Sensing Viral Infections in Insects: A Dearth of Pathway Receptors

Loïc Talide¹, Jean-Luc Imler¹,² and Carine Meignin¹,²*

¹Université de Strasbourg, Centre National de la Recherche Scientifique, Modèles Insectes de l’Immunité Innée (M3i; UPR9022), Strasbourg, France.
²Université de Strasbourg, Faculté des Sciences de la Vie, Strasbourg, France.
*Correspondence: c.meignin@ibmc-cnrs.unistra.fr

Abstract
Insects, the most diverse group of animals, can be infected by an extraordinary diversity of viruses. Among them, arthropod-borne viruses can be transmitted to humans, while bee and silkworm viruses cause important economic losses. Like all invertebrates, insects rely solely on innate immunity to counter viral infections. Protein-based mechanisms, involving restriction factors and evolutionarily conserved signalling pathways regulating transcription factors of the NF-κB and STAT families, participate in the control of viral infections in insects. In addition, RNA-based responses play a major role in the silencing of viral RNAs. We review here our current state of knowledge on insect antiviral defence mechanisms, which include conserved as well as adaptive, insect-specific strategies. Identification of the innate immunity receptors that sense viral infection in insects remains a major challenge for the field.

Introduction
With more than 1 million known species, insects are the largest group of multicellular organisms, representing over 70% of animal species. Dating back to the Early Ordovician (about 480 million years ago, Mya), they were among the first animals to colonize terrestrial and freshwater ecosystems and have undergone major expansions, culminating in the spectacular diversification of holometabolous insects (Hymenoptera, Diptera, Lepidoptera) during the Early Cretaceous (about 120 Mya) (Misof et al., 2014). Insects can be credited with major adaptations, such as flight and establishment of social groups. Importantly, they have undoubtedly contributed to shape the planet’s biota and actively interact with other multicellular eukaryotes such as plants and vertebrates. Like them, insects are exposed to a large panel of infectious microorganisms, which they control through their innate immune system (Hoffmann et al., 1999).

Among infectious microbes, viruses represent a particular threat because they offer few intrinsic targets for inhibition by antiviral molecules. This is because they consist in their simplest form of a nucleic acid encapsidated in a protein shell, and hijack molecular machineries from host cells to complete their replication cycle. Therefore, viruses exert great selective pressure on their host to evolve resistance pathways. These, in turn, favour the adaptation of viruses to escape antiviral mechanisms. This arms race results in the diversification of both host-defence and virus escape mechanisms. As a result, it can be highly instructive to broaden the study of antiviral immunity to non-mammalian models. In light of their diversity, insects represent an interesting group of animals for this type of comparative study (Marques and Imler, 2016; Martins et al., 2016).

Recent advances in high-throughput sequencing technologies have opened the way to the
characterization of the virome of insects (i.e. the genetic diversity of viruses in a biological sample; see also Chapter 1). In a landmark article, Yong-Zhen Zhang and colleagues analysed the transcriptome of more than 220 invertebrate species covering nine animal phyla and reported the identification of close to 1500 new viruses (Shi et al., 2016). Thus, infection by one or several viruses is common in invertebrates. In addition, the genetic diversity of these viruses surpassed that described previously. Many newly identified viruses fell between families and genera from the current virus classification, filling major phylogenetic gaps and revealing that viruses form a continual spectrum of phylogenetic diversity (Shi et al., 2016). A more detailed analysis focusing on 70 arthropod species and negative-sense RNA viruses, which include important pathogens causing a variety of diseases in humans (flu, rabies, encephalitis, haemorrhagic fever), led to the discovery of 112 new viruses (Li et al., 2015). This study revealed that much of the diversity of negative-sense RNA viruses found in plants and vertebrate animals further supports the hypothesis that negative-sense RNA viruses, including vertebrate-specific ones, are derived from arthropod dependent viruses (Li et al., 2015).

There are several specific reasons to study virus–host interactions in insects. First, infection of insects can cause important economic losses (e.g. viral diseases of silkworms; contribution to colony-collapse in honey bees) (Bradshaw et al., 2016; Carrillo-Tripp et al., 2016). Second, haematophagous insects such as Aedes or Anopheles mosquitoes can transmit viral diseases to mammalian hosts. These viruses, the arthropod-borne viruses or arboviruses, include Dengue (DENV), Yellow fever (YFV), and West Nile virus (WNV) (Molina-Cruz et al., 2016; Powers and Waterman, 2017). Third, microbial pathogens (e.g. baculoviruses) can be used as biological control agents against insect pests, which necessitates some knowledge of the host response to these microorganisms (Popham et al., 2016). Fourth, insects such as the genetically tractable model organism Drosophila melanogaster can be used to decipher evolutionarily conserved innate immune mechanisms.

**Antiviral immunity in insects**

**NF-κB and STAT dependent inducible responses**

Innate immunity is the first line of defence that multicellular organisms deploy to limit pathogen infections. In vertebrates, the innate immune response also regulates the production of cytokines and co-stimulatory molecules, which shape the subsequent adaptive immune response (Hoffmann et al., 1999). Studies on innate immunity in Drosophila initially focused on bacterial and fungal infections and revealed that the systemic production of antimicrobial peptides (AMPs) (humoral response) plays an important role in host defence (Steiner et al., 1981; Immler and Bulet, 2005). Also, in the haemolymph, proteolytic cascades involving sequential activation of serine proteases participate in the clotting and melanisation responses to wounding (Binggeli et al., 2014; Theopold et al., 2014). In addition, cellular responses involving both circulating and sessile haemocytes participate in antimicrobial host defence in flies, in particular via phagocytosis of bacteria or virus infected cells by macrophage-like plasmatocytes and, in larvae, encapsulation of parasitic wasp eggs by lamellocytes (Gold and Brückner, 2015; Letourneau et al., 2016; Weavers et al., 2016). In the case of viruses, it is now well established that the cell intrinsic mechanism of RNA interference (RNAi) plays a central role in the control of viral infections in insects, as it does in plants and other invertebrates (see below) (Ding, 2010). In addition, inducible responses and restriction factors also contribute to resistance to viral infections (reviewed in Mussabekova et al., 2017).

Expression of AMPs is controlled by the evolutionarily conserved signalling pathways Toll and IMD (immune deficiency), which regulate the activity of transcription factors of the NF-κB family (reviewed in Hoffmann, 2003). These pathways are activated by pattern recognition receptors (PRRs) that sense components of the bacterial or fungal cell wall such as peptidoglycan in the case of bacteria or β-glucans in the case of fungi (Steiner, 2004; Royet
et al., 2011; Rao et al., 2018). Based on the transcriptomic signature of virus-infected insects and the phenotype of flies or mosquitoes with genes encoding important components of the pathway mutated or silenced, both of these pathways are proposed to also participate in antiviral immunity (Avadhanula et al., 2009; Costa et al., 2009; Ferreira et al., 2014; McFarlane et al., 2014; Paradkar et al., 2014; Carissimo et al., 2015; Lamiable et al., 2016a; Fig. 3.1). A third evolutionarily conserved pathway connected to inflammation in mammals, the Jak/STAT pathway, has also been proposed to play a role in insect antiviral immunity (Dostert et al., 2005; Fragkoudis et al., 2008; Souza-Neto et al., 2009; Paradkar et al., 2012; Kemp et al., 2013; Baribeau et al., 2015; Carissimo et al., 2015; Merkling et al., 2015a; Zhang et al., 2016; Jupatanakul et al., 2017; West and Silverman, 2018). This pathway is activated by cytokines of the Unpaired (Upd) family, which are upregulated by viral infection or stress in Drosophila (Jiang et al., 2009; Kemp et al., 2013; Gordon et al., 2018; Fig. 3.1). In Culex mosquitoes and bumble bees, a single von Willebrand factor type C domain secreted factor, related to the Drosophila antiviral factor Vago (Dedduche et al., 2008), appears to activate the Jak/STAT pathway and antiviral immunity (Paradkar et al., 2012; Wang et al., 2017). Two major questions in the field arise at this stage: (1) How are these pathways activated by viruses? And (2) what are the antiviral effectors they regulate and how do they counter viruses?

**Activation of the Toll and IMD pathways by viruses**

The Toll and the IMD pathways are activated by PRRs of the peptidoglycan recognition protein (PGRP) and β-glucan recognition protein (GNBP) families, which sense components of the bacterial and fungal cell walls (reviewed in Ferrandon et al., 2007; Royet et al., 2011). Interestingly, both pathways can also be activated by virulence factors. In the case of the Toll pathway, the circulating zymogen Persephone senses protease activity independently of microbial patterns and triggers a proteolytic cascade that culminates in the processing of the cytokine Spaetzle to generate an active Toll ligand (Gottar et al., 2006; El Chamy et al., 2008; Issa et al., 2018; Fig. 3.1). In a conceptually similar manner, deamidation by bacterial toxins of a critical glutamine residue in the Rho GTPase Rac2 can be sensed by IMD, resulting in activation of the pathway independently from PRRs (Boyer et al., 2011).

For the moment, it remains unclear how viruses activate the Toll and IMD pathways, and whether this involves PRRs or other sensors. In mammals,

---

**Figure 3.1** Antiviral transcriptional responses in *Drosophila melanogaster*. The signalling pathways activated during viral infection in *Drosophila* are illustrated.
PRRs belonging to different structural families [e.g. Toll like receptors (TLRs), RIG-1-like receptors (RLRs), nucleotide binding domain and leucine-rich repeat containing receptors (NLRLs)] sense viral nucleic acids and trigger expression of the antiviral cytokines of the interferon family (Goubau et al., 2013; Roers et al., 2016). At this stage, the only identified receptor for viruses in insects is Dicer-2, which interacts with viral double-stranded (ds) RNA. Interestingly, Dicer-2 has been proposed to serve a dual function in antiviral immunity, activating both antiviral RNAi and signalling leading to expression of antiviral molecules (e.g. Vago) (Dedouche et al., 2008; Paradkar et al., 2012; Wang et al., 2017; Asad et al., 2018; Fig. 3.1). Viruses also affect the physiology of the cells, which can trigger a response. For example, infection by Drosophila C virus (DCV) induces a heat shock response, possibly in reaction to the accumulation of unfolded viral proteins in the cytosol of infected cells (Merkling et al., 2015b). Viruses are also notorious for inhibiting cellular translation or altering cellular membranes, which may trigger a cellular response, such as apoptosis or autophagy (Shelly et al., 2009; Nainu et al., 2015; Lamiable et al., 2016b; Khong et al., 2017).

Of note, it is becoming apparent that components of the canonical Toll or IMD pathways may participate in novel pathways, responding to different cues and activating distinct transcriptional programs. For example, two closely related NF-κB proteins, Dorsal and DIF, are regulated by the Toll pathway in fruit flies (TANJI et al., 2010). In spite of strong sequence conservation, DIF lacks important features of Dorsal-mediated pattern formation in the Drosophila embryo (Stein et al., 1998). Conversely, DIF, but not Dorsal, is required to regulate expression of the antifungal peptide Drosomycin to resist fungal infections in adult flies, although both proteins are expressed (Lemaître et al., 1995; Rutschmann et al., 2000a). Interestingly, in a context of oral infection of adult flies, it is Dorsal, rather than DIF, that is required for resistance to DCV (Ferreira et al., 2014; Fig. 3.1). The other components of the Toll pathway, up to the cytokine Spaetzle, are also involved in resistance to DCV. An intriguing unresolved question pertains to the proteases acting upstream of Spaetzle in the context of DCV oral infection (Ferreira et al., 2014). Interestingly, although not essential in the context of DCV infection, DIF is required for the strong up-regulation of the cytokine Diedel (see below) triggered by Sindbis virus (SINV) and Vesicular stomatitis virus (VSV). However, MyD88, a key signalling adapter protein of the Toll pathway, is dispensable for induction of Diedel (Lamiable et al., 2016a). This suggests that, in the context of some viral infections, DIF can be activated in a MyD88-independent manner and, presumably, by a mechanism distinct from the canonical Toll pathway.

The IMD pathway can be activated by SINV in Drosophila, by a mechanism that remains unclear (Avadhanula et al., 2009; Lamiable et al., 2016a). Some antimicrobial peptides regulated by this pathway have been proposed to participate in the control of this virus (Huang et al., 2013). The relevance of this pathway in the context of viral infections is highlighted by the observation that members of different families of DNA viruses have hijacked a gene called Diedel, which encodes an immunomodulatory cytokine down-regulating the IMD pathway and antagonizing apoptosis (Coste et al., 2012; Lamiable et al., 2016a; Mih et al., 2018). This observation prompted a reanalysis of the contribution of the IMD pathway to antiviral immunity. Unexpectedly, this study revealed that two components of the pathway, the kinase IKKβ and the NF-κB transcription factor Relish, restrict replication of the dicistroviruses DCV and Cricket paralysis virus (CrPV) in Drosophila (Goto et al., 2018). Strikingly, the other components of the canonical IMD pathway, including IKKγ, the regulatory subunit of the IKK complex, are not required for virus suppression. Further analysis revealed that in the context of viral infection, the kinase IKKβ is activated by a different pathway, involving the Drosophila homologue of the mammalian gene STING (Stimulator of Interferon genes), a critical component of the cytosolic DNA sensing pathway in mammals (Fig. 3.1). The genes regulated by this alternative new pathway are different from the antibacterial peptides regulated by IMD, suggesting that Relish interacts with an additional transcription factor (Goto et al., 2018). This would be conceptually similar to the cooperation of NF-κB with IRF3 to regulate antiviral genes in mammals (Ikushima et al., 2013). Also pointing to an involvement of components of the IMD pathway in antiviral immunity, in Culex mosquitoes, a Dicer-2-dependent pathway regulates...
expression of the gene Vago upon activation of a TRAF factor and the homologue of Relish, REL2 (Paradkar et al., 2014). In summary, it appears that NF-κB pathways are more diverse than initially thought, and that these evolutionarily conserved transcription factors can be activated by alternative branches of the canonical Toll and IMD pathways initially characterized in the context of bacterial and fungal infections. Consistent with these findings, vankyrins form a family of 1κB-like molecules encoded by polydnaviruses that can antagonize NF-κB-dependent responses (Kroemer and Webb, 2005, 2006; Gueguen et al., 2013).

Control of viruses by restriction factors

Constitutively expressed restriction factors also participate in the control of viruses in insects. Some of them are evolutionarily conserved and the functions of their mammalian homologues point to mechanisms for virus inhibition. For example, Drosophila ref(2)P, a restriction factor for Sigma virus (i.e. a Rhabdovirus that is a natural pathogen of Drosophila), is a homologue of human p62/sequestosome-1, pointing to possible involvement of autophagy in the control of this virus (Carré-Mlouka et al., 2007; Ktistakis and Tooze, 2016). Interestingly, other restriction factors are not conserved and represent insect-specific adaptations. One example is the gene CHKov1, which encodes another restriction factor for Sigma virus (Magwire et al., 2011). The function of CHKov1 is unknown, but the protein contains a choline kinase domain, which is intriguing in light of the function of choline kinase as a host factor for Hepatitis C virus (HCV) in human hepatocytes (Wong and Chen, 2016, 2017). The gene pastrel is another example of a non-conserved gene having a potent restricting activity on DCV and CrPV in Drosophila (Magwire et al., 2012; Cao et al., 2017). How this protein functions to block viral replication, and if its action involves interaction with viral RNA remains unknown.

In summary, it is clear that protein-based mechanisms, involving evolutionarily conserved genes and pathways, are involved in insect antiviral immunity. However, the receptors that sense viruses and trigger these responses are unknown. As a result, Dicer-2 remains the only well characterized sensor for viral infection in insect cells (Fig. 3.1).

The siRNA pathway of RNA interference: mechanism and regulation

RNAi pathways in insects

Fire, Mello and coworkers coined the term RNA interference, or ‘RNAi’, to describe the observation that dsRNA can block gene expression when introduced into Caenorhabditis nematodes (Fire et al., 1998). This discovery was rapidly followed by biochemical characterization of RNAi in fruit flies using embryos and the S2 cell line (Hammond et al., 2000; Zamore et al., 2000; Bernstein et al., 2001; Liu et al., 2003). We currently know that different RNAi mechanisms play important regulatory roles in development, maintenance of genome stability, gene expression and antiviral defence. These RNA-based mechanisms involve proteins of the Argonaute family combined with small regulatory RNAs ranging from 20–30 nt length (Treiber et al., 2019). Argonaute proteins associate with small RNAs that guide them towards target mRNAs leading to inhibition of their translation or direct cleavage catalysed by their RNAse-H like domain (Song et al., 2004; Ma et al., 2005; Yuan et al., 2005).

In insects, three RNAi pathways have been well documented, involving small RNAs of 21–22 nt (small interfering RNAs or siRNAs), 22–23 nt (micro RNAs or miRNAs) or 24–30 nt (Piwi-interacting RNAs or piRNAs). Whereas siRNAs and miRNAs are produced from dsRNA precursors by RNAseIII proteins such as Dicers and Drosha, piRNAs are processed independently of these enzymes (see section on piwi pathway below for more detail). Drosophila genetics defined two well-separated pathways involving Dicer-1/AGO1 for miRNAs and Dicer-2/AGO2 for siRNAs (Lee et al., 2004; Okamura et al., 2004). A different clade of AGO proteins involving PIWI, AGO3 and Aubergine (Aub) in Drosophila regulates the production and activity of piRNAs (reviewed in Huang et al., 2017).

miRNAs are essentially produced from nuclear precursors and participate in tight regulation of gene expression during development or cellular homeostasis (e.g. Posadas and Carthew, 2014). Of note, some DNA viruses use virus-encoded or cellular miRNAs to regulate their own gene expression or to modulate host cell transcriptome (Müller and Imler, 2007; Hussain and Asgari, 2014; see also Chapter 4). However, control of viral infections in
insects mostly relies on siRNAs and, in some cases might also involve piRNAs.

Production of small non-coding RNAs by Dicer enzymes

The antiviral RNAi pathway, the siRNA pathway, is triggered by long dsRNAs, which are processed in insects by Dicer-2 (Dcr-2). This cytoplasmic enzyme is composed of a N-terminal Duplex RNA activated ATPase (DRA) domain, a central double-stranded RNA-binding domain (dsRBD), a platform-PAZ domain, two RNAseIII domains and a C-terminal dsRBD (Fig. 3.2, left). Interestingly, Dicer enzymes share a common phylogenetic origin with the vertebrate RLRs, which sense viral nucleic acids in the cytosol and trigger synthesis of interferons. Although Dicer-2 and RLRs have different sizes and domain composition, they all contain a conserved DRA domain (Paro et al., 2015). Our understanding of the contribution of the DRA domain of Dicer-2 to virus sensing remains limited and is largely based on in vitro studies.

In vitro, Dicer-2 is efficient at processing long dsRNAs in addition to shorter structured RNAs such as pre-miRNAs which are bona fide targets of Dicer-1 (Cenik et al., 2011). In vivo, it appears that the activity of Dicer enzymes is regulated by cofactors containing two or three dsRBDs, e.g. R2D2 and Loquacious (Loqs) (Fig. 3.2, right). Dicer-2 forms a stable heterodimer with R2D2, which restricts its cellular localization to cytoplasmic D2 bodies (Nishida et al., 2013) and is mandatory for the efficient loading of siRNAs on AGO2 (Liu et al., 2003; Lee et al., 2004; Tomari et al., 2004a). Based on in vitro experiments, a role for R2D2 in preventing processing of pre-miRNAs by Dicer-2 has also been proposed (Cenik et al., 2011). However, this was not confirmed by in vivo experiments (Marques et al., 2013). Furthermore, inorganic phosphate has been proposed as a factor restricting the panel of targets for Dicer-2 in vitro, precluding it from processing aberrant pre-miRNA targets and short dsRNAs (Fukunaga et al., 2014). The other cofactor, Loqs has two main isoforms, Loqs-PB and Loqs-PD. Whereas Loqs-PB functions as a cofactor of Dicer-1 in the miRNA pathway (Forstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005), Loqs-PD enhances production of siRNAs by Dicer-2, especially from synthetic dsRNA and endogenously encoded dsRNAs derived from structured loci.

![Figure 3.2](https://caister.com/cimb/Curr-Issues-Mol-Biol-2020-Vol-34-caister.com-cimb)
sense–antisense pairs and transposable elements (Zhou et al., 2009; Marques et al., 2010; Miyoshi et al., 2010a).

Once produced, the siRNA duplex is loaded on the effector protein AGO2 in a highly coordinated but energetically unfavourable series of events (Fig. 3.3). In a first step, the Dcr-2/R2D2 complex, which displays intrinsically low affinity for duplex siRNAs, associates with TAF11, a TATA-box binding protein Associated Factor. Colocalized with Dicer-2 and R2D2 in D2 bodies, TAF11 acts as a chaperone facilitating the tetramerization of the Dicer-2–R2D2 heterodimer and increasing affinity for siRNAs by tenfold (Liang et al., 2015). R2D2 tends to preferentially bind the extremity of the siRNA duplex showing the strongest stability, creating an asymmetry in the complex (Tomari et al., 2004b). This asymmetry determines the preferential loading of the strand featuring the least stable 5' extremity in AGO2 to serve as guide siRNA (Fig. 3.3B and C).

**Structure–function of AGO proteins**

AGO proteins are composed of four globular domains named N, PAZ, MID and PIWI (Fig. 3.2, centre). They adopt a closed, flexible and unstable conformation not suitable to accept the siRNA duplex in their ‘Apo’ form. A chaperone machinery composed of the Hsp70 system (Hsp40+Hsp70) and Hsp90 system (Hop, Hsp90 and p23) is required for the efficient loading of siRNA duplexes. Briefly, Hsp70 opens the structure of AGO2 while Hsp90 is required to extend the duration of this opened state to allow sufficient time for the recognition of the 5’ phosphate at the extremity of the guide strand and subsequent loading of the entire duplex (Miyoshi et al., 2010b; Iwasaki et al., 2010, 2015; Tsuboyama et al., 2018). Both processes require ATP hydrolysis. The coordinated action of the heat shock proteins was proposed as the trigger for Dicer-2–R2D2 tetramer destabilization and transfer of the siRNA duplex (Fig. 3.3C). Intriguingly, Hsp proteins can be induced by viral infections in *Drosophila*. However, the siRNA pathway remains functional in flies mutant for the Heat Shock factor (Merkling et al., 2015b). In spite of this progress, the exact mechanism of AGO2 loading remains unclear especially because a 3D structure of the protein in its ‘Apo’ form is lacking. The semi-closed AGO2 protein loaded with the siRNA duplex is rigid and stable, and constitutes the pre-RNA Induced Silencing Complex (RISC). Studies with AGO proteins from eukaryotic (human) or prokaryotic (*Pyrococcus furiosus*) systems indicate that the strand of the duplex showing the less stable 5’ phosphate extremity is anchored to the phosphate binding pocket of the AGO2 MID domain while its 3’ extremity is bound to the hydrophobic cavity of the PAZ domain (Ma et al., 2004; Song et al., 2004; Fig. 3.3C). In the pre-RISC, the passenger strand occupies the same position as the future target RNAs (Kim et al., 2007). The full-length crystal structure of human AGO2 complexed with RNA reveals large structural differences between Argonautes from different kingdoms of life, even if individual domains superimpose reasonably well (Schirle and MacRae, 2012).

**Maturation and slicing**

After loading, the pre-RISC is matured through two essential steps required for downstream RNA silencing activity. First, the passenger strand is discarded. This is achieved by the coordinated action of the N-terminal domain of AGO2, acting as a wedge to unwind the siRNA duplex (Kwak and Tomari, 2012) and the PIWI domain RNAseH-like catalytic core, which cleaves the passenger strand in two small RNAs of 9 and 12 nt (Kim et al., 2007; Matranga et al., 2005; Miyoshi et al., 2005; Rand et al., 2005). The Component 3 Promoter Of RISC (C3PO) complex then helps AGO2 to get rid of the unstable cleavage products (Liu et al., 2009; Ye et al., 2011; Mo et al., 2018) (Fig. 3.3C).

A slicer-independent ejection model, relying on the thermal dynamics of the PAZ domain may also participate in discarding of the passenger strand (Gu et al., 2012; Park and Shin, 2015; Nakanishi, 2016). The final maturation step resides in the 2’O methylation of the 3’ extremity guide of the siRNA by the Hen1 enzyme (Horwich et al., 2007). This methylation step is crucial for the protection of the small guide RNA from 3’ uridylation and further 3’–5’ degradation (Li et al., 2005). The mature RISC, programmed with a guide RNA, functions as a Mg$^{2+}$-dependent, multiple-turnover enzyme that will recognize its mRNA target by perfect base-paired complementarity. The RISC will then slice the mRNA target and release the degradation product in an ATP assisted manner (Hutvágner and Zamore, 2002; Tang et al., 2003; Haley and Zamore,
Figure 3.3 The siRNA pathway in Drosophila melanogaster. (A) The siRNA pathway can be activated by diverse dsRNA substrates encompassing endogenous transcription of Natural Antisense Transcripts (NATs), transposons, convergent transcripts or structured RNA (hairpins). Experimentally delivered synthetic dsRNAs and virus intermediates of replication represent an exogenous source of dsRNA molecules and triggers RNAi. Blue and brown strands represent RNAs while black strands represent DNA. (B) Endogenous and exogenous long dsRNAs are sensed and diced by Dicer-2. Because Dicer-2 preferentially accesses its substrates at extremities, dsRNA termini are crucial determinants of the mode of action of Dicer-2. Indeed, while blunt end dsRNA triggers an efficient, DRA domain and ATP-dependent processive activity of Dicer-2, a dsRNA molecule with 3′ overhang termini promotes a slow, ATP-independent distributive activity. Generated siRNA duplexes are dsRNA molecules with a characteristic 2nt 3′ overhang signature and are 21nt long. The relative thermostability of siRNA extremities determines the orientation of the RNA duplex in Dicer-2. (C) Loading of AGO2 relies on a highly coordinated series of events. The closed structure of Apo-AGO2 is opened by two ATP-dependent chaperone complexes (Hsp70 and Hsp90). The siRNA duplex is transferred from the Dicer-2–R2D2–TAF11 complex to the opened form of AGO2, forming the pre-RISC. Maturation of the RISC includes cleavage and removal of the passenger strand. The 3′ extremity of the remaining guide strand is 2′O-methylated. (D) Mature RISC sequence specifically targets and cleaves complementary mRNA targets.
2004; Schwarz et al., 2004) (Fig. 3.3D). In the case of viruses, viral mRNAs seem to be preferentially targeted compared to the viral genome (Marques et al., 2013).

Control of viruses by the siRNA pathway

Genetic evidence in Drosophila

The discovery that small RNAs are produced in plants after viral infection predates the characterization of RNAi pathways in flies (Hamilton and Baulcombe, 1999). Evidence that RNAi operates against viruses in Drosophila was initially provided in S2 cells infected with Flock House virus (FHV) (Li et al., 2002). Of note, null mutants for the three main components of the siRNA pathway in flies, namely Dicer-2, AGO2 or R2D2 are homozygous viable (Liu et al., 2003; Lee et al., 2004; Okamura et al., 2004). As a result, adult mutant flies can be infected with viruses and both the survival rate and the viral load can easily be monitored. These experiments established that flies mutant for the siRNA pathway are susceptible to a variety of viruses with RNA or DNA genomes (e.g. Galiana-Arnoux et al., 2006; van Rij et al., 2006; Wang et al., 2006; Muller et al., 2010; Bronkhorst et al., 2012; Kemp et al., 2013). Interestingly, Dicer-2, R2D2 and AGO2 are among the 3% fastest evolving of all Drosophila genes highlighting the high selection pressure on the RNAi pathway (Obbard et al., 2006). Population genomic analysis in multiple invertebrates shows that RNAi genes display a greater rate of adaptive protein substitution than other genes, most likely reflecting their function at the forefront of defence against viruses and transposable elements (TEs) (Palmer et al., 2018).

Control of viruses by the siRNA pathway in other insects

Antiviral RNAi has been investigated in other insects, including vector insects. Many mosquito-borne viruses are associated with human and animal diseases, raising interest in mosquito antiviral immunity (Aguiar et al., 2016; Powers and Waterman, 2017). Over the past 50 years, the Aedes albopictus cell line C6/36, isolated from larvae, has been commonly used for amplification of arboviruses but also to study virus–vector interactions. Recently, the genome of C6/36 cells has been sequenced and null mutations in the dicer-2 gene were identified, which makes them incompetent for production of siRNAs (Morazzani et al., 2012; Miller et al., 2018). As a result, C6/36 cells support viral replication to high titres, confirming the important antiviral function of Dicer-2 in mosquitoes.

Injection of dsRNA in the body cavity efficiently silences gene expression in a sequence-specific manner and has been used in pioneer experiments to knock-down expression of components of the siRNA pathway in Aedes aegypti. This led to significantly increased SINV and DENV titres, without compromising insect survival (Campbell et al., 2008; Sánchez-Vargas et al., 2009; Khoo et al., 2010). Interestingly, in genotype–phenotype association studies, Lambrechts et al. (2013) found that the dicer-2 genotype is associated with resistance to DENV in a virus isolate-specific manner. By contrast, no such association is found for flanking loci, suggesting that the dicer-2 gene and the siRNA pathway are important determinants of virus suppression in Aedes. Mutants of the core RNAi components have subsequently been established in Ae. aegypti mosquitoes using genome editing approaches such as TALEN (Transcription Activator-Like Effector Nuclease) or CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) mutagenesis (Basu et al., 2015). In the context of viral infections, dicer-2 null mutant mosquitoes exhibit a decrease in survival rate compared to wild-type mosquito infection, that correlates with high Yellow fever virus titres (Samuel et al., 2016). The midgut epithelium is a critical barrier for viruses in insect vectors, as it becomes infected after acquiring an infectious blood meal from the host. After successfully replicating in the midgut, viruses reach the haemolymph and disseminate systemically. Viruses then reach the salivary glands from which they are transmitted to a naïve host upon blood feeding. An early study revealed that silencing dicer-2 in the midgut of female mosquitoes increased infection and dissemination of SINV (Khoo et al., 2010). Insects can also be vectors for plant viruses (Chapter 6). For example, the small brown planthopper (SBPH) is an incompetent vector for Southern rice black streaked dwarf virus (SRBSDV), a plant virus. SRBSDV is restricted in the midgut epithelium of SBPH. Knock-down of
either dicer-2 or AGO2 in SBPH results in the dissemination of the virus to the salivary glands and facilitates transmission to rice plants, revealing the importance of the siRNA pathway in the control of vector competence (Lan et al., 2016). However, the siRNA pathway was found to be largely dispensable for antiviral immunity in the midgut of Anopheles mosquitoes exposed to the O’nyong-nyong virus, and to become operative only during the systemic stage of infection (Carissimo et al., 2015). Interestingly, a recent study indicates that the siRNA pathway also fails to efficiently silence DENV in the midgut of Aedes aegypti, even though the canonical components of the pathway are expressed in this tissue, and that the pathway is functional when triggered by endogenous (e.g. control of TE by endo siRNAs) and exogenous (e.g. intrathoracic injection of long dsRNA) dsRNAs (Olmo et al., 2018). This discrepancy between antiviral and conventional siRNA pathways in the midgut results from the lack of expression in this tissue of Loqs2, an Aedes-specific parologue of Loqs and R2D2. Importantly, ectopic expression of Loqs2 in the midgut results in restriction of DENV replication and dissemination (Olmo et al., 2018). Altogether, these results point to an additional level of complexity in the insect siRNA pathway when it comes to the control of viruses (see below).

**siRNA as a footprint of antiviral immunity**

As mentioned above, virus-derived 21–22nt-siRNAs (vsiRNAs) produced in the course of viral infection in insects, which can be revealed by High Throughput Sequencing (HTS), provide an excellent read-out of Dicer-2 activity (Box 3.1). These vsiRNAs are strongly reduced or abolished in dicer-2 mutant flies (Aliyari et al., 2008; Muel ler et al., 2010; Bronkhorst et al., 2012; Kemp et al., 2013; Marques et al., 2013). For many RNA viruses, vsiRNAs cover the whole viral genome and the ratio between the number of siRNAs matching the (+) strand and the (–) strand of the genome is close to one (Aliyari et al., 2008; Myles et al., 2008; Box 3.1 High-throughput sequencing

The emergence of next-generation sequencing technologies set a milestone in the development of a tremendous number of high-throughput ‘omics’ approaches. Until 2010, RNA sequencing methods were all relying on sequencing by synthesis or semi-conductor technologies (second generation). The advent of the third generation of sequencing with PacBio and Oxford Nanopore opened the way to direct DNA and RNA sequencing without any amplification step. These ever-evolving technologies allow the detection of events that can only be bioinformatically predicted with second generation sequencing methods such as splicing or defective genomes in the case of viruses. However, to detect the specific signature of Dicer-2 (siRNA duplexes), small RNA sequencing HTS (second generation) remains the method of choice because of the high-throughput number of reads generated.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Limitations</th>
<th>Main technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>• High-throughput number of reads (4M – 20B)</td>
<td>• Library preparation bias including RT, PCR and optional size exclusion steps</td>
<td>• iSeq – NovaSeq 6000 (Illumina)</td>
</tr>
<tr>
<td>• Low error rate (~0.1 - 1%)</td>
<td>• Short reads (max ~400bp)</td>
<td>• Ion PGM (Life Technologies)</td>
</tr>
<tr>
<td>• Low cost</td>
<td>• Sample bleeding due to multiplexing (2-5%: check numbers)</td>
<td>• SOLID DNA sequencer (Thermo Fisher Scientific)</td>
</tr>
<tr>
<td>2nd generation</td>
<td></td>
<td>• 454 GS FLX (Roche)</td>
</tr>
<tr>
<td>• Direct RNA/DNA sequencing</td>
<td>• High error rate (5 – 15%)</td>
<td>• PacBio</td>
</tr>
<tr>
<td>3rd generation</td>
<td>• Long reads (up to 2Mb so far) allowing precise detection of splicing events for instance</td>
<td>• Oxford Nanopore</td>
</tr>
<tr>
<td>• Portability</td>
<td>• Low number of reads (50k – 1M)</td>
<td></td>
</tr>
</tbody>
</table>
Mueller et al., 2010; Marques et al., 2013; Ferreira et al., 2018), as would be expected from the processing of long dsRNA formed during viral replication. By contrast, in the case of DNA viruses, hotspots of vsiRNAs are observed on specific regions of the viral genome. These vsiRNAs match both strands of the genome, suggesting the siRNA pathway targets regions transcribed on both strands and producing dsRNA (Bronkhorst et al., 2012, 2013; Jayachandran et al., 2012; Kemp et al., 2013). Secondary structures in the genome or antigenome may also be targeted by Dicer-2 in some viruses (Sabin et al., 2013).

Based on in vitro experiments, AGO2 has been suggested to impair viral replication by cleaving the viral RNA through its slicer activity (van Mierlo et al., 2012a). In vivo experiments further suggest that viral mRNAs are the primary target of vsiRNA-loaded AGO2 (Marques et al., 2013). As genetics are still tricky in most non-model organisms, HTS provides a convenient readout to monitor activity of RNAi pathways in insects infected with viruses (e.g. Chejanovsky et al., 2014; Zografidis et al., 2015; Ferreira et al., 2018). For example, HTS of small RNAs isolated from SINV-infected Ae. aegypti mosquitoes revealed the presence of 21 nt-long vsiRNAs and brought the first evidence that the siRNA pathway is activated during viral infection in vector mosquitoes (Myles et al., 2008). This technique has also been successfully used for virus identification in both plants and insects (Kreuze et al., 2009; Wu et al., 2010). Indeed, large fragments of viral genomes can be reconstituted upon assembly of contigs from sequenced vsiRNAs. Furthermore, such contigs are enriched for viral sequences compared to long RNA sequencing reads, because they are by-products of the detection of viral replication by the insect immune system. Thus, HTS of small RNAs can be used to determine the virome of laboratory and wild populations of insects and possibly also other multicellular eukaryotes (Aguiar et al., 2015; Waldron et al., 2018). As such, HTS of small RNAs represents a powerful tool for virus surveillance in populations of vector insects.

**Viral suppressors of RNAi (VSRs)**

The study of viruses themselves can provide interesting insight into antiviral defence in insects. Indeed, in the course of their interaction with host cells, viruses have evolved to counter antiviral defence. Thus, many insect viruses, including *Drosophila* viruses, encode viral suppressors of RNAi (VSRs) (Bronkhorst and van Rij, 2014). Some VSRs (e.g. DCV-1A, FHV-B2, IVV-340R, *Drosophila* X virus VP3 and Culex Y virus VP3) directly bind long dsRNA through canonical dsRNA binding domains, dsRBDs and prevent processing by Dicer-2 (Li et al., 2002; van Rij et al., 2006; Bronkhorst et al., 2014; van Cleef et al., 2014; see also Fig. 3.4A and B). Interestingly, with the exception of DCV-1A, these VSRs bind siRNA duplexes as well, suggesting that they also inhibit the pathway after long dsRNAs have been processed into siRNAs (Morazzani et al., 2012; Valli et al., 2012; Bronkhorst and van Rij, 2014). Other VSRs (e.g. CrPV-1A and Nora-VP1) bind directly to AGO2 and inhibit its endonuclease activity (Nayak et al., 2010; van Mierlo et al., 2012b). In addition, CrPV-1A also targets AGO2 to the proteasome through the K48 polyubiquitination pathway (Nayak et al., 2018). This is reminiscent of the mode of action of P0, a VSR from poleroviruses, which triggers degradation of AGO1 in plant cells (Baumberger et al., 2007; Bortolamiol et al., 2007).

The importance of VSRs for viruses has been particularly well illustrated in the case of FHV-B2. *Nodaviridae* have small bipartite RNA genomes that are easy to manipulate genetically. One segment of the genome, RNA1, encodes the replicase, whereas the second, RNA2, encodes the capsid proteins. A third RNA transcript, RNA3, is also produced from RNA1 once it has replicated and encodes the VSR B2 (Chao et al., 2005). Whereas wild-type FHV is highly pathogenic upon injection into the body cavity of flies, viral mutants unable to express B2 are completely attenuated (Galiana-Arnoux et al., 2006; Wang et al., 2006; Fig. 3.4C). As expected, the virus regains virulence when injected into *dicer-2* or AGO2 mutant flies (Han et al., 2011; Petrillo et al., 2013). Similarly, mutation of residue Phe 114 into Ala in the flexible loop of CrPV-1A, which is involved in the interaction with AGO2 results in an attenuated virus in wild type flies but not in AGO2 mutant flies (Nayak et al., 2018). Of note however, even in RNAi mutants, the B2 deficient virus exhibits reduced virulence compared to wild-type FHV. Indeed, an additional function of B2 is to bind double stranded regions of RNA2, thus preventing its recruitment into poorly characterized cytoplasmic RNA granules where its
translation would be repressed. As a result, translation of the capsid protein is impaired in the absence of B2, even in RNAi deficient cells (Petrillo et al., 2013). Interestingly, the FHV-B2 VSR can be used to neutralize the siRNA pathway upon ectopic expression. For example, when Ae. aegypti mosquitoes were challenged with SINV recombined with FHV-B2 (SINV-B2) either by injection or infected blood meal, the virus titre was increased compared to the control virus. The recombinant SINV-B2 caused high mortality among the mosquitoes at 4–6 days post-infection, highlighting that the RNAi pathway is essential to control arboviruses replication (Myles et al., 2008; Cirimotich et al., 2009). In Ae. aegypti mosquitoes, the constitutive and ubiquitous expression of FHV-B2 impaired the siRNA pathway and enhanced replication of both SINV and DENV (Khoo et al., 2013).

One consequence of the direct interaction of some VSRs with protein factors of the siRNA pathway to alter their ability to neutralize viral RNAs or trigger their degradation (Singh et al., 2009; Nayak et al., 2010, 2018; van Mierlo et al., 2012b, 2014) is that both VSRs and components of the siRNA pathway evolve rapidly (Obbard et al., 2006). As a result, the activity of viral suppressors can be host-specific. For example, the VP1 protein from a divergent Nora virus isolated from D. immigrans interacts with and suppresses D. immigrans AGO2, but not D. melanogaster AGO2 (van Mierlo et al., 2014). This provides an excellent example for the co-evolution of the host RNAi machinery and viral suppression mechanisms. Because of their intimate association with key components of the siRNA pathway, VSRs provide promising tools to decipher the regulation and molecular mechanisms of antiviral RNAi. For example, development of single-molecule approaches can shed light on VSR mode of action and how they discriminate viral from cellular RNA (Fareh et al., 2018). VSRs may also be used to visualize and track dsRNA in live plant and animal cells, as recently shown for an FHV-B2–GFP fusion protein (Monsion et al., 2018).

Systemic RNAi in insects?
In both plants and C. elegans, systemic RNAi contributes to the control of viral infections. The mechanism at play involves spreading of siRNAs generated in infected cells to neighbour healthy cells. There, these siRNAs prime the synthesis of dsRNAs by host-encoded RNA-dependent RNA polymerases (RdRPs) (Ding, 2010). Insect genomes do not encode such RdRPs and clonal analyses in Drosophila revealed that the siRNA pathway is a cell autonomous pathway (Roignant et al., 2003). Nevertheless, the spread of antiviral RNAi has been proposed to contribute to the control of viral infections in insects. Indeed, injection of exogenous dsRNA in the body cavity of most insects, or even feeding dsRNA, leads to gene silencing through the siRNA pathway (reviewed in Zotti et al., 2018). Thus, viral infection, which is known to trigger transcriptional responses, may induce a mechanism of systemic RNAi. In Drosophila, uptake of exogenous dsRNA is mediated by the endocytic pathway (Saleh et al., 2006). Nanotube like structures have also been reported to transfer dsRNA and components of the RNAi machinery between cells (Karlikow et al., 2016). Thus, dsRNA released from infected dying cells may trigger RNAi in distant cells, upon internalization by the dsRNA uptake pathway (Saleh et al., 2009). Surprisingly, however, infection of flies with a sublethal dose of DCV, which should prime the antiviral siRNA pathway, did not induce protection against a challenge with a lethal dose of virus (Longdon et al., 2013). Of note, viral RNA can be reverse transcribed into DNA (vDNA) in Drosophila and in mosquitoes, through the action of the reverse transcriptase from transposable elements (TEs) (Goic et al., 2013, 2016), acting together with Dicer-2 (Poirier et al., 2018) or AGO2 (Tassetto et al., 2017). Transcription of vDNA has been proposed to result in the production of secondary siRNAs bearing a 5’-triphosphate mark, associated with systemic antiviral effect (Tassetto et al., 2017). However, the existence of such secondary siRNAs was not confirmed by an independent study (Mondotte et al., 2018). Overall, the mechanisms involved in systemic antiviral RNAi remain to be characterized and genetic evidence for the importance of the contribution of vDNA and secondary siRNAs in antiviral immunity is still lacking.

In summary, it now appears that the antiviral siRNA pathway in insects is more complex than previously thought. A number of host and viral factors, which affect the stability, the binding affinity to viral nucleic acids, and even the cytoplasmic localization of ribonucleoprotein complexes mediating RNA silencing can influence antiviral defence. Biochemistry, genetics and live imaging will clarify the mode of action of the components of the siRNA pathway in the context of infected cells, in different types of tissues and in various insects.

The piRNA pathway in antiviral immunity

Production of primary and secondary piRNAs
Another small RNA pathway, the piRNA pathway is able to sense foreign nucleic acids and has been proposed to participate in antiviral immunity (Miesen et al., 2016a). This pathway involves 24–30 nt small RNAs and Argonaute proteins from the PIWI clade. The founding member of this subfamily, PIW1, was initially characterized as a Drosophila gene essential for male fertility (Lin and Spradling, 1997). This phenotype results from derepression and mobilization of transposable elements in the germline. It was subsequently found that PIWI works with two other Argonaute proteins from the same clade, namely Aub and AGO3 (Saito et al., 2006; Vagin et al., 2006; Brennecke et al., 2007; Gunawardane et al., 2007). The PIWI-mediated mechanism of genome maintenance in the germline is conserved in all animals. The piRNAs were initially identified in mouse testis and in Drosophila, where they were first called repeat-associated small interfering (rasi)RNAs. They interact with PIWI proteins (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Vagin et al., 2006). Drosophila genetics coupled with HTS, have shed light on the molecular mechanism generating piRNAs. This involves a primary processing pathway that primes the production of phased piRNAs during the so-called ping-pong amplification mechanism (Czech and Hannon, 2016; Fig. 3.5). Of note, HTS of cell lines or mosquitoes revealed the existence of virus-derived piRNAs (Wu et al., 2010; Hess et al., 2011; Morazzani et al., 2012; Vodovar et al., 2012; Léger et al., 2013; Aguiar et al., 2015).
The primary processing pathway targets Pol II-dependent transcripts generated from genomic loci rich in transposon remnants known as piRNA cluster (Bucheton, 1995; Brennecke et al., 2007; Pélisson et al., 2007). These precursors are processed by the endonuclease Zucchini, which preferentially cleaves 5′ of a uridine residue. As a result, the piRNA intermediates produced are enriched for 5′ uridine residues (1U) (Pane et al., 2007; Han et al., 2015; Mohn et al., 2015; Fig. 3.5A). In fruit flies, these piRNA intermediates are loaded onto PIWI and Aub in an electron dense perinuclear region of the germline called ‘nuage’.

Indeed, the binding pocket of the MID domain of Hen antisense piRNA RNA Pol II piRNA cluster
these two proteins preferentially accommodates 5’ uridine residues (Cora et al., 2014; Matsumoto et al., 2016; Fig. 3.5B). piRNA-loaded PIWI translocates to the nucleus where it participates in transcriptional silencing of transposons through deposition of repressive histone modifications (Sienski et al., 2012; Dönertas et al., 2013; Le Thomas et al., 2013; Ohtani et al., 2013; Sienski et al., 2015; Yu et al., 2015; Fig. 3.5C). By contrast, piRNA-loaded Aub remains in the nuage and initiates the ping-pong amplification mechanism (Fig. 3.5D). The Aub pi-RISC recognizes and cleaves complementary transposon mRNAs. The resulting cleavage product corresponds to the precursor of a secondary piRNA and associates with AGO3 (Brennecke et al., 2007; Gunawardane et al., 2007; Lim and Kai, 2007). Of note, because cleavage mediated by enzymes of the PIWI clade occurs specifically between nucleotide 10 and 11, AGO3 bound piRNAs are enriched for adenosine residues in position 10 (10A). piRNAs loaded into AGO3 target and cleave antisense piRNA precursors, thus generating the 5’ end of new sense piRNAs and resulting in the ping-pong amplification cycle (Fig. 3.5D). In summary, the combination of slicer activity of PIWI proteins together with the activity of endo- and exonucleases explains the Dicer-independent production of 24–30 nt long piRNAs. Of note, the 1U/10A signature characteristic of the ping-pong amplification provides a convenient way to monitor the activity of this pathway by HTS and was instrumental in revealing the existence of virus-derived piRNAs.

Virus-derived piRNA

Transposable elements share with viruses the property of being selfish genetic units encoding proteins that enable their proliferation and spread. Hence, one can wonder whether the piRNA pathway also participates in antiviral immunity. Indeed, transposable elements are targeted by the siRNA pathway in Drosophila somatic tissues as revealed by accumulation of siRNAs matching transposable elements in HTS analysis (Chung et al., 2008). Conversely, virus-derived piRNAs could be observed in the OSS cell line derived from Drosophila ovarian tissue (Wu et al., 2010). However, in Drosophila, activity of the piRNA pathway is restricted to the germline and neither genetics nor HTS support an antiviral function of the piRNA pathway in Drosophila (Petit et al., 2016; van den Beek et al., 2018).

Nonetheless, there are significant differences between piRNA pathways in Drosophila and other insects. Indeed, a recent study investigating 20 species across the arthropod phylum revealed that TEs are commonly targeted by somatic piRNAs unlike in Drosophila (Lewis et al., 2018). Consistent with this observation, the piRNA pathway components differ between insect species. Notably, it is apparent that the PIWI clade has significantly expanded in Ae. aegypti mosquitoes, where it contains seven PIWIs (instead of two in Drosophila, Piwi and Aub) and one AGO3 (reviewed in Miesen et al., 2016a). This, together with identification of virus derived piRNAs in mosquito Aag2 and C6/36 cell lines, led to the suggestion that the piRNA pathway could form a second layer of antiviral defence in mosquitoes (Wu et al., 2010; Hess et al., 2011; Morazzani et al., 2012; Vodovar et al., 2012; Léger et al., 2013; Aguiar et al., 2015). Depletion in AGO3 and PIWI5 resulting in the decreased production of viral piRNAs revealed that these enzymes mediate recognition and processing of SINV RNAs, although the functional consequence of this processing on viral replication was not reported (Miesen et al., 2015). A subsequent report focusing on DENV indicated that the knockdown of PIWI proteins did not significantly affect viral RNA levels (Miesen et al., 2016b). An independent study proposed a role for PIWI4 in antiviral immunity in Aag2 cells, an Aedes aegypti derived cell line. Interestingly, PIWI4 behaves as an atypical member of the PIWI family and is not involved in piRNA production but associates with the siRNA pathway core components AGO2 and Dicer-2 (Varjak et al., 2017a,b, 2018). This protein may reveal a cross-talk between piRNA and siRNA pathways in the context of viral infections. The presence of virus derived somatic piRNAs in other arthropods was tested by Jiggins and colleagues in their landmark paper (Lewis et al., 2018) and virus derived siRNAs could be identified in 9 of the 20 species investigated. Among them, 5 species also produced 24–30 nt 5’U biased small RNAs derived from viruses. However, only in Aedes aegypti did these piRNAs bare the 1U/10A signature of ping-pong amplification. In the four other species, piRNAs mapped to one strand only, similar to primary piRNAs (Lewis et al., 2018). Therefore, the predominant role of siRNAs compared to piRNAs in antiviral defence observed in Drosophila and Lutzomyia is probably relevant across...
Sensing viral RNA in insects

Viral nucleic acid sensors

In mammals a major molecular pattern associated with viral infection is long dsRNA, generated as a by-product of viral replication. The Toll receptor TLR3 is localized in the endosome compartment and is activated upon binding of long dsRNA (Kawai and Akira, 2011). This receptor probes the content of the endosomes for the presence of endocytosed signs of viral infection in the extracellular milieu. Cross-linking of two TLR3 subunits by dsRNA triggers TRIF-dependent interferon activation. Of note, two other endosomal TLRs, TLR7 and TLR8 detect UG rich short ssRNA fragments (Maeda and Akira, 2016). In the cytosol another family of pattern recognition receptors, the RIG-like receptors, sense viral RNA (Goubau et al., 2013). As mentioned above, a hallmark of these receptors is the presence of a DRA domain phylogenetically related to the Dicer proteins (Paro et al., 2015). This domain is followed by a C-terminal domain (CTD), which participates in RNA binding. In addition, two of the three RLRs, RIG-I and MDA5 contain amino-terminal caspase recruitment domains (CARDs) (Kawai and Akira, 2011). This enables them to recruit the signalling adaptor MAVS and to activate expression of interferon genes. Although the three RLRs can bind dsRNA in vitro, other molecular features found on viral RNA are necessary to activate RIG-I. Indeed, the CTD of RIG-I, which forms a tighter pocket than the one from MDA5, detects the presence of 5′ di- or triphosphate at the extremity of viral RNAs (Kowalinski et al., 2011; Goubau et al., 2014). This provides an efficient means to discriminate between capped cellular mRNAs and uncapped viral RNAs. This biochemical distinction is supported by genetic data that point to non-redundant functions of RIG-I and MDA5. Indeed, RIG-I mutant mice have impaired interferon responses following infection by viruses such as influenza, VSV, or Japanese encephalitis virus (Kato et al., 2005, 2006). Of note, most of these viruses have ssRNA genomes of negative polarity and do not produce detectable amounts of long dsRNA in infected cells (Weber et al., 2006). By contrast, RIG-I mutant mice respond normally to picornaviruses. The S′ end of the genome and antigenome strand of these viruses is covalently linked to a VpG protein, which prevents recognition of the termini by RIG-I. On the other hand, MDAS mutant mice are highly susceptible to picornaviruses, which generate large quantities of dsRNA (Kato et al., 2006). The third RLR, LGP2, functions together with MDA5 (Dedouche et al., 2014).

In insects, the only viral nucleic acid sensor identified so far is Dicer-2, suggesting that long dsRNA is the major molecular pattern used to detect viral infection (Fig. 3.3A). Importantly, even in the case of the negative strand RNA virus VSV, the profile of virus-derived siRNAs reveals a typical long dsRNA signature, with siRNAs covering in equal amounts the whole length of both genome and antigenome strands (Ferreira et al., 2018; Marques et al., 2013; Mueller et al., 2010). However, sensing of viral RNA by Dicer-2 appears not to be limited only to the recognition of dsRNA.

Sensing viral RNA by Dicer-2

In vitro dicing assays using recombinant versions of Dicer-2 with or without its cofactors incubated with diverse dsRNA showed that Dicer-2 preferentially accesses its substrate at the extremities (Sinha et al., 2015). Furthermore, dsRNA termini are crucial determinants of Dicer-2 mode of action in vitro (Fig. 3.3B). Indeed, while blunt end dsRNA triggers an efficient, DRA domain and ATP-dependent processive activity of Dicer-2, a dsRNA molecule with 3′ overhang termini promotes a slow, ATP-independent distributive activity (Cenik et al., 2011; Welker et al., 2011). Cryo-EM studies revealed that dsRNAs presenting 3′ overhangs are repeatedly bound by Dicer-2 platform-PAZ domain to be sequentially cleaved while blunt dsRNAs are threaded through the DRA domain in an ATP-dependent manner and successively diced to generate many phased siRNAs (Sinha et al., 2018a). This distinct mode of dicing is driven by an extensive conformational change of the protein upon binding of a blunt dsRNA extremity, which cannot occur if the extremity harbours a 3′ overhang (Fig. 3.3B). Loqs-PD interacts with the DRA domain of Dicer-2 in an RNA independent manner through its C-terminal FDF motif and allows the enzyme to process RNA substrates normally refractory.
to cleavage, such as dsRNA with blocked, structured or frayed ends (Sinha et al., 2015; Trettin et al., 2017). In summary, in vitro experiments with purified recombinant proteins point to a model where subtle changes in the substrate result in tremendous differences in the dicing mechanism. How this model can be reconciled with the complexity of Dicer-2 natural substrates in the cellular context is an important challenge for future studies. This is particularly relevant in the context of viral infection as the RNA extremities of viruses and viral replication complexes are highly variable. A striking example is the case of picornaviruses and dicistroviruses, which, as mentioned above, display a covalently linked VpG protein at the 5′ extremities of their genome and antigenome (Virgen-Slane et al., 2012). This covalent modification at the extremities of the genome from *Dicistroviridae* such as DCV or CrPV is expected to impact sensing by Dicer-2. One asset to solve this important question is the characteristic siRNA signature of Dicer-2, which is amenable to bioinformatic analysis following small RNA HTS. Bioinformatic analysis of the pattern of vsiRNAs produced in wild-type or mutant flies (e.g. inactivated ATP binding site in the DRA domain of Dicer-2) may provide insights on an alternative access point of Dicer-2 on DCV or CrPV RNA (see Box 3.1) (e.g. Aliyari et al., 2008; Mueller et al., 2010; Bronkhorst et al., 2012; Sabin et al., 2013; Aguiar et al., 2015).

Another indication that features other than double strandedness are sensed lies in the importance of the cofactors acting together with Dicer-2. For example, Loqs-PD is not required for silencing viral RNA, in spite of its essential role in RNA interference triggered by endogenous or in vitro synthesized dsRNA (Marques et al., 2010, 2013). This points to the existence of differences between viral RNAs and other dsRNAs, produced from endogenous sources or synthesized in vitro. Of note, similar observations were made in the nematode *Caenorhabditis elegans*, where a RIG-I-like factor known as DRH1 is essential for antiviral RNAi yet dispensable for the other silencing pathways (Ashe et al., 2013; Guo et al., 2013), but also in *Ae. aegypti* mosquitoes. As mentioned before, the siRNA pathway silences targets of endogenous and exogenous dsRNAs in the midgut of these mosquitoes but fails to suppress viruses (Olmo et al., 2018). These results confirm the existence of differences between silencing triggered by exogenous or endogenous dsRNA and viral RNA, with the latter specifically requiring the dsRBP Loqs2. We note that an isoform of Staufen that evolved in coleopterans is another example of a species-specific dsRBP that regulates the siRNA pathway (Yoon et al., 2018). The fact that *Drosophila* (Loqs-PD), *Aedes* (Loqs2) and coleopteran insects (StaufenC) all require specific cofactors to define the activity of siRNA pathway is intriguing and highlights that an important facet of this pathway remains ill characterized.

The discrimination between the dsRNA precursors of endo- or exo-siRNAs and the dsRNA generated during viral infection may reflect differences either in the receptor complex sensing dsRNAs or in the viral RNA itself. On the protein side, it will be important to characterize biochemically the role of Loqs and R2D2 proteins in *Drosophila* and *Aedes*. On the RNA side, the persuasiveness of RNA modifications begs the question of how much they contribute to the discrimination of viral RNA by the siRNA pathway (Gokhale and Horner, 2017; Helm and Motorin, 2017; Fig. 3.6). For example, ADAR is an RNA editing enzyme catalysing the conversion of A to I within dsRNA regions of cellular RNAs, which prevents unwanted activation of immunity in mice and *C. elegans* (Lidicé et al., 2015; Reich et al., 2018). In mammals, the presence of epitranscriptomic marks such as 2′O-methylation on viral RNA affects sensing by TLR7 or MD2 whereas N6-methyladenosine (m6A) in dsRNA reduces activation of TLR3 (Karió et al., 2005; Gonzales-van Horn and Sarnow, 2017). Therefore, it will be interesting to investigate the presence of post transcriptional modifications in viral RNAs in insect cells and their impact on antiviral RNAi. This will be particularly interesting in the case of flaviviruses such as dengue and Zika viruses, which have been shown to contain m6A modified nucleosides when grown in mammalian cells (Gokhale et al., 2016; Fig. 3.6).

**Future directions**

A great deal of progress was made over the past decade on the genetic characterization of antiviral innate immunity in insects. Yet, lots of questions remain, paramount among them the identification of the receptors that sense viral infections. Among insects, the fruit fly *D. melanogaster* offers a fantastic
model for unbiased, large scale, mutagenesis screens (Wieschaus and Nüsslein-Volhard, 2016). Indeed, such screens were instrumental in defining the components of the IMD and Toll pathways of innate immunity, and in identifying the PRRs activating them in response to bacterial or fungal infections (e.g. Leulier et al., 2000; Rutschmann et al., 2000a,b; Lu et al., 2001; Vidal et al., 2001; Choe et al., 2002; Gottar et al., 2002, 2006). Genetic screens are however time consuming, especially when viruses must be injected into the body cavity of the flies (Merkling and van Rij, 2015). This limitation could, to some extent, be bypassed by using natural infections with fly pathogens, although this would require production of large amounts of virus and the control of viral uptake would be challenging. Alternatively, transgenic viral replicons expressing fluorescent proteins could be used as a proxy for infection, with the caveat that critical steps of the viral cycle (binding, entry, uncoating, assembly and budding of viral particles) would be bypassed (Avadhanula et al., 2009; Wernet et al., 2014). The recent advances in mass spectrometry (MS) technologies provide other opportunities to decipher antiviral innate immunity in insects. These methods can be used to further characterize the antiviral siRNA pathway by defining the interactome of the canonical components Dicer-2, R2D2 and AGO2, but also to identify PRRs sensing nucleic acids.

The genetic characterization of antiviral RNAi in both Drosophila and Aedes mosquitoes points to differences between the siRNA-dependent response triggered by dsRNA of endogenous (endo-siRNA pathway) or exogenous (exo-siRNA pathway) origin on one hand, and the antiviral siRNA pathway on the other (Marques et al., 2013; Olmo et al., 2018). Cell imaging of the canonical components Dicer-2, AGO2, R2D2 and Loqs in the context of cells treated with exogenous dsRNA or infected with viruses is likely to reveal differences in the dynamics of these proteins. This information can subsequently be used to define the interactome of the siRNA pathway at critical steps of the infection cycle, to identify novel regulatory co-factors of the pathway participating in the sensing or neutralization of viral RNAs. Of note, this approach requires synchronized infections, and will have to be carried out in tissue culture cells, a caveat considering the possible existence of tissue specific regulators (e.g. Olmo et al 2018).

MS can also be used to identify host proteins that sense nucleic acids of viral origin. Some foreign nucleic acids bear specific marks absent in cellular nucleic acids, such as, in the case of RNA, long double strandedness, the presence of

---

**Figure 3.6** Overview of cellular and viral RNAs. Schematic representation of self and non-self RNAs adapted from Gebhardt et al. (2017). 2′O methylation, Inosine and N6-methyladenosine are the three main RNA modifications described on viral RNAs affecting their recognition by the innate immune system.
immunity in insects. Protein binding to these marked, viral-like nucleic acids can be identified by nucleic acid affinity purification coupled to MS-based identification of proteins. For example, such an approach in mammals resulted in the identification of the host proteins IFIT1 and AIM2 as nucleic acid binding proteins involved in antiviral function and a DNA sensor of the innate immune system, respectively (Bürckstümmer et al., 2009; Pichlmair et al., 2011). Applied to insect cell lines or whole animals, this strategy, coupled to RNAi or CRISPR/Cas9 functional screens may reveal novel nucleic acid binding proteins involved in antiviral immunity in insects.

Finally, a fascinating question for future work pertains to the sensing of DNA in insect cells. Indeed, activation of the IMD pathway in flies mutant for the enzyme DNaseII (significant but moderate compared to a bacterial infection) suggests that cytosolic DNA can activate an NF-κB-dependent pathway in insect cells (Mukae et al., 2002; Liu et al., 2012). In mammals, a dedicated pathway, the cGAS-STING pathway, activates synthesis of interferons upon detecting cytosolic DNA (Hornung et al., 2014). So far, the data available indicate that viral DNA produces dsRNAs and activation of the siRNA pathway, akin to the sensing of herpes virus derived dsRNA by TLR3 in mammals (Tabeta et al., 2004; Zhang et al., 2007; Bronkhorst et al., 2012; Kemp et al., 2013). Yet, important questions remain. For example, is viral DNA, which is expected to be transcribed in the nucleus possibly by a host RNA polymerase, activating the endo-siRNA pathway (Loqs-PD dependent) or the antiviral siRNA pathway (Loqs-PD independent)? Is there another pathway sensing DNA in the cytosol, as in mammals? The cGAS-like enzymes identified in insect genomes so far do not contain the zinc ribbon motif mediating interaction with DNA in mammalian cGAS (Wu et al., 2014; Margolis et al., 2017). However, the recent discovery that STING carries immune functions in some insects, in particular against DNA viruses in Bombyx mori, raises the possibility that dedicated receptors operate in the cytosol of insect cells to sense DNA (Goto et al., 2018; Hua et al., 2018; Liu et al., 2018; Martin et al., 2018). Characterization of the ligands activating the insect STINGs now at hand, which do not appear to bind cyclic dinucleotides (Kranzusch et al., 2015), should help to identify these receptors.

Acknowledgements

We thank Ludmila dos Santos Silva and Joao Marques for critical reading of the manuscript and insightful suggestions. Work in our laboratory is supported by the CNRS, the ANR, the Fondation ARC and the Programme Investissements d’Avenir (Labex NetRNA and Equipex I2MC). LT acknowledges financial support from the Hoffmann Infinitus Program.

References


Immunol. 27, 357–368. https://doi.org/10.1016/j.smim.2016.03.010


Host Microbe 24, 57–68.e3 https://doi.org/10.1016/j.chom.2018.05.022


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

Schirle, N.T., and MacRae, I.J. (2012). The crystal structure of human Argonaute2. Science 336, 1037–1040. https://doi.org/10.1126/science.1212151


Wong, M.T., and Chen, S.S. (2017). Hepatitis C Virus Subverts Human Choline Kinase-a To Bridge Phosphatidylinositol-4-Kinase IIIα (PI4KIIIα) and


