestar. **Mamá**, sin la resiliencia que tuve que aprender para sobreponerme a tu marcha, nunca lo habría logrado.

Maribi, la tía abuela modelo que es envidia de cualquier abuela. Has jugado un papel clave en mi desarrollo, enseñándome a ser ordenado y limpio, a programarme mis tareas, a ser generoso, a cuidar de los demás. Junto con mi padre, sin duda fueron los responsables de que me fuera a Holanda y, por ende, de que este libro exista.

Hermanita, gracias por tu generosidad, detallismo y diversión. Por salir de tu zona de confort para venir a verme y por compartir intimidades de las que muchas hermanas se tapan los oídos y saldrían corriendo.

Asun, muchas gracias por cuidarme y quererme. Sin querer sustituir a mi madre has ocupado un papel paralelo, haciéndome sentir protegido. Gracias por estar!

Jorge, el hermanito que no tuve del que el "ito" queda ya poco. Estoy muy orgulloso del tiarraco en el que te has convertido, sigue así! Me alegra ver que saboreas tu independencia y la disfrutas, porque de que te organizarías bien tu parte profesional nunca tuve dudas.

Pino, mi tía loca, MUY LOCA, más cuerda que nadie. Gracias por tu visión particular y pertinente de la vida, la familia y los animales. Por señalar lo que importa y hacer gala de ello.

Ali, compañera de juegos, excursiones y guía en el Siam Park cuando era niño. Aunque algunas cosas han cambiado, lo que remarco es que ahora te lo piensas dos veces antes de empezar una guerra de cosquillas!

Samy, fuiste el primer bebé que sostuve en brazos y ahora eres ya toda una mujercita simpática, dulce y divertida. Sigue poniendo la pasión y perseverancia que te caracterizan en lo que hagas, llegarás a donde quieras! Te quiero prima!



Aroa, mi prima pequeña que vence el miedo! Me impresionó cómo te sobrepusiste a tus miedos en el Siam Park! Inteligente y observadora, parece que no moja pero empapa. Contra pereza diligencia, y, como tu hermana, conseguirás lo que quieras. Muchísimas gracias por ser tan detallista con la familia. Te quiero mucho!

Garbi, gracias a tu generosidad (y la de **Camile**) me aventuré a ir a Holanda por primera vez; me gustó y decidí quedarme para este capítulo de mi vida que ahora se cierra. Son unos padrazos ejemplares! **Aray**, espero poder jugar contigo pronto con la bici y darte algún paseo en moto! Besotes también para **Ico**. **Pau**, persona bella donde las haya, gracias por apoyarme en momentos difíciles y hacer los buenos más felices. **Delia** y **Joan**, me encanta cómo disfrutan de la jubilación, siempre con ganas de comerse el mundo! Me alegra infinito que esta experiencia haya servido para reconectar con todos ustedes, y que así se quede para lo que nos quede aquí!

Miguel, gracias por tu incondicionalidad, presencia y amistad. Nadie explica mi tesis de una forma más amena que tú!

Pablo, gracias por estar ahí como un apoyo constante, y por dejarme cuestionar alguno de los planteamientos que en mi opinión limitan tu vida.

En cuanto a mi familia andaluza, agradecer a **Joaquín** por escuchar, comprender y acompañarme siempre, a **José** por abrirme las puertas de su casa, a **Blanca** por su entusiasmo, su energía y los buenos ratos, y a **Irene** por hacer las interminables horas de biblioteca divertidas e (intentar) enseñarme a tocar las palmas.

Curriculum vitae

Óscar Morales Lezcano was born in Las Palmas de Gran Canaria on the 6th of April 1995, to Carlos Morales Pérez and Carmen Lezcano Jaén. He grew up on the subtropical island, where he attended public secondary education and graduated in 2013.

Driven by a globetrotting spirit and a passion for molecular biology, Óscar moved to Seville to pursue a Bachelor's degree in Biochemistry at the University of Seville, graduating in 2017. During his studies, he also spent time backpacking through India. In 2018, he completed a Master's degree in Secondary Education Teaching at the same university.

Far from ready to settle down, he moved to the Netherlands looking for a PhD but, not finding it yet, he ended up selling bikes at Decathlon. He returned to Spain to obtain an MSc diploma in Regenerative Biomedicine at the University of Granada in 2020. There, he completed his first Master's internship in the laboratory of Dr. Elena Gómez-Díaz at the Spanish National Research Council (CSIC), where he studied the transcriptional responses of malaria parasite-infected mosquitoes. For his second internship, he switched to virology and joined the laboratory of Prof. Dr. Ronald van Rij at Radboud University Medical Center in Nijmegen, the Netherlands, where he investigated protein interactions between dengue virus and its mosquito vector.

As a result of his fascination with virus-host interactions, Óscar remained in the same laboratory to pursue a PhD focusing on RNA virus evolution. In particular, he investigated the influence of host immune pathways on virus adaptation, using fruit flies as an experimental model. During his PhD, he presented his work at several (inter)national conferences and was involved in numerous teaching and supervision activities.



List of publications

1. **Lezcano, O. M.**, Fuhrmann, L., Bos, R. T., Wang, H., Stevens, M., Beerenwinkel, N., Huynen, M. A., & van Rij, R. P. (2026). A recurrent adaptive mutation in the transmembrane 2B protein of an insect picorna-like virus in a nonnative host. *J Virol*, e0123925. <https://doi.org/10.1128/jvi.01239-25>
2. **Lezcano, O. M.**, Fuhrmann, L., Ramakrishnan, G., Beerenwinkel, N., Huynen, M. A., & van Rij, R. P. (2023). Parallel evolution and enhanced virulence upon in vivo passage of an RNA virus in *Drosophila melanogaster*. *Virus Evol*, 9(2), vead074. <https://doi.org/10.1093/ve/vead074>
3. Besson, B., **Lezcano, O. M.**, Overheul, G. J., Janssen, K., Spruijt, C. G., Vermeulen, M., ... & van Rij, R. P. (2022). Arbovirus-vector protein interactomics identifies Loquacious as a co-factor for dengue virus replication in *Aedes* mosquitoes. *PLoS Pathog*, 18(9), e1010329. <https://doi.org/10.1371/journal.ppat.1010329>
4. **Lezcano, O. M.**, Sánchez-Polo, M., Ruiz, J. L., & Gómez-Díaz, E. (2020). Chromatin structure and function in mosquitoes. *Front Genet*, 11, 602949. <https://doi.org/10.3389/fgene.2020.602949>

PhD portfolio of Óscar Morales Lezcano

Department: **Medical Microbiology**

PhD period: **2021-04-01 – 2026-09-18**

PhD Supervisors: **Prof. dr. ir. R.P. van Rij and Prof. dr. M.A. Huynen**

Training activities	Hours
Courses	
• RIMLS - Introduction course "In the lead of my PhD" (2021)	15.00
• Fundamentals of RNA viruses (2021)	28.00
• Extraction of viruses and bacteria from clinical samples (2022)	56.00
• Nucleic acid interaction in silico (2022)	28.00
• Using computational pipelines to analyse viral NGS data: From SNV calling to quasispecies assembly (2022)	20.00
• Phylodynamics and phylogenetics (2022)	36.00
• RU - Poster Pitching (2023)	28.00
• Making the most of your presentation (2023)	8.00
• Radboudumc - Scientific integrity (2023)	20.00
• RU - Mindfulness Based Stress Reduction (2024)	45.00
• RU - Effective Writing Strategies (2025)	75.00
Seminars	
• RIMLS Meet the Expert Illustrator (2021)	2.00
• ViroInf Midterm Meeting (2022)	28.00
• Research Integrity Round (2023)	1.50
• Research Integrity Round (2024)	1.50
Conferences	
• Dutch Arbovirus Research Network (2021)	8.00
• Dutch Belgian Drosophila Meeting (2021)	8.00
• PhD retreat (2022) *	21.00
• Dutch Arbovirus Research Network (2022)	8.00
• PhD retreat: Discovery Science & Data Science (2023) *	21.00
• Dutch Annual Virology Symposium (2023) #	16.00
• Insect Models for Infection Biology (2023) #	40.00
• Dutch Annual Virology Symposium (2024)	8.00
• International Virus Bioinformatics Meeting (2024) *	28.00
• Radboudumc Community for Infectious Diseases Science Day (2024) #	16.00
• Lowlands Flymeeting 2024 (2024)	8.00
• Dutch Annual Virology Symposium (2025) #	16.00



Training activities	Hours
Teaching	
• Lecture and Practicum in the MSc Molecular Mechanisms of Disease course "Computational Biology and Bioinformatics" (2023)	16.00
• Lecture in the MSc Medical Microbiology course "Medical Microbiology" (2023)	2.00
• Lecture and Practicum Supervision in the MSc Biomedical Sciences course "OMICS Data Analysis for Systems Biology" (2023)	16.00
• Lecture and Practicum in the MSc Molecular Mechanisms of Disease course "Computational Biology and Bioinformatics" (2024)	16.00
• Journal Club Supervision in the BSc Biomedical Sciences course "RNA Structure and Function" (2024)	10.00
• Lecture in the MSc Medical Microbiology course "Medical Microbiology" (2024)	2.00
• Practicum supervision in the MSc Biomedical Sciences course "OMICS Data Analysis for Systems Biology" (2024)	12.00
Supervision	
• Supervision of a Research Project in the BSc Biomedical Sciences (2022)	70.00
• Supervision of master student (2022)	140.00
• Supervision of a Research Project in the BSc Biomedical Sciences (2023)	45.00
• Meet the PhD (2024)	20.00
• Supervision of master student (2024)	140.00
Total hours	1,079.00

Poster and oral presentations are indicated with * and #, respectively.

