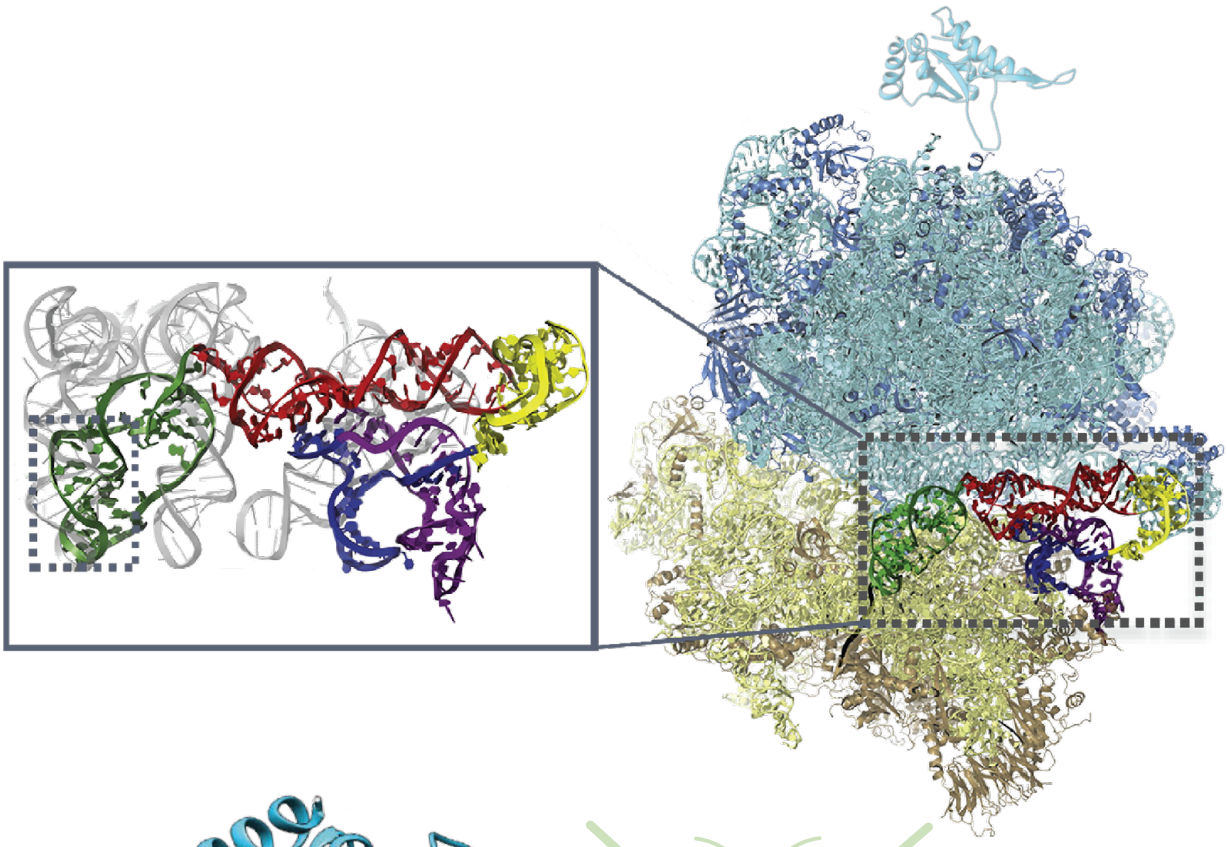
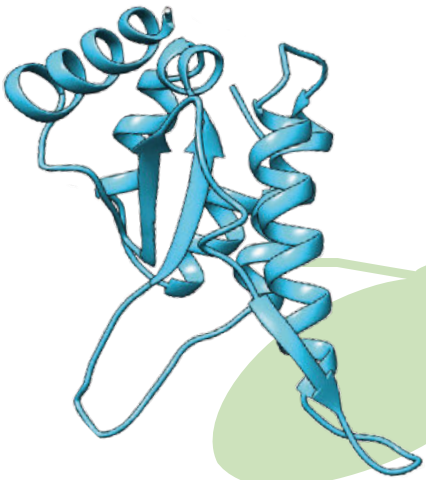
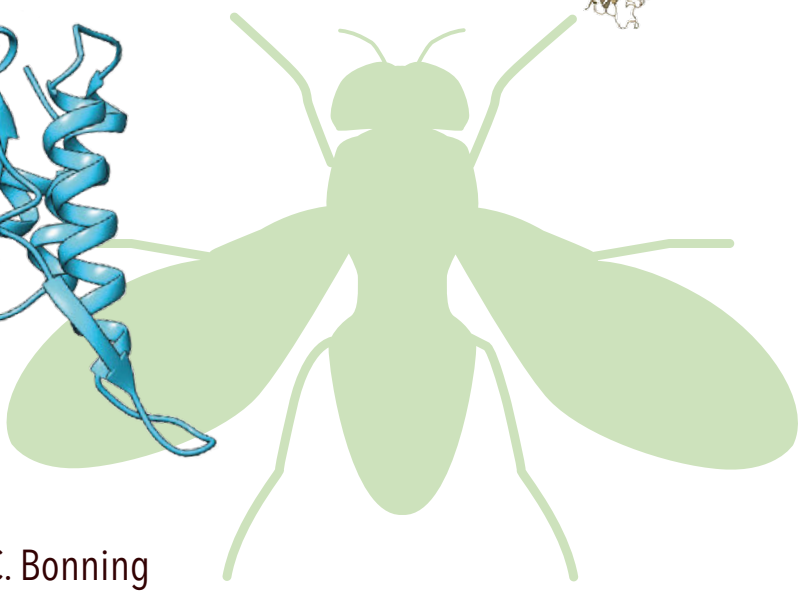
INSECT MOLECULAR VIROLOGY

Advances and Emerging Trends







Edited by Bryony C. Bonning

Caister Academic Press

Chapter 2 from:

Insect Molecular Virology

Advances and Emerging Trends

Edited by Bryony C. Bonning

ISBN: 978-1-912530-08-3 (paperback)
ISBN: 978-1-912530-09-0 (ebook)

The Widespread Occurrence and Potential Biological Roles of Endogenous Viral Elements in Insect Genomes

2

Carol D. Blair[1], Ken E. Olson[1] and Mariangela Bonizzoni[2*]

[1]Arthropod-borne and Infectious Diseases Laboratory, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO, USA.
[2]Department of Biology and Biotechnology, University of Pavia, Pavia, Italy.

*Correspondence: mariangela.bonizzoni@unipv.it

https://doi.org/10.21775/9781912530083.02

Abstract

Modern genomic sequencing and bioinformatics approaches have detected numerous examples of DNA sequences derived from DNA and RNA virus genomes integrated into both vertebrate and insect genomes. Retroviruses encode RNA-dependent DNA polymerases (reverse transcriptases) and integrases that convert their RNA viral genomes into DNA proviruses and facilitate proviral DNA integration into the host genome. Surprisingly, DNA sequences derived from RNA viruses that do not encode these enzymes also occur in host genomes. Non-retroviral integrated RNA virus sequences (NIRVS) occur at relatively high frequency in the genomes of the arboviral vectors *Aedes aegypti* and *Aedes albopictus*, are not distributed randomly and possibly contribute to mosquito antiviral immunity, suggesting these mosquitoes could serve as a model system for unravelling the function of NIRVS. Here we address the following questions: What drives DNA synthesis from the genomes of non-retroviral RNA viruses? How does integration of virus cDNA into host DNA occur, and what is its biological function (if any)? We review current knowledge of viral integrations in insect genomes, hypothesize mechanisms of NIRVS formation and their potential impact on insect biology, particularly antiviral immunity, and suggest directions for future research.

Introduction: definition of endogenous viral elements and the surprising discovery of those from non-retroviral RNA viruses

Interactions between viruses and their hosts occur at many levels. Viruses evolve rapidly through genetic drift and/or natural selection and can influence the genetic structures of host populations, especially if viral infection results in disease and mortality. A classic example of virus and host co-evolution is the natural experiment initiated by the release of highly pathogenic myxoma virus into the naive Australian wild rabbit population in 1950, followed by rapid development of resistance in rabbits coincident with natural attenuation of the virus (Best and Kerr, 2000; Kerr, 2012). Another mechanism through which viruses affect the genetics of the host is viral genome integration. Integrations of viral nucleic acid sequences can occur in both somatic and germline cells. Somatic integrations are frequently associated with host genome instability (e.g. Chen *et al.*, 2014).

Additionally, complete viral sequence integration into host genomes can favour the persistence of the infection (Cohn *et al.*, 2015). If integrations occur in germline cells, they can be inherited by the next generation. The persistence and outcome of these integrations in host populations depend on their effects on host fitness. If deleterious, integrations are likely lost. Alternatively, viral sequences can be functionally adopted by the host and exert beneficial functions (Frank and Feschotte, 2017). For instance, the product of the murine retrovirus restriction gene, *Fv1*, protects mice against infection with murine leukaemia virus (MLV). *Fv1* genes in *Mus musculus* and other *Mus* spp. were derived from the *gag* genes of ancient endogenous retroviruses, some of which are distantly related to MLV, and protect against other retroviruses, including lentiviruses and spumaviruses (Yap *et al.*, 2014). A third outcome is found when integrations are in chromosomal locations that are not transcribed or lack regulatory regions. In this case, viral integrations may persist but accumulate mutations at the host rate, which is typically much slower than that of exogenous viruses. These endogenous viral elements (EVEs) thus represent a magnifying lens into the past, allowing researchers to more closely examine the evolutionary history of viruses and virus–host co-evolution (Aiewsakun and Katsourakis, 2015; Katzourakis, 2017).

Integration of genome sequences from DNA viruses and retroviruses is a relatively common phenomenon, as evidenced by the abundance of virus-derived sequences in the genomes of various organisms. For instance, genetic code from retroviruses constitutes about 8% of the human genome (Griffiths *et al.*, 2001). Genome integrations are essential events in the retrovirus life cycle. Because non-retroviral RNA viruses lack coding for reverse transcriptase, the machinery needed for successful conversion to DNA and integration into host genomes, their potential to endogenize has been considered minimal. However, an increasing number of studies show genome integrations into both somatic and germline host cells from non-retroviral RNA viruses, including both single-stranded (positive and negative) and double-stranded RNA viruses. The acronym non-retroviral integrated RNA virus sequences (NIRVS) has been proposed to emphasize the unusual viral origin of these EVEs (Ballinger *et al.*, 2012; Kondo *et al.*, 2013; Ballinger *et al.*, 2014; Olson and Bonizzoni, 2017; Palatini *et al.*, 2017).

Here we describe current knowledge on the origin, widespread occurrence and genomic context of EVEs in insect genomes. Additionally, we report examples of the influence of EVEs on both vertebrate and insect host physiology and immunity.

Integration of genomes from DNA viruses

Polydnaviruses (PDV) provide a well-known example of DNA virus integration into insect genomes (Table 2.1A; for a complete, authoritative review of polydnaviruses and their interactions with their insect hosts, see Chapter 8). PDV are recognized by the International Committee on Taxonomy of Viruses (ICTV; https://talk.ictvonline.org/ictv-reports/ictv_9th_report/dsdna-viruses-2011/w/dsdna_viruses/127/polydnaviridae) as the family *Polydnaviridae*, divided into the genera *Bracovirus* (BV) and *Ichnovirus* (IV), which are associated with braconid and ichneumonid parasitoid wasps, respectively (Strand and Burke, 2015; Gauthier *et al.*, 2018). Bracoviruses have been domesticated by wasps, and their genomes are integrated as proviral segments dispersed throughout the genome in every cell of both female and male wasps (Desjardins *et al.*, 2008). They replicate by producing defective virus-like particles (VLPs) in the calyx of the wasp ovary. Bracovirus genomes in VLPs consist of multiple, diverse, circular dsDNAs depleted of replication genes, with one DNA molecule per VLP. Parasitoid wasps reproduce by depositing their eggs in or on other insects, which are consumed by juvenile stages of the wasp as they develop (Burke *et al.*, 2014). The defective VLPs are deposited into the wasp's hemipteran or dipteran host along with its eggs, their defective genomes migrate to the host cell nucleus, where they subsequently express virulence factors that alter host physiology and allow the development of wasp progeny (Strand and Burke, 2015; Gauthier *et al.*, 2018). PDVs cannot be transmitted horizontally; rather, the integrated proviruses are transmitted vertically in wasps. Transcript studies suggest that BVs evolved from a single whole genome integration of an ancient ancestor of the insect virus family *Nudiviridae* (Bezier *et al.*, 2009; Burke and Strand, 2012). Lineages of ichneumonid wasps with integrated IVs are polyphyletic

and integrated IVs are diverse, suggesting multiple evolutionary integration events (Gauthier *et al.*, 2018).

Analyses of a recently discovered endogenization involving an alphanudivirus and the braconid wasp *Fopius arisanus* are shedding light on the mechanisms of integration. Alphanudiviruses are members of the family *Nudiviridae* with a single circular dsDNA genome that usually cause virulent infections in insects. *Fopius arisanus* belongs to the Opiinae subfamily of Braconidae. Wasp species in this subfamily usually lack integrations from bracoviruses (Burke *et al.*, 2018). Consequently, the identification of integrated alphanudivirus sequences in the genome of *F. arisanus* suggests that this integration is a recent event. This event shows that endogenization is accompanied with gene-loss in the viral genome, and that changes in its architecture occur before expression of the endogenized virus genes falls under the control of the wasp genome (Burke *et al.*, 2018).

Genome endogenization has also been observed for the betanudivirus HzNV-1, which naturally infects the corn earworm *Helicoverpa zea.* HzNV-1 can exhibit a latent phase of infection during which its DNA is integrated into the host genome (Gauthier *et al.*, 2018). Moreover, HzNV-1 genome integrations were observed in persistently infected cells of other lepidopteran species (*Trichoplusia ni* and *Spodoptera frugiperda*). Interestingly, cells of the fall armyworm carrying endogenous HzNV-1 appeared resistant to superinfections (Lin *et al.*, 1999).

Integration of non-retroviral RNA virus genomes: discovery and association with retroelements

The first discovery of integrations of non-retroviral RNA virus genome sequences in insect genomes was the unexpected finding in 2004 of four regions similar to RNAs of cell fusing agent (CFAV) and Kamiti River (KRV) viruses in the genomes of *Aedes aegypti* and *Aedes albopictus.* These sequences, which were amplified using degenerate flavivirus primers, were called cell silent agents (CSA) because of their similarity to CFAV (Crochu *et al.*, 2004). Both CFAV and KRV are insect-specific viruses (ISVs) that are genetically related to arboviruses but exclusively infect insect cells. In 2009, additional unexpected integrations from ISVs were detected in wild-caught *Aedes* and *Ochlerotatus* genomes by PCR using flavivirus-specific primers (Roiz *et al.*, 2009; Rizzo *et al.*, 2009; Sanchez-Seco *et al.*, 2010).

Descriptions of RNA virus genome sequence integrations into insect genomes increased with the advent of next-generation sequencing technologies and metagenomic analyses in the early 2000s. Bioinformatics-based identification of viral integrations uses BLAST+ in the form of BLASTx or tBLASTn (Altschul *et al.*, 1990). Researchers parse positive BLAST hits using cut-off values of between 10^{-3} and 10^{-6} (Kondo *et al.*, 2019; Whitfield *et al.*, 2017; ter Horst *et al.*, 2018). Additional criteria (i.e. presence of viral open reading frames or minimum size restrictions) are applied to further reduce false positive results that occur primarily due to low-complexity reads with short tandem repeats or homopolymers. The complexity of the viral database used to search for viral integrations may influence the number of identified BLAST hits. Researchers have used a number of virus genome databases to identify viral integrations in host genomes. These viral sequence databases include sequences from single virus families (Fort *et al.*, 2012; Chen *et al.*, 2015), from plant viruses transmitted by insects (Cui and Holmes, 2012), all known RNA viruses, including those most-recently identified (i.e. Whitfield *et al.*, 2017; Palatini *et al.*, 2017) and from a mixture of DNA and RNA viruses (Katzourakis and Gifford, 2010; ter Horst *et al.*, 2018). These analyses provided the number of viral integrations in sequenced insect genomes and insight into the diversity and richness of their viral origins. The results were influenced by the breadth of viral genomes analysed and the stringency of the criteria used to define an insect sequence as a viral integration (i.e. cut off BLAST e-value; reverse BLAST; analyses of low complexity sequences; length of sequences; inclusion of newly discovered and unclassified viruses).

Currently-known insect genome viral integrations are largely from non-retroviral RNA viruses with sequences also derived from single-stranded DNA viruses, e.g. from densoviruses (*Parvoviridae*) in the genome of *Acyrthosiphon pisum* (ter Horst *et al.*, 2018). NIRVS have been characterized from the genomes of both haematophagous and non-bloodsucking insects (Table 2.1B and C). A review

Table 2.1 Characteristics of EVEs that have been identified in insect genomes including parasitic wasps (A), haematophagous species (B) and non-bloodsucking insects (C)

Host		EVEs					
Family	Genus, species	Viral oigin/R, D	N	Context	Expression	Discovery method	References
A. Parasitic wasps							
Hymenoptera	Ichneumonid and braconid subfamilies	Polydna, D	Full viral genome	Dispersed within wasp genome	Y		Burke and Strand (2012), Gauthier *et al*. (2018)
Lepidoptera	*Helicoverpa*; *Trichoplusia ni* and *Spodoptera frugiperda* cells	*Betanudivirus* HzNV, D	not c	ND			Lin *et al*. (1999)
Hymenoptera	*Fopius arisanus*	*Alphanudivirus*, D	55	Dispersed within host genome		*In silico* screening; PCR	Burke *et al*. (2018)
B. Haematophagous insects							
Culicidae	*Aedes albopictus*, *Aedes aegypti*, C6/36 cell lines	Flavi[1], R	6	Repetitive DNA[3]	Y	PCR	Crochu *et al*. (2004)
Culicidae	*Ae. albopictus*	Flavi[2], R	Not c	ND	ND	PCR	Roiz *et al*. (2009)
Culicidae	*Ochlerotatus* and *Aedes*	Flavi, R	Not c	ND	ND	PCR	Sanchez-Seco *et al*. (2009)
Culicidae	*Ochlerotatus geniculatus*, *Aedes vexans*	Flavi, R	Not c	ND	ND	PCR	Rizzo *et al*. (2009)
Culicidae	*Ae. aegypti*, *Culex quinquefasciatus*	Rhabdo, R	143	ND	ND	*In silico* screening	Katzourakis and Gifford (2010)
		Reo, R	1	ND	ND	*In silico* screening	
		Flavi, R	5	ND	ND	*In silico* screening	
Culicidae	*Ae. vexans*, *Och. caspius*, *Och. detritus*, *Culiseta annulata*	Flavi, R	not c	ND	ND	PCR	Vasquez *et al*. (2012)
Culicidae	*Cx. quinquefasciatus*, *Ae. aegypti*, wild mosquitoes	Rhabdo, R	112 in *Ae. aegypti*, 1 in *Cu. quinquefasciatus*	ND	Y	*In silico* screening; PCR	Fort *et al*. (2012)
Culicidae	*Ae. aegypti*	Plant RNA viruses[6], R	2	*Ace1* gene	ND	*In silico* screening	Cui and Holmes (2012)

Hemiptera	*Rhodnius prolixus*	Benyvirus, R	4	repetitive DNA[7]	ND	*In silico* screening	Kondo *et al*. (2013)
Culicidae	*Aedes aegypti*, *Anopheles spp*.	Chu, Rhabdo[8], Bunya, Phlebovirus-like, *Quaranjavirus*, R	1–12[9]		ND	*In silico* screening, RNA-seq	Li *et al*. (2015)[10]
	Culex quinquefasciatus	Chu, Rhabdo[8], R			ND	*In silico* screening, RNA-seq	
Hemiptera	*Rhodnius prolixus*	Chu, Bunya, MonoN[11]			ND	*In silico* screening, RNA-seq	
Culicidae	*Anopheles minimus*, *Anopheles sinensis*[4]	Flavi, R	2	Repetitive DNA	Y	*In silico* screening	Lequime and Lambrechts (2017)
Culicidae	*Ae. aegypti*, *Ae. albopictus*, wild mosquitoes	Flavi, R	At least 8	Repetitive DNA	Y[12]	PCR; *in silico* screening	Suzuki *et al*. (2017)
Culicidae	*Aedes*, *Culex* and *Anopheles* spp. with genome sequence available	Rhabdo, R	Up to 88 in Aedine; 7 in Anophelinae	Repetitive DNA[13]	Y[14]	*In silico* screening	Palatini *et al*. (2017)
		Flavi, R	Up to 32 in Aedine; 1 in Anophelinae				
		Bunya, R	1				
		Reo, R	1				
Culicidae	*Ae. aegypti* Aag2 cell line	Rhabdo, Flavi, Chu, Virg, Bunya, R	368	repetitive DNA[13]	ND	*In silico* screening	Whitfield *et al*. (2017)
C. Non-blood-sucking insects							
Hymenoptera	*Apis mellifera*	Discistro[15], R	1	Intergenic	Y	PCR	Maori *et al*. (2007)
Drosophilidae	Various species[16]	Sigmavirus, R	<5 per genome	Repetitive DNA[17]	Y[18]	*In silico* screening	Ballinger *et al*. (2012)
Drosophilidae	Various species[16]	Rhabdo, R	1	ND	ND	*In silico* screening	Fort *et al*. (2012)
Homoptera	*Acyrthosiphon pisum*	Rhabdo, R	1	ND	ND	*In silico* screening	Fort *et al.* (2012)
Hymenoptera; Diptera;	Various species[19]	Plant RNA viruses[5], R	<20 per genome	Genes or ND	ND	*In silico* screening	Cui and Holmes (2012)
Drosophilidae	Various species[16]	Bunya, R	2	Intergenic	Y	*In silico* screening	Ballinger *et al*. (2014)

Table 2.1 Continued

Host		EVEs					
Family	Genus, species	Viral oigin/R, D	N	Context	Expression	Discovery method	References
Coleoptera	*Dendroctonus ponderosae*	Chu, Bunya, R				*In silico* screening, RNA-seq	Li *et al*. (2015)
	Tribolium castaneum	Chu, R	3		ND	*In silico* screening, RNA-seq	
Drosophilidae	*Drosophila spp.*	Rhabdo[8,20], Bunya, R			ND	*In silico* screening, RNA-seq	
Isoptera	*Zootermopsis nevadensis*	Chu, R			ND	*In silico* screening, RNA-seq	
Hemiptera	*Acyrthosiphon pisum*	Chu, Rhabdo[8], Phlebovirus-like; *Quaranjavirus*, MonoN[11], R	9		ND	*In silico* screening, RNA-seq	
Hymenoptera	*Atta cephalotes*	MonoN[9], R			ND	*In silico* screening, RNA-seq	
	Acromyrmex echinatior	Chu, MonoN[11]	8		ND	*In silico* screening, RNA-seq	
	Camponotus floridanus	Chu, MonoN[11], Rhabdo[8], R	3		ND	*In silico* screening, RNA-seq	
	Harpegnathos saltator	Chu, R	7		ND	*In silico* screening, RNA-seq	
	Linepithema humile	Chu, R	11		ND	*In silico* screening, RNA-seq	
	Nasonia vitripennis, *N. girault*i, *N. longicornis*	Chu, R	4–20[21]		ND	*In silico* screening, RNA-seq	
	Pogonomyrmex barbarus	Chu, R	23		ND	*In silico* screening, RNA-seq	
	Solenopsis invicta	Chu, MonoN[11], R	12		ND	*In silico* screening, RNA-seq	
Lepidoptera	*Bombyx mori*	Chu, Quaranjavirus, Rhabdo[8], R			ND	*In silico* screening, RNA-seq	
	Melitaea cinxia	Rhabdo[20], Quaranjavirus, R			ND	*In silico* screening, RNA-seq	
	Plutella xylostella	Rhabdo[8], R			ND	*In silico* screening, RNA-seq	
	Spodoptera frugiperda	Phlebovirus-like, R			ND	*In silico* screening, RNA-seq	

Lepidoptera	*Spodoptera frugiperda* cell lines, *Bombyx mori*	Rhabdo[22], R	4[23]	ND	Y	*In silico* screening	Geisler and Jarvis (2016)
Hymenoptera; Thysanoptera, Lepidoptera, Diptera	*Bombus terrestris and B. ignites; Frankliniella occidentalis*, *Heliconius*	Virga/nege-like[24], R	<10 per genome	repetitive DNA	ND	*In silico* screening; PCR	Kondo *et al*. (2017)

Only data from peer-reviewed publications are reported.
[1]Cell fusing agent virus, Kamiti River virus.
[2]Aedes flavivirus.
[3]Repetitive DNA: LTR-retrotransposons of the Pao/Ninja and Copia lineages and an *An. gambiae* putative gene.
[4]*In silico* screening was done across all 24 currently available *Anophelinae* genomes.
[5]Repetitive DNA, LTR-retrotransposons of the Copia class.
[6]Passion fruit mosaic virus (*Virgaviridae*), pelargonium zonate spot virus (*Bromoviridae*), carrot red leaf virus (*Luteoviridae*), beet yellow virus (*Closteroviridae*), cardamine chlorotic fleck virus (*Tombusviridae*), bamboo mosaic virus (*Alphaflexiviridae*), blueberry nectrotic ring blotch virus and citrus leprosis virus c (unclassified).
[7]Repetitive DNA, transposable elements as *mariner* and *mos*.
[8]Dirmarhabdovirus.
[9]One NIRVS in *A. arabiensis*, *An. epiroticus*, *An. atroparvus* and *An. nilii*, two in *An. quandriannulatus*, three in *An. funestus*, *An. farauti*, *An. dirus*, *An. gambiae* and *An. minimus*; four in *An. stephensi*; 12 in *Aedes aegypti*, five in *Cx. Quinquefasciatus*.
[10]NIRVS copy number was shown only for NIRVS from Chuvirus.
[11]Unclassified Mononegavirus.
[12]mRNA expression was detected, but no proteins. Production of small RNA was also observed.
[13]Repetitive DNA, non-LTR retroelements, primarily of the Ty3/Gypsy and Pao Bel lineages.
[14]Limited mRNA expression; piRNA production.
[15]Israeli acute paralysis virus.
[16]Various species within the *Sophophora* and *Drosophila* subgenera.
[17]Repetitive DNA, non-LTR retroelements, primarily of the Ty3/Gypsy lineage.
[18]Expression was verified for one X-linked Sigma virus-P like NIRVS of *Drosophila yakuba*.
[19]NIRVS were identified in 14 out of 53-tested insect genomes, including *Drosophila rhopaloa*, *Dr. ficusphila*, *Dr. anannasse*, *Bombus terrestris*, *B. impatiens*, *Megachile rotundata*, *Pogonomyrmex barbatus*, *Acyrthosiphon pisum*, *Bombyx mori*, *Danaus plexippus*, *Nosonia longicornis*, *Nosonia vitripennis*, *Nosonia girault*i.
[20]Unclassified rhabdovirus.
[21]Four NIRVS in *Nasonia girault*i, five in *N. longcornis*, 20 in *N. vitripennis*.
[22]Sf-rhabdovirus.
[23]Four integrations were found in the genome of *S. frugiperda* and three in that of *B. mori*.
[24]Virga/nege like viruses are newly-recognized insect-specific viruses.
Key to headings: context, genomic context of integration; D, DNA-based genome; N, number; R, RNA-based genome; viral or., viral origin.
Key to table: Bunya, *Peribunyaviridae*; Chu, *Chuviridae*; Dicistro, *Dicistroviridae*; Flavi, *Flaviviridae*; N, no; ND, not determined; not c, not certain; Polydna, *Polydnaviridae*; Reo, *Reoviridae*; Rhabdo, *Rhabdoviridae*; Virg, *Virgaviridae*; Y, yes.

of the widespread occurrence of NIRVS in insect genomes suggests similar mechanisms of biogenesis and integration occurred across different insect species; however, striking differences were found in the number of NIRVS and the viral origins of those in *Aedes* mosquitoes relative to other insects.

In contrast to endogenization of polydnavirus sequences originating from entire rearranged viral genomes, NIRVS consist of complete or interrupted viral RNA open-reading frames (ORFs). Most NIRVS have similarity to coding regions of non-retroviral nucleoproteins, glycoproteins, and/or RNA-dependent-RNA polymerases (Crochu *et al.*, 2004; Katzourakis and Gifford, 2010; Fort *et al.*, 2012; Ballinger *et al.*, 2012; Cui and Holmes, 2012; Kondo *et al.*, 2013; Ballinger *et al.*, 2014; Geisler and Jarvis, 2016; Lequime and Lambrechts, 2017; Suzuki *et al.*, 2017; Palatini *et al.*, 2017; Whitfield *et al.*, 2017).

A striking similarity among NIRVS detected across different insect species and from different virus families is that integrations are frequently found in regions of repetitive DNA, in association with long terminal repeat (LTR) retroelements. Retrotransposons of the Pao/Ninja, BEL/Pao, Copia and Ty3/Gypsy lineages and, less frequently, Tc1/mariner, flank NIRVS derived from flaviviruses and rhabdoviruses in both Culicidae and Drosophilidae (Crochu *et al.*, 2004; Ballinger *et al.*, 2012; Goic *et al.*, 2013; Lequime and Lambrechts, 2017; Palatini *et al.*, 2017; Whitfield *et al.*, 2017). Retroelements also flank NIRVS from plant RNA viruses and newly-characterized alphanudivirus-like ISVs in examples from Hemiptera, Hymenoptera, Thysanoptera, Homoptera, Lepidoptera, and from Drosophilidae (Cui and Holmes, 2012; Kondo *et al.*, 2013, 2019).

The origin of NIRVS in insect genomes requires the formation of cDNA from the non-retroviral RNA template with the help of reverse transcriptases. Since NIRVS come from viruses that lack this enzyme, retrotransposons are a likely source within insect cells. Naturally occurring cDNAs of non-retroviral RNA sequences, vDNAs, have been found in cells infected with these viruses (Goic *et al.*, 2013; Nag and Kramer, 2017). A reduction in formation of *Chikungunya virus* (CHIKV; *Alphavirus*) vDNA in *Ae. albopictus* treated with a reverse transcriptase inhibitor has been observed (Goic *et al.*, 2016). This finding and the observed physical contiguity between NIRVS and retroelements in genomes of insects has led to the hypothesis that integration events in insect genomes occur through non-homologous recombination between the genomes of infecting non-retroviral RNA viruses and LTR retroelements during their reverse transcription (Ballinger *et al.*, 2012; Goic *et al.*, 2013). The hypothesis is supported by similar findings in mammalian genomes (Geuking *et al.*, 2009) and by the fact that LTR retroelement reverse transcription occurs in the cytoplasm, where most non-retroviral RNA viruses complete their life-cycle (Havecker *et al.*, 2004; Geuking *et al.*, 2009; Horie *et al.*, 2010; Ballinger *et al.*, 2012). Furthermore, LTR retroelements are active in germline cells, which could favour inheritance of newly acquired viral sequences (Morozov *et al.*, 2017).

The association between NIRVS and retroelements is also supported by the recent finding that some retroelements harbour viral-like domains (Ballinger *et al.*, 2012). Viral-like helicase domains of recently characterized ISVs were found as part of TRAS and Jockey non-LTR retrotransposon proteins in the genomes of Lepidoptera (i.e. *Plutella xylostella*), Hemiptera (i.e. *Nilaparvata lugens*, *Homalodisca vitripennis*) Orthoptera (species of the *Ceutophilius* genus), Hymenoptera (*Leptopilina boulardi*) and Culicidae (*Ae. aegypti*) species (Lazareva *et al.*, 2015; Shi *et al.*, 2016; Morozov *et al.*, 2017). Interestingly, expression of viral-insect fused proteins in some cases increased the evolutionary fitness of the harbouring retroelements by functioning as viral suppressors of RNA silencing (Lazareva *et al.*, 2015). Helicases containing a viral motif could stabilize retrotransposon transcripts and increase their transposition efficiency. An excess of stabilizing helicase activity could also negatively impact non-retroviral RNA virus infection by altering the balance between virus transcription and replication (Morozov *et al.*, 2017).

Integration of non-retroviral RNA virus genomes: viral origin and number of NIRVS in insect genomes

Because most current discoveries of NIRVS come from *in silico* studies of sequenced insect genomes with limited data from natural specimens,

knowledge of the widespread occurrence of NIRVS in insect genomes is biased by comparisons across species as compared to among populations within species. With this limitation in mind, ISVs and other viruses known to establish persistent infections are well represented in the list of NIRVS (Table 2.1B).

Currently known insect NIRVS are similar to the sequences of viruses from the *Flavivirus*, *Benyvirus*, and *Quaranjavirus* genera; the *Reoviridae*, *Rhabdoviridae* and *Peribunyaviridae* families and the recently characterized family *Chuviridae* and phlebovirus-like and virga/nege-like viruses. *Chuviridae* is a newly identified monophyletic family of negative-sense RNA viruses, with diverse genome structures (i.e. unsegmented, bi-segmented and circular forms) (Li *et al.*, 2015). *Quaranjavirus* and phlebovirus-like viruses have negative-sense, segmented RNA genomes, belong to the *Orthomyxoviridae* and *Phenuiviridae* families, respectively, and have been recently identified in arthropods (Li *et al.*, 2015). 'Virga/nege-like viruses' refers to a still poorly characterized group of positive-strand RNA viruses that appear to belong to the alphavirus-like superfamily (Kondo *et al.*, 2019). *Virgaviridae* is a family of plant alpha-like viruses, while negevirus refers to newly recognized ISVs from sandflies and mosquitoes, which are distantly phylogenetically-related to plant cileviruses (Vasilakis *et al.*, 2013; Carapeta *et al.*, 2015; Nunes *et al.*, 2017).

NIRVS from flaviviruses (F-NIRVS) have been found almost exclusively in the genomes of culicine mosquitoes (Table 2.1). The Culicidae is a large mosquito family that includes arboviral vectors such as *Aedes* spp., *Ochlerotatus* spp. and *Culex* spp. and protozoan vectors such as those in the subfamily Anophelinae (Chen *et al.*, 2015). F-NIRVS are found in the genomes of medically important arboviral vectors (*Ae. aegypti* and *Ae. albopictus*), in the genomes of the minor arboviral vectors *Ochlerotatus* spp. (i.e. *Oc. caspius* and *Oc. detritus*) and *Ae. vexans*, but not in the genome of *Cx. quinquefasciatus* (Vasquez *et al.*, 2012, Palatini *et al.*, 2017; Whitfield *et al.*, 2017; Blagrove *et al.*, 2017). Among the 19 species of Anophelinae for which a genome sequence is available, F-NIRVS were identified only in *Anopheles minimus* and *Anopheles sinensis* (Palatini *et al.*, 2017; Lequime and Lambrechts, 2017). Interestingly, the number of F-NIRVS in the *Ae. aegypti* and *Ae. albopictus* genomes is thirty times greater than that detected in any other culicine genome. Although the precise number of F-NIRVS will vary with improvements in genome annotations and with sequencing of novel culicine genomes, the large number of F-NIRVS in these two mosquito species was obtained by applying the same conservative bioinformatics pipeline across the genomes of *Aedes*, *Culex* and *Anopheles* spp., and hence the analysis was methodologically unbiased.

Current data, which come from studies using different bioinformatics approaches and analysing different insect species, show that NIRVS from reovirus, bunyavirus, sigmavirus, benyvirus, orthomyxovirus and phlebovirus genomes are sporadic, as they are found in both limited numbers and distribution across insect species (Table 2.1). Even with the paucity of data currently available, it is interesting to note that haematophagous species like *Ae. aegypti* and the kissing bug *Rhodnius prolixus* harbour not only arbovirus-derived NIRVS, but also NIRVS with similarity to the genomes of plant RNA viruses (Cui and Holmes, 2012; Kondo *et al.*, 2013). Non-blood-sucking insect species harbour NIRVS primarily from *Rhabdoviridae* and *Chuviridae*, which are also in the genomes of haematophagous insects (Table 2.1). The *Rhabdoviridae* family includes viruses that infect vertebrates, invertebrates and plants and that are extremely variable in their genomic organization (Dietzgen *et al.*, 2017; Geoghegan *et al.*, 2017). The ecological diversity of rhabdoviruses is also reflected in their frequent cross-species transmission between hosts (Geoghegan *et al.* 2017). While the ecology of *Chuviridae* is not completely understood, viruses in this family were found in arthropod species that share the same ecological niche, indicative of possible horizontal transfer (Li *et al.*, 2015). Integrations from *Rhabdoviridae* and *Chuviridae* may appear to be more frequent than NIRVS from other viruses simply because these viruses are ecologically diverse and have wide geographical distributions. Alternatively, the infection capabilities of these viruses and their capacity to transmit between host species, genera, and even host kingdoms in the case of rhabdoviruses, depending on their ecological and geographic proximity, could select for the emergence of generalist protection mechanisms, which may include viral sequence integrations.

Integration of RNA virus genomes: focus on *Aedes* spp.

Among all the insect species in which NIRVS have been identified to date, *Ae. aegypti* and *Ae. albopictus* are notable for their high number of NIRVS and the diversity of their viral origins. In the genomes of these two mosquito species, NIRVS derived from genomes of flaviviruses, rhabdoviruses, reoviruses, bunyaviruses, phleboviruses, and quaranjavirus have been identified and NIRVS have been characterized from the *Virgaviridae* and *Chuviridae* virus families. (Crochu *et al.*, 2004; Roiz *et al.*, 2009; Katzourakis and Gifford, 2010; Fort *et al.*, 2012; Cui and Holmes, 2012; Li *et al.*, 2015; Suzuki *et al.*, 2017; Palatini *et al.*, 2017; Whitfield *et al.*, 2017; ter Horst *et al.*, 2018). All these studies agree on hundreds of individual NIRVS in *Aedes* spp., although the exact number of NIRVS from each viral type differs among studies depending on the bioinformatics parameters, sequencing and assembly method used, and source of data (i.e. genome sequence from the Aag2 cell line or adult mosquitoes). This richness and diversity of NIRVS in *Ae. aegypti* and *Ae. albopictus* is striking because these species are among the most medically important vectors of arboviruses such as dengue, Zika, yellow fever and chikungunya. This observation has prompted ongoing research efforts to understand the distribution of NIRVS in natural populations and whether NIRVS influence mosquito physiology and arbovirus infection and transmission.

NIRVS are not distributed randomly in the genomes of *Ae. aegypti* and *Ae. albopictus*, but rather are statistically significantly enriched in regions corresponding to PIWI-interacting RNA (piRNA)-generating loci, called piRNA clusters (Palatini *et al.*, 2017; Whitfield *et al.*, 2017). NIRVS are also differentially distributed in geographic populations (Pischedda *et al.*, 2018). A difference between F-NIRVS and NIRVS with similarity to rhabdoviruses (R-NIRVS) was seen in the first whole-genome survey of NIRVS across five geographical populations of *Ae. albopictus*: R-NIRVS were more widespread and included more ancient integrations based on accumulation of sequence changes than F-NIRVS (Pischedda *et al.*, 2018), an interesting result considering that negative-strand RNA viruses like rhabdoviruses are thought to be evolutionarily more recent than positive-strand flaviviruses (Koonin *et al.*, 2015).

Exogenous virus resistance mediated by endogenous viral elements

The most comprehensive current data on the effects of NIRVS on host physiology are from studies of endogenous bornavirus-like nucleoprotein-encoding sequences (EBLNs) in human, mouse and squirrel genomes (Honda and Tomonaga, 2016). A recombinant EBLN from the ground squirrel genome was expressed in human oligodendroglioma cells, where it co-localized with bornaviral sites of replication in the nucleus and inhibited the activity of infecting Bornavirus polymerase (Fujino *et al.*, 2014). This dominant negative effect was assumed to be dependent on the relatively high-degree of amino-acid sequence identity (77%) between the protein expressed from the integrated complete viral ORF from the ground squirrel genome and the nucleocapsid protein of the Bornavirus (Fujino *et al.*, 2014). However, the minimum degree of sequence-identity between the NIRVS-expressed protein and that of cognate viruses essential to exert competitive or dominant interaction was not established.

A transcript of an EBLN of the human genome (i.e. hsEBLN-1) has been proposed to function as a long non-coding RNA and control the expression of the neighbouring COMMD3 gene. COMMD3 encodes a protein that inhibits the NFκB immune pathway (Burstein *et al.*, 2005; Sofuku *et al.*, 2015). As such, hsEBLN-1 is considered an epigenetic immune regulator (Sofuku *et al.*, 2015; Honda and Tomonaga, 2016). This type of function has not been observed for insect NIRVS, likely because, in addition to differences between the insect and vertebrate antiviral immune pathways, insect NIRVS are predominantly found in repetitive regions of the genome and correspond to partial viral ORFs.

The first evidence of a biological role for an insect NIRVS came when a ≈ 420 bp sequence from the *Israeli acute paralysis virus* (IAPV) was found integrated into the genome of bees (*Apis mellifera*), resulting in their resistance to subsequent infection with IAPV (Maori *et al.*, 2007). The molecular mechanisms underlying this phenomenon were not

elucidated however (Maori *et al.*, 2007; see Chapter 3 for details of similar observations in *Drosophila*).

Endogenous viral elements and the piRNA pathway in mosquitoes

A possible biological function for NIRVS in mosquitos may be to enhance antiviral defence (Meisen *et al.*, 2016b; Blair and Olson, 2017; Olson and Bonizzoni, 2017). An antiviral function was suggested by two observations. First, detection of RNA virus genome-derived piRNAs was shown in both uninfected and infected mosquitoes and mosquito cells (Scott *et al.*, 2010; Brackney *et al.*, 2010; Arensburger *et al.*, 2011; Morazzoni *et al.*, 2012; Vodovar *et al.*, 2012; Schnettler *et al.*, 2013; Miesen *et al.*, 2016a,b). Second, integration of most NIRVS occurred in or near retroelements in piRNA clusters of *Ae. aegypti* and *Ae. albopictus* genomes. These observations not only suggest a mechanism by which non-retroviral RNA genomes were converted to cDNA and inserted into host genomes, but also that NIRVS function similarly to piRNA clusters, the genomic loci that express transposon-derived piRNAs implicated in defending host genome integrity by restricting transposon activity.

The small-interfering RNA (siRNA) pathway of the RNA interference (RNAi) system is the major antiviral defence system in acute arboviral infections of *Aedes* spp. mosquitoes (Sánchez-Vargas *et al.*, 2009; Scott *et al.*, 2010; Blair, 2011). In mosquitoes, the siRNA pathway is a cell-autonomous, cytoplasmic system in which the dsRNA-specific RNase III family member Dicer 2 (Dcr2) recognizes viral dsRNA produced during replication. Dcr2 binds the dsRNA and cleaves it into 21-nt siRNA duplexes with perfect base-pairing and 2-nt overhangs at the 3′ ends. In complex with the dsRNA-binding protein R2D2, Dcr2 loads the siRNA duplex onto Argonaute 2 (Ago2), an endonuclease that is part of the multi-component RNA-induced silencing complex (RISC). Ago2 cleaves and releases one strand of the siRNA duplex, retaining the second strand as a guide to hybridize to a complementary coding sequence in viral genome/mRNA, which it then cleaves in the centre of the complementary region. Transiently knocking down expression of Dcr2, R2D2, or Ago2 results in greatly increased arbovirus replication, indicating their central role in siRNA-mediated antiviral defence (Keene *et al.*, 2004; Sánchez-Vargas *et al.*, 2009). In either *Ae. aegypti* mosquitoes or *Ae aegypti* embryo-derived Aag2 cell cultures, infection with the flavivirus dengue (DENV) results in production of DENV RNA-derived 21 nt siRNAs (Scott *et al.*, 2010). However, in *Ae albopictus*-derived C6/36 cells, which have a single-nucleotide deletion in their Dcr2 gene (Scott *et al.*, 2010), arbovirus infection drives production of virus RNA-derived 27-nt piRNAs (Scott *et al.*, 2010, Brackney *et al.*, 2010). The more robust replication of DENV in exo-siRNA-deficient mosquito cells emphasizes the importance of this pathway in antiviral defence in *Aedes* spp., but production of piRNAs suggests that this pathway might have a secondary role in anti-arbovirus defence.

In *Drosophila*, which has served as a model for the mosquito RNAi system, production of mature piRNAs is carried out by the three members of the Piwi subfamily of the Argonaute protein family, Piwi, Aubergine (Aub) and Ago3, and is Dcr-independent (Brennecke, 2007, Gunawardane, 2007). These proteins are expressed in germline cells of the fly and their major function is to protect the genome by silencing transposable element (TE) activity. The substrates for production of primary piRNAs are nuclear transcripts from piRNA clusters in the genome that contain defective TE sequences. The transcripts are processed in the nucleus to produce mature, antisense Piwi- or Aub-associated 24–27 nt piRNAs, known as piRNA-induced silencing complexes (piRISC), that hybridize to and cleave complementary sense-strand RNA in the cytoplasm. This 'slicer' action of the piRISC results in positive-sense, Ago3-bound piRNA that can initiate a 'ping-pong' amplification loop (Siomi, 2011).

The Piwi protein subfamily has undergone expansion to eight proteins, Ago3 and Piwi1–7, in *Ae. aegypti* (Campbell *et al.*, 2008), suggesting additional functional roles of these proteins. Also, the piRNA pathway is active in somatic as well as germline tissues of *Ae. aegypti* (Morazzani *et al.*, 2012). Further studies are needed to determine which of the eight *Aedes* Piwi-family proteins may be involved in binding, processing and cleavage of

the piRNA precursor. Determining whether a particular Piwi-family protein affects virus replication could depend on the mosquito tissue or cultured cells analysed, the time post-infection at which analysis is done, and the infecting arbovirus (flavivirus or alphavirus) (Vodovar *et al.*, 2012; Morazzani *et al.*, 2012; Schnettler *et al.*, 2013; Miesen *et al.*, 2015, 2016a; Goic *et al.*, 2016). In addition to piRNAs derived from NIRVS in *Aedes* spp., newly generated piRNAs derived from the infecting arbovirus genome occur during acute infections of *Aedes* spp. cells and mosquitoes by flaviviruses, alphaviruses, and bunyaviruses (Scott, 2010; Morazzani, 2012; Léger, 2013; Schnettler, 2013; Miesen, 2015, 2016a,b). However, their role in antiviral defence is uncertain and the template for transcription of virus genome-derived piRNAs (vpiRNAs) and protein(s) involved in their biogenesis are not fully understood.

Several studies have shown that *Aedes* mosquitoes and mosquito cells are capable of synthesizing short segments of cDNA (vDNA) from alphavirus, flavivirus, or bunyavirus genome templates, using endogenous reverse transcriptase (RT), likely through recombination between viral RNA and a retrotransposon transcript during reverse transcription (Vodovar *et al.*, 2012; Nag *et al.*, 2016; Nag and Kramer, 2017). By analogy to the proviral DNA form of retrovirus genomes, this hybrid DNA is termed 'proviral'. Proviral DNA is hypothesized to enter the nucleus where it is transcribed into primary piRNA precursors from vDNA episomes or after integration into a piRNA cluster in the genome using the retroelement integrase. Although robust production of virus genome-derived piRNAs has been observed in alphavirus-infected mosquitoes (Morazzani *et al.*, 2012), no alphavirus genome-related NIRVS have been reported in mosquitoes. The evolutionary conservation of NIRVS and their variation in different populations of *Aedes* spp. provide support for their acquisition during acute infection. NIRVS acquisition and any long-term antiviral effects for a particular arbovirus would depend on infection of mosquito germ-line tissue and integration into the vector genome of NIRVS derived from the infecting viral RNA. Recent bioinformatics analyses have detected NIRVS derived from *Dengue virus-1* in the genome of a distinct *Ae. albopictus* population from China; this may signify a recent NIRVS acquisition (Chen *et al.*, 2015). Furthermore, germ-line infections and vertical transmission of dengue viruses in *Ae. aegypti* probably occur more frequently than previously thought (Sánchez-Vargas *et al.*, 2018).

Concluding remarks and future directions

The first experimental evidence that non-retroviral RNA viruses can integrate genome sequences into host DNA genomes surfaced in the 1970s (Zhdanov, 1975) but this evidence was ignored until the development of powerful and inexpensive next-generation sequencing technology that enabled the sequencing of many host genomes. Analyses of genome data are revealing an increasing number of viral sequence integrations in different organisms including insects. Although current data are mostly descriptive, preliminary observations are already transforming our perception of virus–host co-evolution, allowing for the generation of hypotheses about connections between virus evolution, insect ecology, insect genomic structure and insect antiviral immunity.

Understanding the significance of the widespread occurrence of viral integrations within insect genomes will require analysis of the genomes of more species (Fig. 2.1). It would be beneficial for the search for – and characterization of – viral integrations to be standard aims during the annotation of insect genomes. This would help answer whether integration of viral sequences is a common phenomenon within insects, or if it depends on the structure of the host genome (e.g. on the content of LTR transposons). It will also be important to verify which viral species/genera can efficiently integrate, especially as non-retroviral RNA viruses have been proposed as delivery tools for gene therapy (Mogler and Kamrud, 2015). This type of medical treatment is based on the transient presence of non-retroviral RNA viruses in host cells. For safety, it is essential to verify whether integrations into host genomes occur because of the properties of the host genome or of the virus. The use of long-read technologies to refine the genome assembly of species where NIRVS have been characterized will help resolve repetitive regions and unravel the landscape of NIRVS, including their numbers and their positioning within the genome (Whitfield *et al.*, 2017; Matthews *et al.*, 2018). Pertinent to this aim, a new,

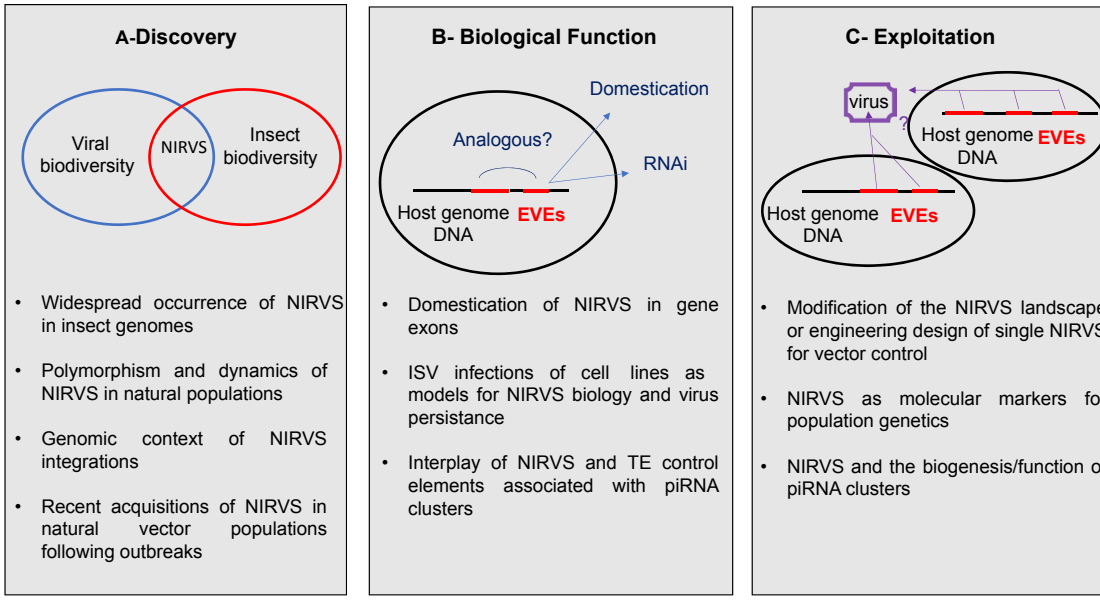


Figure 2.1 Current knowledge gaps and areas for future research on endogenous viral elements from non-retroviral RNA viruses (i.e. non-retroviral integrated RNA virus sequences or NIRVS) in insect genomes. (A) NIRVS Discovery, a better understanding of the widespread occurrence of NIRVS in insect genomes, the timing of integrations and the genomic context in which integrations occur will lead to understanding of the process of endogenization and its impact on insect genomes. (B) Biological Function, indirect evidence on the distribution of NIRVS in the genomes of the mosquitoes *Ae. aegypti* and *Ae. albopictus* and their viral origin point to an interplay with TE control elements associated with piRNA clusters and mainly ISVs, which could be experimentally addressed in cell-line based systems. Cases of NIRVS domestication could also have occurred given the finding of several NIRVS as complete ORFs or part of annotated exons. (C) Exploitation, depending on the differential distribution of NIRVS in insect populations, NIRVS could be co-opted as novel molecular markers to depict population genetic structure or trace recent invasions. Additionally, given that in *Ae. aegypti* and *Ae. albopictus* NIRVS are statistically significantly enriched in piRNA clusters, unravelling the process of endogenization could offer insights on the biogenesis and functions of piRNA clusters in non-model organisms. Finally, NIRVS may be harnessed for vector control by modifying the landscape of NIRVS within the vector genome or using NIRVS sites as anchors to drive expression of effector molecules. In this regard, studies need to determine whether NIRVS have a direct role in anti-viral immunity, how the level of sequence-similarity between NIRVS and incoming viruses can attain antiviral activity, and if it is possible to genetically modify mosquitoes by introducing novel NIRVS into the host genome.

more complete and accurate annotation of the *Ae. aegypti* genome, designated AaegL5, has been released to resolve additional NIRVS and piRNA cluster sites (Matthews *et al.*, 2018). This information could enable testable predictions regarding the mechanisms and entities involved in the integration process (e.g. by confirming NIRVS arrangements with respect to specific TE lineages). At the same time, elucidating the sites of integrations could help design experiments to test for NIRVS functions. If integrations are numerous, if NIRVS are duplicated and located primarily within piRNA clusters, which are composed of TE fragments, using genetic engineering strategies that require unique anchor sequences to modify NIRVS (i.e. knock-out NIRVS or introduce novel viral sequences) may be difficult. Examining the genomic landscape and polymorphism of NIRVS in natural specimens will enable exploration of the pervasiveness and micro-evolution of integrations. Finally, the similarity of most NIRVS to the genomes of ISVs, most of which are easily cultivated in insect cells, point to the promise of using ISVs and insect hosts as 'animal model' systems to elucidate the endogenization process and study the impact of NIRVS on exogenous virus infections.

The major controller of acute arbovirus infection in the vector is the siRNA pathway of RNAi. We (and others) have observed that when the RNAi inhibitor B2 (from *Flock House virus*) is expressed

by the arbovirus Sindbis (genus *Alphavirus*), Sindbis virus infection in the vector is greatly increased and the fitness of the vector is significantly lowered (Myles *et al.*, 2008; Cirimotich *et al.*, 2009). We hypothesize that NIRVS act by expressing virus genome-derived piRNAs in the vector to target and control infections. Thus, NIRVS may represent a fortuitous co-adaptation between arboviruses and vectors (Blair and Olson, 2015; Olson and Blair, 2016; Palatini *et al.*, 2017). NIRVS probably do not prevent or ablate infections but help maintain virus persistence in the vectors.

We have developed RNAi- or ribozyme-based antiviral effector genes. Integration and expression of either of these effector genes in *Ae. aegypti* at the time of infection profoundly lowers or eliminates vector competence for several arboviruses (Franz *et al.*, 2006; Mishra *et al.*, 2016). However, these antiviral transgenes are integrated randomly in the genome using mariner or piggyBAC transposable elements, and their expression varies according to their insertion site. As we improve our understanding of NIRVS integration, expression, and biological functions, NIRVS could provide valuable insights into how we might optimally engineer antiviral transgenes into the vector genome to maintain antiviral gene cassettes for multiple generations, as observed for piRNA clusters. As NIRVS reveal naturally selected genetic engineering of the vector to control infection outcomes, NIRVS may provide the perfect road map for how to successfully engineer antiviral transgenes.

References

Aiewsakun, P., and Katzourakis, A. (2015). Endogenous viruses: connecting recent and ancient viral evolution. Virology *479–480*, 26–37. https://doi.org/10.1016/j.virol.2015.02.011

Altschul, S.F., Gish. W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local 645 alignment search tool. J. Mol. Biol. *215*, 403–410.

Arensburger, P., Hice, R.H., Wright, J.A., Craig, N.L., and Atkinson, P.W. (2011). The mosquito *Aedes aegypti* has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. BMC Genomics *12*, 606. https://doi.org/10.1186/1471-2164-12-606

Ballinger, M.J., Bruenn, J.A., and Taylor, D.J. (2012). Phylogeny, integration and expression of sigma-virus-like genes in *Drosophila*. Mol. Phylogenet. Evol. *65*, 251–258. https://doi.org/10.1016/j.ympev.2012.06.008

Ballinger, M.J., Bruenn, J.A., Hay, J., Czechowski, D., and Taylor, D.J. (2014). Discovery and evolution of bunyavidids in artic phantom midges and ancient bunyavirid-like sequences in insect genomes. J. Virol. *88*, 8783–8794. https://doi.org/10.1128/JVI.00531-14

Best, S.M., and Kerr, P. (2000). Coevolution of host and virus: the pathogenesis of virulent and attenuated strains of myxoma virus in resistant and susceptible european rabbits. Virology *267*, 36–48.

Bézier, A., Annaheim, M., Herbiniere, J., Wetterwald, C., Gyapay, G., Bernard-Samain, S., Wincker, P., Roditi, I., Heller, M., Belghazi, M., *et al.* (2009). Polydnaviruses of braconid wasps derive from an ancestral nudivirus. Science *323*, 926–930. https://doi.org/10.1126/science.1166788.

Blagrove, M.S., Sherlock, K., Chapman, G.E., Impoivil, D.E., McCall, P.J., Medlock, J.M., Lycett, G., Solomon, T., and Baylis, M. (2017). Evaluation of vector competence of a native UK mosquito *Ochlerotatus detritus* (*Aedes detritus*) for dengue, Chikungunya and West Nile viruses. Parasit. Vectors *9*, 452. https://doi.org/10.1186/s13071-016-1739-3

Blair, C.D. (2011). Mosquito RNAi is the major innate immune pathway controlling arbovirus infection and transmission. Future Microbiol. *6*, 265–277. https://doi.org/10.2217/fmb.11.11

Blair, C.D., and Olson, K.E. (2015). The role of RNA interference (RNAi) in arbovirus–vector interactions. Viruses *7*, 820–843. https://doi.org/10.3390/v7020820

Blair, C.D., and Olson, K.E. (2017). Chapter 12 – Targeting dengue virus replication in mosquitoes. In Arthropod Vector: Controller of Disease Transmission, Volume 1, Wikel, S.K., Aksoy, S., Dimopoulos, G., eds. (Academic Press, San Diego), pp. 201–217.

Brackney, D.E., Scott, J.C., Sagawa, F., Woodward, J.E., Miller, N.A., Schilkey, F.D., Mudge, J., Wilusz, J., Olson, K.E., Blair, C.D., *et al.* (2010). C6/36 *Aedes albopictus* cells have a dysfunctional antiviral RNA interference response. PLOS Neglect. Trop. Dis. *4*, e856. https://doi.org/1 0.1371/journal.pntd.0000856

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089–1103.

Burke, G.R., and Strand, M.R. (2012). Polydnaviruses of parasitic wasps: domestication of viruses to act as gene delivery vectors. Insects *3*, 91–119. https://doi.org/10.3390/insects3010091

Burke, G.R., Walden, K.K.O., Whitfield, J.B., Robertson, H.M., and Strand, M.R. (2014). Widespread genome reorganization of an obligate virus mutualist. PLOS Genetics *10*, e1004660. https://doi.org/10.1371/journal.pgen.1004660

Burke, G.R., Simmonds, T.J., Sharanowski, B.J., and Beib, S.M. (2018). Rapid viral symbiogenesis via changes in parasitoid wasp genome architecture. Mol. Biol. Evol. *35*, 2463–2474. https://doi.org/10.1093/molbev/msy148

Burstein, E., Hoberg, J.E., Wilkinson, A.S., Rumble, J.M., Csomos, R.A., Komarck, C.M., Maine, G.N., Wilkinson, J.C., Mayo, M.W., and Duckett, C.S. (2005). COMMD proteins, a novel family of structural and functional homologs of MURR1. J Biol Chem *280*, 22222–22232.

Campbell, C.L., Black, W.C. IV, Hess, A.M., and Foy, B.D. (2008). Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. BMC

Genomics *9*, 425. https://doi.org/10.1186/1471-2164-9-425

Carapeta, S., do Bem, B., McGuinness, J., Esteves, A., Abecasis, A., Lopes, Â., de Matos, A.P., Piedade, J., António, P.G., de Almeida, A.P.G., and Parreira, R. (2015). Negeviruses found in multiple species of mosquitoes from southern Portugal: isolation, genetic diversity, and replication in insect cell culture. Virology *483*, 318–328.

Chen, X.G., Jiang, X., Gu, J., Xu, M., Wu, Y., Deng, Y., Zhang, C., Bonizzoni, M., Dermauw, W., Vontas, J. *et al.* (2015). Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. Proc. Natl. Acad. Sci. U.S.A. *112*, E5907–5915. https://doi.org/10.1073/pnas.1516410112

Chen, Y., Willimans, V., Filippova, M., Filippov, V., and Duerksen-Hughes, P. (2014). Viral carcinogenesis: factors inducing DNA damage and virus integration. Cancers *6*, 2155–2186. https://doi.org/10.3390/cancers6042155

Cirimotich, C.M., Scott, J.C., Phillips, A.T., Geiss, B.J., and Olson, K.E. (2009). Suppression of RNA interference increases Alphavirus replication and virus-associated mortality in *Aedes aegypti* mosquitoes. BMC Microbiol. *9*, 49. https://doi.org/10.1186/1471-2180-9-49

Cohn, L.B., Silva, I.T., Oliveira, T.Y., Rosales, R.A., Parrish, E.H., Learn, G.H., Hahn, B.H., Czartoski, J.L., McElrath, M.J., Lehmann, C., *et al.* (2015). HIV-1 integration landscape during latent and active infection. Cell *160*, 420–432. https://doi.org/10.1016/j.cell.2015.01.020

Crochu, S., Cook, S., Attoui, H., Charrel, R.N., De Chesse, R., Belhouchet, M., Lemasson, J.J., de Micco, P., and de Lamballerie, X. (2004). Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of *Aedes* spp. mosquitoes. J. Gen. Virol. *85*, 1971–1980.

Cui, J., and Holmes, E.C., (2012). Endogenous RNA viruses of plants in insect genomes. Virology *427*, 77–79. https://doi.org/10.1016/j.virol.2012.02.014

Desjardin, C.A., Gundersen-Rindal, D.E., Hostetler, J.B., Tallon, L.J., Fadrosh, D.W., Fuester, R.W., Pedroni, M.J., Haas, B.J., Schatz, M.C., Jones, K.M., *et al.* (2008). Comparative genomics of mutualistic viruses of *Glyptapanteles* parasitic wasps. Genome Biol *9*, R183. https://doi.org/10.1186/gb-2008-9-12-r183

Dietzgen, R.G., Kondo, H., Goodin, M.M., Kurath, G., and Vasilakis, N. (2017). The family Rhabdoviridae: mono- and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. Virus Res. *227*, 158–170. https://doi.org/10.1016/j.virusres.2016.10.010

Fort, P., Albertini, A., Van-Hua, A., Berthomieu, A., Roche, S., Delsuc, F., Pasteur, N., Capy, P., Gaudin, Y., and Weill, M. (2012). Fossil rhabdoviral sequences integrated into arthropod genomes: ontogeny, evolution, and potential functionality. Mol. Biol. Evol. *29*, 381–390. https://doi.org/10.1093/molbev/msr226

Frank, J.A., and Feschotte, C. (2017). Co-option of endogenous viral sequences for host cell function. Curr. Opin. Virol. *25*, 81–89. https://doi.org/10.1016/j.coviro.2017.07.021

Franz, A.W.E., Sanchez-Vargas, I., Adelman, Z.N., Blair, C.D., Beaty, B.J., James, A.A., and Olson, K.E. (2006). Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. Proc. Natl. Acad. Sci. U.S.A. *103*, 4198–4203.

Fujino, K., Horie, M., Honda, T., Merriman, D.K., and Tomonaga, K. (2014). Inhibition of Borna disease virus replication by an endogenous bornavirus-like element in the ground squirrel genome. Proc. Natl. Acad. Sci. U.S.A. *111*, 13175–13180.

Gauthier, J., Drezen, J.M., and Herniou, E.A. (2018). The recurrent domestication of viruses: major evolutionary transitions in parasitic wasps. Parasitology *145*, 713–723. https://doi.org/10.1017/S0031182017000725

Geisler, C., and Jarvis, D.L. (2016). Rhabdovirus-like endogenous viral elements in the genome of Spodoptera fruigiperta insect cells are actively transcribed: implications for adventitious virus detection. Biologicals *44*, 219–225. https://doi.org/10.1016/j.biologicals.2016.04.004

Geoghegan, J.L., Duchene, S., and Holmes, E.C. (2017). Comparative analyses estimates the relative frequencies of co-divergence and cross-species transmission within viral families. PLOS Pathog. *13*, e1006215. https://doi.org/10.1371/journal.ppat.1006215

Geuking, M.B., Weber, J., Dewannieux, M., Gorelik, E., Heidmann, T., Hengartner, H., Zinkernagel, R.M., and Hangartne, L. (2009). Recombination of retrotransposon and exogenous RNA virus results in non-retroviral cDNA integration. Science *323*, 393–396. https://doi.org/10.1126/science.1167375

Goic, B., Vodovar, N., Mondotte, J.A., Monot, C., Frangeul, L., Blanc, H., Gausson, V., Vera-Otarola, J., Cristofari, G., and Saleh, M.-C. (2013). RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. Nat. Immunol. *14*, 396–403. https://doi.org/10.1038/ni.2542

Goic, B., Stapleford, K.A., Frangeul, L., Doucet, A.J., Gausson, V., Blanc, H., Schemmel-Jofre, N., Cristofari, G., Lambrechts, L., Vignuzzi, M., *et al.* (2016). Virus-derived DNA drives mosquito vector tolerance to arboviral infection. Nat. Commun. *7*, 12410. https:/https://doi.org/10.1038/ncomms12410

Griffiths, D.J. (2001). Endogenous retroviruses in the human genome sequence. Genome Biol Rev. *2*, 1017.1–1017.5.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. Science *315*, 1587–1590.

Havecker, E.R., Gao, X., and Voytas, D.F. (2004). The diversity of LTR retrotransposons. Genome Biol. *5*, 225.

Honda, T., and Tomonaga, K. (2016). Endogenous non-retroviral RNA virus elements evidence a novel type of antiviral immunity. Mob. Genet. Elements *6*, e1165785. https://doi.org/10.1080/2159256X.2016.1165785

Horie, M., Honda, T., Suzuki, Y., Kobayashi, Y., Daito, T., Oshida, T., Ikuta, K., Jern, P., Gojobori, T., Coffin, J.M., *et al.* (2010). Endogenous non-retroviral RNA virus elements in mammalian genomes. Nature *463*, 84–87.

Katzourakis, A. (2017). Editorial overview: Paleovirology: the genomic fossil record, and consequences of ancient

viral infections. Curr. Opin. Virol. *25*, ix–xi. https://doi.org/10.1016/j.coviro.2017.08.002

Katzourakis, A., and Gifford, R.J. (2010). Endogenous viral elements in animal genomes. PLOS Genet. *6*, e1001191. https://doi.org/10.1371/journal.pgen.1001191

Keene, K.M., Foy, B.D., Sanchez-Vargas, I., Beaty, B.J., Blair, C.D., and Olson, K.E. (2004). RNA interference acts as a natural antiviral response to O'nyong-nyong virus (*Alphavirus*; *Togaviridae*) infection of *Anopheles gambiae*. Proc. Natl. Acad. Sci. U.S.A. *101*, 17240–17245.

Kerr, P.J. (2012). Myxomatosis in Australia and Europe: A model for emerging infectious diseases. Antiviral Res. *93*, 387–415. https://doi.org/10.1016/j.antiviral.2012.01.009

Kondo, H., Hirano, S., Chiba, S., Andika, I.B., Hirai, M., Maeda, T., and Tamada, T. (2013). Characterization of burdock mottle virus, a novel member of the genus *Benyvirus*, and the identification of benyvirus-related sequences in the plant and insect genomes. Virus Res. *177*, 75–86. https://doi.org/10.1016/j.virusres.2013.07.015

Kondo, H., Chiba, S., Maruyama, K., Andika, I.B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. Virus Res. *262*, 37–47. https://doi.org/10.1016/j.virusres.2017.11.020

Koonin, E.V., Dolja, V.V., and Krupovic, M. (2015). Origins and evolution of viruses of eukaryotes: the ultimate modularity. Virology *479–480*, 2–25. https://doi.org/10.1016/j.virol.2015.02.039

Lazareva, E., Lezzhov, A., Vassetzky, N., Solovyev, A., and Morozov, S. (2015). Acquisition of full-length viral helicase domains by insect retrotransposon-encoded polypeptides. Front. Microbiol. *6*, 1447. https://doi.org/10.3389/fmicb.2015.01447

Léger, P., Lara, E., Jagla, B., Sismeiro., O., Mansuroglu, Z., Coppée, J.Y., Bonnefoy, E., and Bouloy, M. (2013). Dicer-2- and Piwi-mediated RNA interference in Rift Valley Fever Virus-infected mosquito cells. J. Virol. *87*, 1631–1648. https://doi.org/10.1128/JVI.02795-12

Lequime, S., and Lambretchs, L. (2017). Discovery of flavivirus-derived endogenous viral elements in *Anopheles* mosquito genomes supports the existence of *Anopheles*-associated insect-specific flaviviruses. Virus Evol. *3*, vew035. https://doi.org/10.1093/ve/vew035

Li, C.X., Shi, M., Tian, J.H., Lin, X.D., Kang, K.Y., Chen, L.J., Quin, X.C., Xu, J., Holmes, E.C., and Zhang, Y.Z. (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife *4*, e05378. https://doi.org/10.7554/eLife.05378

Lin, C.L., Lee, J.C., Chen, S.S., Wood, H.A., Li, M.L., Li, C.F., and Chao, Y.C. (1999). Persistent Hz-1 virus infection in insect cells: evidence for insertion of viral DNA into host chromosomes and viral infection in a latent status. J. Virol. *73*, 128–139.

Maori, E., Tanne. E., and Sela. I. (2017). Reciprocal sequence exchange between non-retro virus and hosts leading to the appearance of new host phenotypes. Virology *362*, 342–349.

Matthews, B.J., Dudchenko, O., Kingan, S.B., Koren, S., Antoshechkin, I., Crawford, J.E., Glassford, W.J., Herre, M., Redmond, S.N., Rose, N.H. *et al.* (2018). Improved reference genome of *Aedes aegypti* informs arbovirus vector control. Nature *563*, 501–507. https://doi.org/10.1038/s41586-018-0692-z

Miesen, P., Girardi, E., and van Rij, R.P. (2015). Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. Nucleic Acids Res. *43*, 6545–6556. https://doi.org/10.1093/nar/gkv590

Miesen, P., Ivens, A., Buck, A.H., and van Rij, R.P. (2016a). Small RNA profiling in Dengue Virus 2-infected *Aedes* mosquito cells reveals viral piRNAs and novel host miRNAs. PLOS Negl. Trop. Dis. *10*, e0004452. https://doi.org/10.1371/journal.pntd.0004452

Miesen, P., Joosten. J., and van Rij, R.P. (2016b). PIWIs go viral: arbovirus-derived piRNAs in vector mosquitoes. PLOS Pathogens *12*, e1006017. https://doi.org/10.1371/journal.ppat.1006017

Mishra, P., Furey, C., Balaraman, V., and Fraser, M. (2016). Antiviral hammerhead ribozymes are effective for developing transgenic suppression of Chikungunya virus in *Aedes aegypti* mosquitoes. Viruses *8*, E163. https://doi.org/10.3390/v8060163

Mogler, M.A., and Kamrud, K.I. (2015). RNA-based viral vectors. Expert Rev. Vaccines *14*, 283–312. https://doi.org/10.1586/14760584.2015.979798

Morazzani, E.M., Wiley, M.R., Murreddu, M.G., Adelman, Z.N., and Myles, K.M. (2012). Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. PLOS Pathog. *8*, e1002470. https://doi.org/10.1371/journal.ppat.1002470

Morozov, S.Y., Lazareva, E.A., and Solovyev, A.G. (2017). RNA helicase domains of viral origin in proteins of insect retrotransposons: possible source for evolutionary advantages. PeerJ *5*, e3673. https://doi.org/10.7717/peerj.3673

Myles, K.M., Wiley, M.R., Morazzani, E.M., and Adelman, Z.N. (2008). Alphavirus-derived small RNAs modulate pathogenesis in disease vector mosquitoes. Proc. Natl. Acad. Sci. U.S.A. *105*, 19938–19943. https://doi.org/10.1073/pnas.0803408105

Nag, D.K., and Kramer, L.D. (2017). Patchy DNA forms of the Zika virus RNA genome are generated following infection in mosquito cell cultures and in mosquitoes. J. Gen. Virol. *98*, 2731–2737. https://doi.org/10.1099/jgv.0.000945

Nag, D.K., Brecher, M., and Kramer, L.D. (2016). DNA forms of arboviral RNA genomes are generated following infection in mosquito cell cultures. Virology *498*, 164–171.

Nunes, M.R.T., Contreras-Gutierrez, M.A., Guzman, H., Martins, L.C., Barbirato, M.F., Savit, C., Balta, V., Uribe, S., Vivero, R., Suaza, J.D., *et al.* (2017). Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. Virology *504*, 152–167. https://doi.org/10.1016/j.virol.2017.01.022

Olson, K.E., and Blair, C.D. (2015). Arbovirus–mosquito interactions: RNAi pathway. Curr. Opin. Virol. *15*, 119–126. https://doi.org/10.1016/j.coviro.2015.10.001

Olson, K.E., and Bonizzoni, M. (2017). Nonretrovial Integrated RNA viruses in arthropod vectors: an

occasional event or something more? Curr. Opin. Insect Sci. *22*, 45–53. https://doi.org/10.1016/j.cois.2017.05.010

Palatini, U., Miesen, P., Carballar-Lejarazu, R., Ometto, L., Rizzo, E., Tu, Z., van Rij, R.P., and Bonizzoni, M. (2017). Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors *Aedes aegypti* and *Aedes albopictus*. BMC Genomics *18*, 512. https://doi.org/10.1186/s12864-017-3903-3

Pischedda, E., Scolari, F., Valerio, F., Carballar-Lejarazu, R., Catapano, P.l., Waterhouse, R.M., and Bonizzoni, M. (2018). Evolutionary landscape of mosquito viral integrations. bioRxiv. https://doi.org/10.1101/385666

Rizzo, F., Cerutti, F., Ballardini, M., Mosca, A., Vitale, N., Radaelli, M.C., Desiato, R., Prearo, M., Pautasso, A., Casalone, C., *et al.* (2014). Molecular characterization of flavivirus from field-collected mosquitoes in northwestern Italy, 2011–2012. Parasit. Vectors *7*, 395. https://doi.org/10.1186/1756-3305-7-395

Roiz, D., Vazquez, A., Sanchez Seco, M.P., Tenorio, A., and Rizzoli, A. (2009). Detection of novel insect flavivirus sequences integrated in *Aedes albopictus* (Diptera: Culicidae) in northern Italy. Virol. J. *6*, 93. https://doi.org/10.1186/1743-422X-6-93

Sanchez-Seco, M.P., Vazquez, A., Collao, X., Hernandez, L., Aranda, C., Ruiz, S., Escosa, H., Marques, E., Bustillo, M.A., Molero, F., *et al.* (2010). Surveilance of arboviruses in Spanish wetlands: detection of new Flavi- and Phleboviruses. Vector Borne Zoonotic Dis. *10*, 203–206. https://doi.org/10.1089/vbz.2008.0188

Sánchez-Vargas, I., Scott, J.C., Poole-Smith, B.K., Franz, A.W., Barbosa-Solomieu, V., Wilusz, J., Olson, K.E., and Blair, C.D. (2009). Dengue virus type 2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. PLOS Pathog. *5*, e1000299. https://doi.org/10.1371/journal.ppat.1000299

Sánchez-Vargas, I., Harrington, L., Doty, J.B., Black, W.C., and Olson, K.E. (2018). Demonstration of efficient vertical and venereal transmission of dengue virus type-2 in a genetically diverse laboratory strain of *Aedes aegypti*. PLOS Negl. Trop. Dis. *12*, e0006754. https://doi.org/10.1371/journal.pntd.0006754

Schnettler, E., Donald, C.L., Human, S., Watson, M., Siu, R.W.C., McFarlane, M., Fazakerley, J.K., Kohl, A., and Fragkoudis, R. (2013). Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. J. Gen. Virol. *94*, 1680–1689. https://doi.org/10.1099/vir.0.053850-0

Scott, J.C., Brackney, D.E., Campbell, C.L., Bondu-Hawkins, V., Hjelle, B., Ebel, G.D., Olson, K.E., and Blair, C.D. (2010). Comparison of Dengue virus type 2-specific small RNAs from RNA interference-competent and -incompetent mosquito cells. PLOS Negl. Trop. Dis. *4*, e848.

Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P., Eden, J.S., *et al.* (2016). Redefining the invertebrate RNA virosphere. Nature *540*, 539–543. https://doi.org/10.1038/nature20167

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A., 2011. PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. *12*, 246–258. https://doi.org/10.1038/nrm3089

Sofuku, K., Parrish, N.F., Honda, T., and Tomonaga, K. (2015). Transcription profiling demonstrates epigenetic control of non-retroviral RNA virus-derived elements in the human genome. Cell Rep. *12*, 1548–1554. https://doi.org/10.1016/j.celrep.2015.08.007

Strand, M.R., and Burke, G.R. (2015). Polydnaviruses: from discovery to current insights. Virology *479–480*, 393–402. https://doi.org/10.1016/j.virol.2015.01.018

Suzuki, Y., Frangeul, L., Dickson, L.B., Blanc, H., Verdier, Y., Vinh, J., Lambrechts, L., and Saleh, M.C. (2017). Uncovering the repertoire of endogenous flaviviral elements in *Aedes* mosquito genomes. Virol. J. *91*, e00571-17. https://doi.org/10.1128/JVI.00571-17

ter Horst, A.M., Nigg, J.C., and Flak, B.W. (2018). Endogenous viral elements are widespread in arthropod genomes and commonly give rise to piRNAs. bioRxiv. https://doi.org/10.1101/396382

Vasilakis, N., Forrester, N.L., Palacios, G., Nasar, F., Savji, N., Rossi, S.L., Guzman, H., Wood, T.G., Popov, V., Gorchakov, R., Gonzalez, A.V., *et al.* (2013). Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. Virol. J. *87*, 2475–2488. https://doi.org/10.1128/JVI.00776-12

Vasquez, A., Sanchez-Seco, M.P., Palacio, G., Molero, F., Reyes, N., Ruiz, S., Aranda, C., Marques, E., Escosa, R., Moreno, J., *et al.* (2012). Novel flaviviruses detected in different species of mosquitoes in Spain. Vector Borne Zoonotic Dis. *12*, 223–229. https://doi.org/10.1089/vbz.2011.0687

Vodovar, N., Bronkhorst, A.W., van Cleef, K.W.R., Miesen, P., Blanc, H., van Rij, R.P., and Saleh, M.C., 2012. Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. PLOS ONE *7*, e30861. https://doi.org/10.1371/journal.pone.0030861

Whitfield, Z.J., Dolan, P.T., Kunitomi, M., Tassetto, M., Seetin, M.G., Oh, S., Heiner, C., Paxinos, E., and Andino, R. (2017). The diversity, structure, and function of heritable adaptive immunity sequences in the *Aedes aegypti* genome. Curr. Biol. *27*, 3511–3519.e7. https://doi.org/10.1016/j.cub.2017.09.067

Yap, M.W., Colbeck, E., Ellis, S.A., and Stoye, J.P. (2014). Evolution of the retroviral restriction gene *FV1*: inhibition of non-MLV retroviruses. PLOS Pathog. *10*, e1003968. https://doi.org/10.1371/journal.ppat.1003968

Zhdanov, V.M. (1975). Integration of viral genomes. Nature *256*, 471–473.