Caractérisation des interactions entre les défenses antivirales et le contrôle génomique des éléments transposables chez Drosophila

Marlène Roy

▶ **To cite this version:**

Marlène Roy. Caractérisation des interactions entre les défenses antivirales et le contrôle génomique des éléments transposables chez Drosophila. Virologie. Université de Lyon, 2019. Français. NNT : 2019LYSE1126 . tel-02943888

HAL Id: tel-02943888
https://theses.hal.science/tel-02943888

Submitted on 21 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N°d'ordre NNT : 2019LYSE1126

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale ED341
Evolution Ecosystème Microbiologie Modélisation

Spécialité de doctorat : Infections virales, éléments transposables, petits ARN interférents
Discipline : Virologie, Génomique

Soutenue publiquement le 20/09/2019, par :

Marlène Roy

Caractérisation des interactions entre les défenses antivirales et le contrôle génomique des éléments transposables chez *Drosophila*

Devant le jury composé de :

Mouchiroud, Dominique ; Professeure – UCBL	Présidente
García Guerreiro, Maria Pilar ; Professeure – Universitat de Barcelona	Rapporteure
Dupressoir, Thierry ; Directeur d'études EPHE – Université Montpellier 2	Rapporteur
Gilbert, Clément ; Chargé de Recherche CNRS – Université Paris-Sud	Rapporteur
Brasset, Emilie ; Maître de Conférences – GReD - Université Clermont-Ferrand	Examinatrice
Vavre, Fabrice ; Directeur de Recherche CNRS – UCBL	Examinateur
Fablet, Marie ; Maître de Conférences – UCBL	Directrice de thèse
Ratinier, Maxime ; Maître de Conférences – EPHE	Co-directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directeur Générale des Services	M. Damien VERHAEGHE

COMPOSANTE SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G. RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directrice : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. SEUX
Institut des Sciences Pharmaceutiques et Biologiques	Directrice : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directrice : Mme la Professeure A-M. SCHOTT

COMPOSANTE ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directrice : Mme. la Professeure K. GIESELER
Département Chimie Biochimie	Directrice : Mme C. FELIX
Département GEP	Directeur : Mme. la professeure R. FERRIGNO
Département Informatique	Directeur : M. le Professeur B. SHARIAT
Département Mathématiques	Directeur : M. le Professeur I. BEN YAACOV
Département Mécanique	Directeur : M. le Professeur M. BUFFAT
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : Mme. Y. VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. I. DANIEL
Polytech Lyon	Directeur : M. le Professeur E. PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur P. CHAREYRON
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Ce travail a été réalisé au sein de deux laboratoires de recherche :

- le laboratoire de Biométrie et Biologie Evolutive (LBBE) = unité mixte de recherche UMR, CNRS 5558, dans l'équipe « Elements transposables, Evolution, Populations », dirigée par la Pr. Cristina VIEIRA,
- le laboratoire des Infections Virales et Pathologie Comparée (IVPC) = unité mixte de recherche UMR 754, dans l'équipe « Biologie des Phlébovirus », dirigée par le Dr. Frédérick ARNAUD,

Sous la co-direction des Dr. Marie Fablet et Maxime Ratinier.

REMERCIEMENTS

Je tiens tout d'adord à remercier l'ensemble des membres de mon jury d'avoir accepté de participer à ma soutenance de thèse afin d'évaluer mes travaux de recherche. Merci à **Thierry Dupressoir**, **Maria Pilar Garcia-Guerreiro** et **Clément Gilbert** d'avoir accepté d'être rapporteurs de mon manuscrit, et à **Emilie Brasset**, **Fabrice Vavre** et **Dominique Mouchiroud** d'avoir accepté d'être examinateurs de ma soutenance de thèse. Je remercie l'ensemble des membres de mon jury de thèse pour leur intérêt et les discussions très enrichissantes que nous avons eu le jour de ma soutenance. J'ai été très heureuse de partager cette journée inoubliable avec vous.

Je tiens à remercier particulièrement mes directeurs de thèse, tout d'abord pour m'avoir donnée l'opportunité de réaliser cette thèse. Merci **Marie** pour ta bienveillance, ton optimisme, ton expertise scientifique et ta pédagogie, qui fait que c'est toujours très agréable de travailler avec toi. Tu m'as appris beaucoup et je t'en remercie. J'ai apprécié tous nos échanges et je suis vraiment reconnaissante de ton implication, ton écoute et ton aide. Mille mercis Marie ! Merci **Maxime** également pour ta grande implication, ta disponibilité et pour tout ce que tu m'as enseigné. Je me rappelerai toujours de mon entretien de thèse où tu étais très enthousiaste et bienveillant. Merci à toi pour m'avoir toujours poussée à aller plus loin et à faire de mon mieux. Merci de m'avoir permis de travailler sur le projet Toscana qui m'a beaucoup plus, et de m'avoir permis de finir ma thèse dans de meilleures conditions avec le poste ATER que j'ai obtenu grâce à toi.
Merci à tous les deux pour nos interactions très intéressantes et pour m'avoir permis de réaliser une thèse en co-direction dans de bonnes conditions. Je vous suis très reconnaissante pour l'ensemble de ces raisons et vous souhaite une bonne continuation professionnelle mais aussi personnelle.
Je tiens également à dédier ma thèse à **Christophe Terzian**, qui avec Marie est à l'origine de ce projet et avec qui je n'ai malheureusement pas eu la chance de partager mes travaux. Mon implication a été encore plus grande pour atteindre les objectifs qui avait été proposés.

J'aimerais remercier les membres de mon comité de thèse, **Séverine Chambeyron**, **François Leulier**, **Philippe Marianneau** et **Tristan Lefebure** pour nos discussions très intéressantes qui m'ont vraiment aidé au cours de ma thèse.

Je tiens à dire merci particulièrement à chacune des personnes suivantes.

Aux membres du **laboratoire IVPC** (Infections Virales et Pathologie Comparée) où j'ai passé la plupart de mon temps. A **Fabienne** qui dirige l'unité UMR 754 avec efficacité et bienveillance.

A **l'équipe Phlébovirus**. **Fred**, merci de m'avoir accueillie dans ton équipe, et de m'avoir permis de travailler sur le projet Toscana. Merci pour tes conseils et ton esprit critique permettant l'optimisation des recherches.

A **Barbara V**, ma coach de droso, qui m'a appris à les aimer. Merci pour ton aide et ton expertise en drosophile, je suis très réconnaissante de cela car en plus de m'avoir appris beaucoup de choses, tu m'as également transmis ton intérêt particulier pour ce domaine de recherche. Merci aussi de m'avoir écoutée et soutenue de nombreuses fois, pour les moments heureux de ma vie comme pour les moments moins faciles. Merci également de m'avoir fait découvrir le full-contact qui a été une grande source d'apaisment pour moi.

A **Sophie** pour toutes nos discussions, pour ta bienveillance et ton intérêt à mon égard. J'ai vraiment apprécié ta curiosité par rapport à mon projet de thèse qui, c'est vrai, me tenait quelques fois à l'écart de l'équipe. J'ai également été contente de travailler avec toi pour le projet Toscana.

Marie-Pierre, alias MPi, merci mille fois pour ta bienveillance et ton bon état d'esprit de travail. Merci d'avoir pris soin de moi, pour ton aide, ta gentillesse et ta disponibilité. Sache qu'une personne comme toi, c'est rare ! Je t'en suis reconnaissante car cela a eu impact important sur mon bien-être au labo.

A **Saw-See** pour ta bonne humeur et ton intérêt pour mon projet de recherche.

A **François** d'avoir été là au début de ma thèse car j'ai vraiment apprécié nos discussions scientifiques ou non. Merci de m'avoir initiée à la culture de cellules d'insectes. En espérant que tu profites bien de ta retraite.

Aux étudiants qui ont été là pendant ma thèse, particulièrement **Justine** avec qui le feeling est tout de suite très bien passé, je suis contente de t'avoir rencontrée et espère continuer à rester en contact ; et **Etienne** pour ton investissement et ta bonne humeur. Je vous souhaite que du bonheur pour la suite et un épanouissement professionnel et personnel.

Aux autres membres de l'IVPC, qui n'ont pas moins compté pour moi. Merci à **Marie**, **Lucie et Maryline** pour leur gentillesse, leur bonne humeur et leur soutien en tant que doctorantes. Merci de m'avoir accueillie dans le labo, pour les soirées etc. Merci pour les pauses glaces et autres moments détente bien mérités. Je ne vous souhaite que du bonheur.

A **Carine** pour ta bienveillance à mon égard et nos discussions que nous avons pu avoir. A **Catherine**, **Céline** et ma voisine de bureau **Kathy** pour votre gentillesse et les moments de détente plus fun.

A **Sylvie** pour ton efficacité et ton aide pour tout le côté administratif. Sans parler de tes encouragements et nos discussions qui font du bien à 7 heures du matin quand on est seule. A **Angélique**, pour ta présence à mes côtés en salle droso, ta curiosité et toutes nos discussions agréables et intéressantes. Tu es ma gardienne de drosophile, merci d'en avoir pris soin pendant ma rédaction. Tu sais que je t'apprécie énormément et j'espère qu'on restera en contact.

A **Barbara G**, **Alexandra**, **Caroline** et **Christine** pour leur disponibilité et nos discussions. A **Jean-Michel** pour nos discussions qui étaient très agréables et intéressantes, et de m'avoir partagée ta passion professionnelle.

Aux anciens doctorants : **Margot**, d'avoir pris le temps pour m'expliquer certaines techniques, je te souhaite que du positif pour la suite, que ce soit professionnel ou personnel ; **Wilhelm**, je te souhaite un épanouissement dans ton post-doc à Glasgow ainsi que dans ta vie personnelle parce que même si nous n'étions pas d'accord sur tout, tu le mérites.

Je tiens également à remercier très sincèrement les membres de **l'équipe « Elements Transposable, Evolution, Population » du LBBE** (Laboratoire de biométrie et biologie évolutive). Tout d'abord, je sais que la situation était assez étrange puisque je n'ai pas souvent été avec vous. Pour autant, vous m'avez tout de suite intégrée et tous mes passages courts ou longs au LBBE étaient un véritable plaisir.

A **Cristina** pour ton intérêt et ton enthousiasme concernant mon projet de thèse. Je suis très contente d'avoir travailler dans ton équipe et te remercie pour cela. Merci pour ton aide, ton expertise scientifique et ta bienveillance.

A **Nelly** de m'avoir aidé tout au long de ma thèse. Merci pour ta disponibilité, ta gentillesse et ton efficacité. A **Judit** de m'avoir transmis le pouvoir d'isoler les petits ARN ! Merci pour ton temps et tous tes précieux conseils.

A **Angelo** de m'avoir accueillie dans l'équipe et de m'avoir fait découvrir le Domus, pour partager un bon repas et une pause très agréable. Merci aussi pour nos discussions et ton aide. J'espère que tu va continuer à t'épanouir dans ton nouveau travail et je te souhaite le meilleur pour la suite car tu le mérites.

Mille mercis à **Justine** et **Vincent** qui m'ont épaulée pour mes analyses bio-infomatiques, mes débuts dans ce monde que je ne connaissais pas. Merci pour votre temps, votre patience et pour votre sympathie.

A **Pierre**, pour ta bonne humeur et ton dynamisme. Merci aussi pour ton aide au labo.

A **Cecilia** pour nos discussions et ta solidarité. Je suis très heureuse de t'avoir rencontrée.

Je vous souhaite à tous le meilleur pour la suite. Je suis très impressionnée par vos recherches et je suis sûre que vous allez cartonner pour la suite. Bon courage en tout cas !

A **Matthieu**, **Emmanuelle**, **Anabelle** pour vos conseils, notamment en réunion d'équipe, et pour votre bienveillance.

A **Bianca** avec qui j'étais toujours très contente de discuter et de partager des idées sur nos travaux de recherche.

A **Rita** ! J'adore ton dynamisme et ton humanité. Tu es une personne extraordinaire, toujours prête à donner de soi. Merci pour ton aide au labo et aussi pour ton soutien lorsque j'étais enfermée dans la grotte du confocale.

J'aurais aimé avoir plus de temps pour ma thèse et être par conséquent plus au LBBE car l'ambiance de travail était vraiment agréable grâce à vous tous.

Merci à l'équipe de **Vincent Lotteau** et **Patrice André** où j'ai réalisé mes stages de Master, qui m'ont donnée l'envie de continuer dans la recherche. Sachez que ma déception était immense lorsqu'aucune opportunité n'est arrivée à temps pour ma thèse. Merci **Olivier** pour tout ce que tu m'as appris et que j'ai essayé au mieux d'appliquer pendant ma thèse. Merci à **Marine**, **Anne**, **Laure**, **Caroline**, **Christophe et Marie** de m'avoir permis de débuter dans la recherche dans une agréable ambiance. Merci à **Clémence**, ma Clem qui est maintenant une amie en or, merci pour tout ce que tu as fait pour moi au labo et bien sûr en dehors ! Et comme tu m'as dit en partant du labo, on ne se quitte plus de toute façon. Merci **Karim** pour ton aide au labo, pour les minutes mouz', les cours d'anglais et ta joie de vivre.

Bien évidemment je remercie du fond du cœur toute ma **famille** et ma **belle-famille** qui sont toujours là pour me soutenir dans mes projets et pour me changer les idées en organisant plein de choses formidables. Merci d'être là, je vous aime !

Merci à mon petit chat, **Spartacus** qui a été à mes côtés pendant toute mes études, ma petite glue tellement importante pour moi. Et je n'oublie pas son demi-frère **Corto** qui met un peu plus d'ambiance à la maison ☺.

Je remercie également sincèrement ma deuxième famille, mes amis, mon clan ! ☺ Dix millions de mercis à **Chou**, **Claire**, **Clem**, **Débo**, **Drap-housse**, **Drinou**, **Flo**, **Guitoune**, **Jess**, **Kimo**, **Mimie**, **Van Van**, **Yum** ! Merci pour tout ce qu'on a partagé, ce qu'on partage et ce qu'on partagera. Merci pour les moments de fête mais aussi pour les moments de réconfort.

Votre présence compte beaucoup pour moi et je suis heureuse de vous avoir. Votre aide et vos encouragements ont été précieux durant toutes ces années. Chou et Claire, Je suis persuadée que vous allez finir en beauté votre thèse. Encore félicitations Karim pour ta thèse, je te souhaite de t'épanouir durant ton post-doct aux USA même si tu vas nous manquer. Je pourrais bien évidemment écrire une dizaine de pages pour vous remercier chacun mais vous le savez, comme je l'ai souvent dit le soutien de mes amis n'a pas de prix à mes yeux !

Un grand merci à **Kéké**. On se suit maintenant depuis plus de 6 ans. Tout a commencé sur les bancs de la fac où je n'étais pas vraiment sûre de moi. Tu m'as beaucoup apporté depuis ce temps, et notamment avec notre super expérience en Angleterre que je n'aurais jamais faite sinon, merci ! Tu m'as vraiment impressionnée lors de ta soutenance de thèse, encore bravo et j'espère que tu vas t'épanouir dans ta poursuite professionnelle. Merci aussi pour nos discussions très agréables, nos soirées, et pour tes concerts formidables. Je vous souhaite que du bonheur avec **Cindy** et je suis certaine qu'on a encore plein de choses à vivre ensemble !

Renaud, merci pour ton soutien tout au long de ma thèse et bien avant. Merci pour nos discussions très intéressantes qui m'ont permis d'en apprendre plus sur moi-même et de gagner en confiance. Je te souhaite un épanouissement total dans ton travail qui te va à merveille et te souhaite également tout le bonheur du monde pour ta vie personnelle.

Le meilleur pour la fin, mon **Félix**, avec qui je partage ma vie depuis un certain nombre d'années ! ☺ Je ne sais même pas comment exprimer ce que je ressens aujourd'hui en voyant notre histoire. Je suis tellement heureuse de vivre tout ça avec toi, tous ces projets et ces belles choses que nous partageons. Tu m'as toujours soutenue dans tous mes projets, et bien entendu tout au long de ma thèse ! Merci de m'avoir supportée, aidée, encouragée et bien plus. Merci pour ton optimisme et ton enthousiasme qui me fascine encore maintenant ! ^^ J'ai conscience de la chance que j'ai de t'avoir. C'est grâce à toi que j'en suis là aujourd'hui et je ne t'en serais jamais assez reconnaissante ! Ne change rien surtout <3.

C'est par le travail que la femme a en grande partie franchi la distance qui la séparait du mâle ; c'est le travail qui peut seul lui garantir une liberté concrète.

Simone de Beauvoir

RESUME

Les éléments transposables (ET) sont des parasites génomiques présents dans tous les génomes, dont une partie présente une structure similaire à celle de certains virus. Dans les cellules ovariennes d'insecte, l'abondance des transcrits d'ET est contrôlée par ARN interférence (ARNi), et plus particulièrement par la voie piARN (Piwi-interacting RNA). Une autre voie d'ARNi, la voie siARN (Small interfering RNA), constitue l'une des principales réponses immunitaires des insectes contre les infections virales, et est aussi dédiée au contrôle somatique des ET. Ces deux voies d'ARNi sont dirigées par des effecteurs moléculaires distincts et décrites comme indépendantes. Cependant, des similitudes structurales et de mécanisme de contrôle entre les ET et les virus suggèrent la possibilité d'une interaction. Nous avons utilisé la drosophile comme modèle et infecté l'organisme avec le virus Sindbis (SINV), un arbovirus (Arthropod-Borne virus), ou le virus Sigma (SV), un virus spécifique de drosophile. À l'aide d'un séquençage à haut débit, nous avons caractérisé les répertoires d'ARN et de petits ARN interférents produits par les drosophiles infectées, à partir des tissus de carcasses ou d'ovaires. Globalement, nos résultats démontrent un impact de l'infection virale sur les quantités de transcrits d'ET via la modulation des répertoires piARN et siARN, et représentent la première démonstration des effets d'infections virales sur l'activité des ET. Plus précisément, l'infection par SINV favorise une diminution globale des quantités de transcrits d'ET alors que l'infection par SV réactive de nombreux ET. Nos données suggèrent que la modulation résulte de substrats d'ARN partagés et de trans-régulateurs communs des voies de l'ARNi. Ces résultats ouvrent un nouvel axe de recherche en génomique, suggérant que les épidémies virales ou les infections chroniques peuvent avoir un impact sur l'activité des ET, et donc sur le taux de diversification génétique ultérieure.

Titre : Caractérisation des interactions entre les défenses antivirales et le contrôle génomique des éléments transposables chez *Drosophila*

Mots-clés : virus à ARN, élément transposable, ARN interférence, piARN, siARN, *Drosophila*

ABSTRACT

Transposable elements (TEs) are genomic parasites that are found in all genomes, a part of which display structure similarity to some viruses. In insect ovarian cells, TE transcript abundance is controlled by RNA interference (RNAi) machinery and more particularly by the piRNA pathway (Piwi-interacting RNA). Another RNAi pathway, named siRNA pathway (Small interfering RNA), is one of the key insect immune responses against viral infections and is also dedicated to TE somatic control. These two interfering RNA pathways are led by distinct molecular effectors and described as independent. However, similarities in structure and control mechanisms across TEs and viruses suggest that an interplay may exist. We used *Drosophila* as a model, and infected the organism with Sindbis virus (SINV), an arbovirus (Arthropod-Borne virus), or Sigma virus (SV), a *Drosophila*-specific virus. Using high-throughput sequencing, we characterized the RNA and small-RNA repertoires produced by *Drosophila* flies from carcass or ovary tissues. Overall, our results demonstrate an impact of viral infection on TE transcript amounts via modulation of the piRNA and siRNA repertoires and represent the first demonstration of the effects of viral infection on TE activity. More precisely, SINV infection promotes a global decrease of TE transcript levels while SV infection reactivates many TEs. Our data suggest that the modulation results from shared RNA substrates and common trans-regulators of the small RNA pathways. These results open new avenue of research in genomics, suggesting that viral epidemics or chronic infections can impact TE activity and thus the rate of subsequent genetic diversification.

Title: Characterization of the interplay between RNA interference-type antiviral defenses and genomic control of transposable elements in *Drosophila*

Keywords: RNA viruses, transposable element, RNA interference, piRNA, siRNA, *Drosophila*

TABLE DES MATIERES

REMERCIEMENTS 5
RESUME 10
ABSTRACT 11
TABLE DES MATIERES 12
ABREVIATIONS ET ACRONYMES 16
LISTE DES FIGURES 19
LISTE DES FIGURES SUPPLEMENTAIRES 22
LISTE DES TABLEAUX 24
INTRODUCTION GENERALE ET OBJECTIFS 26
SYNTHESE BIBLIOGRAPHIQUE 31
I. La drosophile 31
1. Le modèle 31
2. *Drosophila simulans versus melanogaster* 32
II. Les virus à ARN 34
1. Les virus infectant naturellement la drosophile 35
A. Le virus sigma 36
B. Le virus C de drosophile (DCV) 39
C. Le virus X de drosophile (DXV) 40
D. Le virus A de drosophile (DAV) 40
E. Le virus Nora (NV) 40
2. Les virus n'infectant pas naturellement la drosophile 41
3. Arbovirus 42
A. Le virus Sindbis 43
B. Interaction entre le virus et le vecteur insecte 49
III. Lignes de défense antivirale chez les insectes 50
1. La voie Jak/STAT 50
2. Les voies immunitaires Imd et Toll 52
3. Les voies Jak/STAT, Imd et Toll lors d'infections virales 53
IV. Les éléments transposables 55
1. Introduction générale 55

2. Rétrotransposons, classe I 56
A. Rétrotransposons à LTR 56
B. Rétrotransposons sans LTR 60
3. Transposons à ADN, classe II 62
4. Paysage des ET selon l'hôte 63
5. Impacts de la transposition des ET 65
6. Impact d'un stress sur la transposition des ET 66
V. Le mécanisme d'ARN interférence 68
1. siARN 69
A. siARN exogènes 70
B. Réponse ARN interférence systémique chez *Drosophila* 72
C. endo-siARN 75
2. miARN 77
3. piARN 79
A. Introduction générale 79
B. piARN germinaux et somatiques dans les structures ovariennes 79
C. piARN somatiques, non gonadiques 86
VI. Interférences entre les piARN et les siARN 87
1. Le rôle des ET dans l'immunité anti-virale 88
A. Stimulation de l'immunité antivirale par la voie des siARN 88
B. Induction de l'immunité antivirale par la voie des piARN 89
2. Production de piARN viraux et activité antivirale 90
A. Production de piARN viraux 90
B. piARN viraux chez le moustique 92
MATERIEL ET METHODES 96
I. Souches de drosophiles 96
II. Validation du génotype chez *D. melanogaster Dcr-2*$^{[L811fsX]}$ 96
III. Validation du génotype chez *D. melanogaster CG4572*^{c05963} 97
IV. Détection et élimination de *Wolbachia* 97
V. Détection et élimination des virus des drosophiles 98
VI. Production du virus Sindbis par génétique inverse 100
1. Préparation du plasmide pour la transcription *in vitro* 100
2. Transcription *in vitro* 100
3. Electroporation des cellules BHK-21 100

VII. Titration des stocks viraux de SINV 102
VIII. Production du virus sigma (DMelSV) 102
IX. Titration du stock viral de DMelSV 103
X. Infection des drosophiles 103
XI. Courbes de survie et titration virale de SINV 104
XII. Titration virale de DMelSV 105
XIII. Dissection des drosophiles 105
XIV. Séquençage du transcriptome (RNA-seq) 105
XV. Analyse de données de séquençage publiques 107
XVI. Séquençage des petits ARN interférents (Small RNA-seq) 108
XVII. Titrations virales de SINV à partir des différents tissus de drosophiles 111
XVIII. Détection de DAV par RT-PCR à partir des drosophiles choniquement infectées par DMelSV Hap23 112
XIX. Transcription inverse et PCR quantitative (RT-qPCR) 112
XX. Immunomarquage d'ovaires et de coupes de drosophiles infectées par le virus Sindbis et microscopie confocale 114
1. Préparation des blocs d'OCT contenant les tissus à couper 114
2. Préparation des coupes de tissus 114
3. Immunomarquage : détection de la capside de SINV 114
4. Observation des coupes au microscope confocal 115
RESULTATS EXPERIMENTAUX (I) 117
I. Les voies d'ARN interférence siARN et piARN interagissent et modulent la transcription des éléments transposables lors de l'infection virale chez *Drosophila* 117
1. Introduction 117
2. Infection par le virus Sindbis chez *D. melanogaster* 118
A. Identification et élimination de *Wolbachia* et des infections virales chroniques chez *D. melanogaster* 118
B. Etude de la réplication du virus Sindbis chez *D. melanogaster* w^{1118} 120
C. Modulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis chez *D. melanogaster* w^{1118} 122
3. Infection par le virus Sindbis chez *D. simulans* 127
A. Identification des infection chroniques et élimination de *Wolbachia* chez *D. simulans* 128
B. Etude de la réplication du virus Sindbis chez *D. simulans*, Makindu 129
C. Modulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis chez *D. simulans*, Makindu 133

4. Mécanismes impliqués dans la dérégulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis 141

A. Impact de la voie d'absorption d'ARN double brin sur la dérégulation des ET et des petits ARN interférents ciblant les ET 141

B. Impact de la voie des siARN sur la dérégulation des ET et des petits ARN interférents ciblant les ET 150

C. Analyse des petits ARN interférents ciblant les ET chez *D. melanogaster* 163

D. Résumé 166

Discussion (I) 170

DONNEES SUPPLEMENTAIRES (I) 179

1. Infection par le virus Sindbis de *D. melanogaster* 180

2. Infection par le virus Sindbis chez *D. simulans* 183

3. Mécanismes impliqués dans la dérégulation des ET 186

..... 199

RESULTATS EXPERIMENTAUX (II) 201

II. Réactivation d'éléments transposables lors de l'infection par les virus Sigma et DCV chez *Drosophila simulans* 202

1. Introduction 202

2. Etude de la réplication du virus Sigma (DMelSV) chez *D. simulans* 203

3. Production des petits ARN interférents viraux chez *D. simulans* 205

4. Dérégulation des voies des petits ARN interférents et de l'expression des éléments transposables lors de l'infection par DMelSV chez *D. simulans* 207

5. Dérégulation des gènes impliqués dans la voie Toll lors de l'infection par DMelSV chez *D. simulans* 214

6. Dérégulation des quantités de transcrits des éléments transposables lors de l'infection par DCV chez *D. melanogaster* 218

A. Réplication de DCV et tolérance virale 218

B. Modulation des quantités de transcrits des ET 219

Discussion (II) 222

DONNEES SUPPLEMENTAIRES (II) 226

DISCUSSION GENERALE 233

BIBLIOGRAPHIE 239

ANNEXES 266

ABREVIATIONS ET ACRONYMES

Ac	Anticorps
ADN	Acide Désoxyribonucléique
ADNc	ADN complémentaire
ADNv	ADN viral
Ago1	Argonaute 1
Ago2	Argonaute 2
Ago3	Argonaute 3
AMP	Antimicrobial peptides
ARN	Acide RiboNucléique
ARNi	ARN interférence
ARNm	ARN messager
Aub	Aubergine
BTV	Bluetongue virus
CHIKV	Chikungunya virus
CO_2	Dioxyde de carbone
CrPV	Cricket paralysis virus
DAV	*Drosophila* A virus
Dcr-1	Dicer-1
Dcr-2	Dicer-2
DCV	*Drosophila* C virus
DENV	Dengue virus
DID50	50% *Drosophila* infectious dose
DMelSV	*Drosophila melanogaster* Sigma Virus
DMEM	Dulbecco's Modified Eagle Medium
DXV	*Drosophil*a X virus
endo-siRNA	Endogenous small interfering RNA

Env	Enveloppe
ERV	Endogenous Retroviruses
ET	Eléments transposables
EVE	Endogenous Viral Elements
exo-siRNA	Exogenous small interfering RNA
FHV	Flock House virus
ICTV	International Committee on Taxonomy of Viruses
IIV-6	Invertebrate Iridescent virus 6
Imd	Immune deficiency
Jak/STAT	Janus kinase/signal transducer and activator of transcription
LACV	La Crosse virus
LINEs	Long interspersed nuclear elements
LTR	Long terminal repeats
miARN	micro-ARN
NIRVS	Non-Retroviral Integrated RNA Virus Sequences
NRAMP2	Natural Resistance-Associated Macrophage Protein 2
nsP	Non-structural proteins
NV	Nora virus
ONNV	O'nyong nyong virus
ORF	Open Reading Frame
OROV	Oropouche virus
PAMP	Pathogen Associated Molecular Patterns
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PFU	Plaque-forming unit
PGRP	Peptidoglycan recognition protein
piARN	piwi-interacting RNA
pre-miRNA	MicroRNA precursor-miRNA

pri-miRNA	MicroRNA primary-miRNA
PRR	Pattern recognition receptor
RdRp	ARN-polymérase ARN-dépendante
RISC	RNA interfence silencing complex
RT-(q)PCR	Reverse-transcriptase (quantitative) polymerase chain reaction
RVFV	Rift Valley fever virus
SBV	Schmallenberg virus
SFV	Semliki Forest virus
SGP	Subgenomic promoter
SINEs	Short Interspersed Nuclear Elements
SINV	Sindbis virus
siRNA	small interfering RNA
SV	Sigma virus
SVF	Sérum de Veau Fœtal
TCID50	50% Tissue culture infectious dose
TIR	Terminal inverted repeats
TOSV	Toscana virus
TSD	Target Site Duplications
UTR	untranslated region
VEE	Venezuelan Equine Encephalitis virus
VLP	Virus Like-Particle
VSV	Vesicular Stomatitis Virus
WNV	West Nile Virus
ZIKV	Zika virus
Zuc	Zucchini

LISTE DES FIGURES

Figure 1. D. melanogaster et D. simulans présentent des différences de distribution et de prévalence des virus qui les infectent naturellement 33
Figure 2. Une grande diversité de virus à ARN infecte les insectes 35
Figure 3. Schéma du virion et du génome du virus sigma 37
Figure 4. Schéma du cylce de réplication du virus sigma 38
Figure 5. Transmission virale de SINV 44
Figure 6. Schéma de la structure du génome de SINV 45
Figure 7. Cycle de réplication de SINV 47
Figure 8. Réplication du génome de SINV 48
Figure 9. La voie Jak/STAT chez les insectes 51
Figure 10. Voies immunitaires relatives à NF-κB chez les insectes 52
Figure 11. Schéma de la transposition d'un rétrotransposon impliquant un mécanisme de type « copier-coller » 56
Figure 12. Schéma de la structure des différentes classes et sous-classes des ET 58
Figure 13. Schéma de la transposition d'un rétrotransposon à LTR 59
Figure 14. Schéma de la transposition d'un rétrotransposon sans LTR 61
Figure 15. Schéma de la transposition d'un transposon à ADN impliquant un mécanisme de type « couper-coller 63
Figure 16. Proportions des ET dans différents génomes eucaryotes séquencés 64
Figure 17. Mécanisme d'acquisition de la résistance au DMelSV 67
Figure 18. La voie de l'immunité antivirale des siARN chez les insectes 71
Figure 19. Diffusion systémique de siARN secondaires par les hémocytes chez la drosophile 73
Figure 20. siARN et diffusion systémique chez les insectes 74
Figure 21. Production d'endo-siARN chez la drosophile 76
Figure 22. La voie des miARN chez les insectes 78
Figure 23. Schéma de l'ovogénèse et structure du germarium 81
Figure 24. Voie primaire de biogenèse des piARN dans les cellules somatiques ovariennes chez la drosophile 83
Figure 25. Boucle d'amplification des piARN en ping-pong 85
Figure 26. Signature des piARN impliqués dans la boucle d'amplification en ping-pong 86
Figure 27. Modèle de la persistance virale chez les insectes 89
Figure 28. Schéma de l'origine de la production de piARN viraux chez le moustique 93
Figure 29. Production de SINV par électroporation des cellules BHK-21 101
Figure 30. Analyse des données de séquençage du transcriptome (RNA-seq) 107
Figure 31. Isolement des petits ARN interférents 109
Figure 32. Analyse des données de séquençage des petits ARN interférents (Small RNA-seq) 111
Figure 33. Absence de détection de transcrits de virus spécifiques de D. melanogaster (w^{1118} et Dcr-$2^{[L811fsX]}$) après javellisation des œufs de drosophile 119
Figure 34. Cinétique de réplication de SINV chez D. melanogaster 121
Figure 35. Absence de détection de SINV dans les ovaires de D. melanogaster w^{1118} 122
Figure 36. Analyse des siARN et des piARN ciblant SINV chez D. melanogaster w^{1118} 124
Figure 37. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez D. melanogaster (w^{1118}) lors de l'infection par SINV 126
Figure 38. Signature ping-pong des piARN ciblant les ET dans les carcasses de D. melanogaster w^{1118} 127
Figure 39. Détection des transcrits viraux chez (D. simulans) Makindu 128

Figure 40. Réplication de SINV chez (D. simulans) Makindu à 6 jours après infection 130
Figure 41. Cinétique de réplication de SINV chez (D. simulans) Makindu 130
Figure 42. Absence de détection de SINV dans les ovaires de D. simulans Makindu 131
Figure 43. SINV est principalement détecté au niveau du thorax et des têtes chez Makindu 133
Figure 44. Analyse des siARN et des piARN ciblant SINV chez D. simulans Makindu 134
Figure 45. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez Makindu lors de l'infection par SINV 137
Figure 46. Diminution de l'expression des ET chez Makindu lors de l'infection par SINV 138
Figure 47. Signature ping-pong des piARN ciblant les ET dans les carcasses de D.simulans Makindu 138
Figure 48. Analyse des siARN et des piARN ciblant NV chez Makindu lors de l'infection expérimentale par SINV 140
Figure 49. Absence de différence de survie entre les mutants CG4572^{c05963} et les drosophiles contrôles w^{1118} 143
Figure 50. Absence de différence de réplication entre les mutants CG4572^{c05963} par rapport aux contrôles w^{1118} 144
Figure 51. Réplication de SINV chez D. melanogaster w^{1118}, CG4572^{c05963} et Dcr2$^{[L811fsX]}$ 145
Figure 52. Analyse des siARN et des piARN ciblant SINV chez D. melanogaster w^{1118} et CG4572^{c05963} 146
Figure 53. Analyse de la modulation des productions de petits ARN interférents ciblant les ET chez D. melanogaster CG4572^{c05963} (CG4572-/-), lors de l'infection par SINV 149
Figure 54. Mortalité des mutants D. melanogaster Dcr-2$^{[L811fsX]}$ induite par SINV 151
Figure 55. Réplication de SINV dans les mutants Dcr2$^{[L811fsX]}$ et les drosophiles w^{1118} 152
Figure 56. Détection de SINV seulement en périphérie des ovaires des mutants Dcr-2$^{[L811fsX]}$ 153
Figure 57. Détection de SINV dans le thorax des drosophiles w^{1118} et Dcr-2$^{[L811fsX]}$ 154
Figure 58. Détection de SINV dans l'abdomen des drosophiles w^{1118} et Dcr-2$^{[L811fsX]}$ 155
Figure 59. Détection de SINV dans le corps gras des drosophiles Dcr-2$^{[L811fsX]}$ 156
Figure 60. Absence de production des brins antisens des petits ARN interférents ciblant SINV chez les mutants Dcr-2$^{[L811fsX]}$ 158
Figure 61. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez D. melanogaster Dcr-2$^{[L811fsX]}$, lors de l'infection par SINV 162
Figure 62. Analyse des siARN et des piARN ciblant les ET chez D. melanogaster 165
Figure 63. Résumé de la modulation de l'activité des ET et des petits ARN interférents ciblant les ET dans les carcasses de D. melanogaster 167
Figure 64. Résumé de la modulation de l'activité des ET et des petits ARN interférents ciblant les ET dans les ovaires de D. melanogaster 169
Figure 65. Modèle de l'impact de l'infection par SINV sur les ET 177
Figure 66. Réplication de DMelSV chez Makindu 204
Figure 67. Détection des transcrits viraux chez Makindu 205
Figure 68. Analyse des siARN et des piARN ciblant DMelSV chez D. simulans Makindu 206
Figure 69. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez Makindu lors de l'infection par DMelSV 210
Figure 70. Les ERV sont significativement plus modulés que les autres classes d'ET chez Makindu, lors de l'infection par DMelSV, dans les carcasses 211
Figure 71. Comparaison des modulations de l'activité des ET et des petits ARN interférents ciblant les ET entre les infections par DMelSV et SINV 214
Figure 72. Parmi les gènes différentiellement exprimés lors de l'infection par DMelSV, on trouve un enrichissement en gènes impliqués dans les processus de régulation de la voie Toll 216
Figure 73. Surexpression de gènes impliqués dans la voie Toll et de certains AMP chez Makindu, lors de l'infection par DMelSV, dans les carcasses 217
Figure 74. Modulation de l'abondance des transcrits des ET chez D. melanogaster, lors de l'infection par DCV dans les drosophiles entières (A ; C) et dans le corps gras (B ; D) 220

Figure 75. Comparaison des quantités de transcrits d'ET entre les w^{1118} et les mutants G9a _____ *221*
*Figure 76. Modèle de l'impact de l'infection par DMelSV et DCV sur les ET*_____ *223*

LISTE DES FIGURES SUPPLEMENTAIRES

Figure S1. Absence de détection des virus naturels chez D. melanogaster après javellisation des œufs de drosophiles 180
Figure S2. Absence de détection de Wolbachia après traitement à la tétracycline chez D. melanogaster 181
Figure S3. Absence de signature ping-pong des piARN ciblant SINV chez D. melanogaster w^{1118} 182
Figure S4. La diminution de l'abondance des transcrits d'ET est dépendante de l'augmentation des productions de piARN et siARN ciblant les ET, lors de l'infection par SINV, dans les carcasses w^{1118} 182
Figure S5. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN et siARN chez D. melanoagster w^{1118}, lors de l'infection par SINV 183
Figure S6. Détection de NV chez (D. simulans) Makindu 184
Figure S7. Absence de détection de Wolbachia après traitement à la tétracycline chez D. simulans 185
Figure S8. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN (A) et siARN (B) chez Makindu, lors de l'infection par SINV 185
Figure S9. Parmi les gènes différentiellement exprimés lors de l'infection par SINV, on trouve un enrichissement en gènes associés au métabolisme de l'ARN 186
Figure S10. Validation du génotype $CG4572^{c05963\ -/-}$ par la sélection du phénotype associé 187
Figure S11. Absence de détection de SINV dans les ovaires des mutants D. melanogaster $CG4572^{c05963}$ 187
Figure S12. Production de siARN ciblant SINV et DAV chez D. melanogaster $CG4572^{c05963}$ 188
Figure S13. Comparaison des quantités de petits ARN interférents ciblant les ET entre les w^{1118} et les mutants $CG4572^{c05963}$, dans les carcasses 189
Figure S14. Comparaison des quantités de petits ARN interférents ciblant les ET entre les w^{1118} et les mutants $CG4572^{c05963}$, dans les ovaires 190
Figure S15. Génotype des mutants $Dcr\text{-}2^{[L811fsX]}$ 191
Figure S16. Absence de production de petits ARN interférents ciblant les virus naturels persistants chez D. melanogaster $Dcr\text{-}2^{[L811fsX]}$ 192
Figure S17. Comparaison des quantités de transcrits des ET et des piARN ciblant les ET entre les w^{1118} et les mutants $Dcr\text{-}2^{[L811fsX]}$, dans les carcasses 193
Figure S18. Comparaison des quantités de transcrits des ET et des piARN ciblant les ET entre les w^{1118} et les mutants $Dcr\text{-}2^{[L811fsX]}$, dans les ovaires 194
Figure S19. Pourcentage des petits ARN interférents ciblant les ET chez D. melanogaster 195
Figure S20. Signature ping-pong des piARN ciblant les ET chez D. melanogaster 196
Figure S21. Réplication de SINV dans les têtes de drosophiles chez Makindu 197
Figure S22. Absence de modulation significative de l'expression des ET lors de l'infection par SINV dans les têtes et le reste des carcasses des Makindu 197
Figure S23. Réplication de SINV chez Makindu après javellisation des œufs 198
Figure S24. Absence de détection de SINV dans les ovaires chez Makindu après javellisation des œufs 199
Figure S25. Absence de modulation significative de l'expression des ET chez Makindu, après élimination des infections persistantes, lors de l'infection par SINV 199
Figure S26. Corrélation positive entre les petits ARN interférents de 23-30 nucléotides et ceux de 21 nucléotides dans toutes les conditions 200
Figure S27. Titration in vivo de DMelSV chez D. melanogaster 227
Figure S28. Analyse des siARN et des piARN viraux chez D. simulans Makindu 228

Figure S29. Absence de détection des transcrits de DAV dans les drosophiles chroniquement infectées par DMelSV _____ *229*
Figure S30. L'augmentation de l'abondance des transcrits d'ET n'est pas significativement dépendante de la diminution des productions de piARN et siARN ciblant les ET, lors de l'infection par DMelSV, dans les carcasses _____ *230*
Figure S31. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN (A) et siARN (B) chez Makindu, lors de l'infection par DMelSV _____ *230*
Figure S32. Le régulateur épigénétique G9a est nécessaire à la tolérance des drosophiles lors de l'infection par DCV (Merkling et al., 2015) _____ *231*

LISTE DES TABLEAUX

Tableau 1. Virus étudiés expérimentalement chez Drosophila _____ *42*
Tableau 2. Propriétés des petits ARN interférents _____ *69*
*Tableau 3. Amorces utilisées pour la détection des infections par les virus persistants et par la bactérie endosymbiotique Wolbachia dans les lignées de drosophiles*_____ *99*
Tableau 4. Préparation de milieu de broyage des drosophiles _____ *105*
Tableau 5. Amorces utilisées pour la détection des transcrits des ET et des virus SINV et DMelSV _____ *113*

INTRODUCTION GENERALE ET OBJECTIFS

INTRODUCTION GENERALE ET OBJECTIFS

Ces dernières années, nous avons assisté à l'émergence de maladies infectieuses responsables d'importantes problématiques en matière de santé humaine et animale. Certaines d'entre elles sont causées par des arbovirus (ARthropod-BOrne VIRUS), incluant les membres de la famille des *Togaviridae* et du genre Alphavirus, tels que le virus Sindbis (SINV), le virus de la forêt de Semliki (SFV) ou encore le virus Chikungunya (CHIKV). Ce que l'on entend par arbovirus correspond à des virus ayant pour vecteurs les arthropodes hématophages, tels que les moustiques, les phlébotomes et les tiques, et ainsi n'infectant pas naturellement la drosophile. Nous nous intéresserons davantage au SINV qui est le principal modèle viral de mes travaux de thèse. Des maladies humaines causées par SINV ont été rapportées en Afrique du Sud et dans le Nord de l'Europe, incluant des symptômes tels que de la fièvre, des malaises, des éruptions cutanées et des douleurs musculo-squelettiques (Adouchief et al., 2016). Une problématique importante de cette infection est la persistance de ces symptômes pendant plusieurs années, sans traitement spécifique disponible. La transmission du virus SINV aux hôtes vertébrés implique des moustiques ornithophiles, principalement du genre *Culex* (Hesson et al., 2015). Lors de toute infection virale chez les insectes, la machinerie d'ARN interférence (ARNi), et principalement la voie des siARN (*small interfering RNA*) représente une réponse essentielle de l'immunité innée antivirale (Gammon and Mello, 2015).

Par ailleurs, la machinerie d'ARNi est également impliquée dans le contrôle des éléments transposables (ET). Ces derniers sont des parasites du génome, présents chez tous les organismes vivants et représentent un des constituants importants des génomes eucaryotes. Les ET sont des séquences d'ADN capables de transposer (c'est-à-dire de se déplacer) tout au long des chromosomes et de multiplier leur nombre de copies. L'intégration génomique de ces séquences peut conduire à des mutations délétères. La sélection naturelle a permis d'élaborer des mécanismes de contrôle, permettant la régulation de l'activité des ET. Au niveau des cellules ovariennes des insectes, l'abondance des transcrits d'ET est contrôlée notamment par la machinerie d'ARNi, mais d'une classe différente de celle impliquée dans le contrôle antiviral, la voie des piARN (*Piwi-interacting RNA*). Cependant, le contrôle des ET au niveau somatique peut être assuré par la voie des siARN chez *Drosophila* (Ghildiyal et al., 2008). Les voies des siARN et des piARN impliquent des effecteurs moléculaires distincts et sont décrites de façon indépendante. Par ailleurs, des similarités de structure et de mécanisme de contrôle suggèrent une interférence entre ces deux voies d'ARNi. De plus, certains ET, appelés rétrovirus endogènes (ERV), sont des vestiges d'infections rétrovirales, qui ont surmonté les barrières antivirales et atteint les cellules germinales. L'exemple d'ERV le plus caractéristique est *gypsy*, qui est transmis verticalemement, en suivant les lois de Mendel, et qui présente un pouvoir infectieux (Kim et al., 1994). Par ailleurs, les piARN sont également

effectifs dans les cellules somatiques des gonades de la drosophile, favorisant l'ovogénèse (C. Li et al., 2009; Malone et al., 2009). En dehors des gonades, des piARN somatiques fonctionnels ont été identifiés dans des types cellulaires spécifiques de *Drosophila melanogaster* tels que dans le corps gras (Jones et al., 2016) ou le cerveau (Perrat et al., 2013). Des piARN somatiques, ciblant de nombreux virus, sont aussi décrits chez les Arthropodes, tels que les moustiques (Lewis et al., 2018; Miesen et al., 2015; Morazzani et al., 2012), dont un grand nombre présentent une activité antivirale (Varjak et al., 2018b).

L'objectif principal de mes travaux de thèse consiste à explorer chez les insectes les interférences entre les réponses de l'hôte face aux infections virales et l'immunité génomique des ET, à travers les voies d'ARNi. Comprendre davantage les interactions hôte-pathogène est d'une grande importance dans l'étude de l'évolution des génomes des insectes, de la persistance virale ou de la compétence vectorielle. Ainsi, nous cherchons principalement à savoir si l'expression des ET est impactée lors d'infections virales chez *Drosophila*, et si les quantités de piARN et de siARN ciblant les ET sont également modulées lors de ces infections virales. Pour répondre à ces questions, nous avons profité des avantages du modèle drosophile, incluant notamment la connaissance des séquences complètes des génomes de drosophiles et la disponibilité de divers outils génétiques. SINV a été sélectionné pour mes travaux de thèse, notamment car il se réplique expérimentalement chez *Drosophila* et permet ainsi l'activation initiale de la machinerie d'ARNi. Cependant, SINV ne semble pas se répliquer dans les ovaires des drosophiles. Nous avons également étudié des virus naturels de *D. melanogaster*, tels que le virus sigma (SV) et le virus C de drosophile (DCV), afin d'élargir nos recherches et de déterminer si les interférences identifiées sont virus-spécifiques ou non. SV est également un modèle de choix puisqu'il se transmet verticalement chez la drosophile, de façon transovarienne.

Nous avons effectué des analyses de séquençage à haut débit, pour l'étude du transciptome (RNA-seq) et du répertoire des petits ARN interférents (Small RNA-seq) à partir d'échantillons d'ovaires de drosophiles et de carcasses (correspondant au corps de la drosophile sans les ovaires). L'analyse transcriptomique de 237 familles et sous-familles d'ET, nous a notamment permis d'identifier une réduction globale de l'abondance des transcrits d'ET, lors de l'infection par SINV, par rapport à la condition non infectée, dans les carcasses de *D. melanogaster* et *D. simulans*. Ces modulations sont associées à une augmentation globale de la production des piARN et des siARN ciblant ces ET. Ces résultats représentent la première démonstration de modulation des transcrits d'ET et des petits ARN interférents ciblant les ET lors de l'infection virale.

Dans une seconde partie, nous avons utilisé des mutants de *D. melanogaster*, permettant de décrypter les mécanismes sous-jacents mis en jeux lors de ces modulations.

Pour cela, nous avons utilisé des mutants de la voie de l'*uptake* ou import de l'ARN double brin (*CG4572*^{c05963}), impliquée dans la diffusion systémique des siARN, et également des mutants de la voie des siARN (*Dcr2*$^{[L811fsX]}$). Chez les mutants *CG4572*^{c05963}, la modulation de la production des petits ARN interférents lors de l'infection par SINV a quasiment disparu, démontrant l'importance de la voie de l'import d'ARN double brin dans la modulation des ET. Ces résultats suggèrent que lors de l'infection virale, la voie de l'import d'ARN double brin induirait la production systémique de siARN dirigés contre les ET, diminuant les quantités de transcrits d'ET dans les carcasses. Dans un contexte non infectieux, les mutants *Dcr2*$^{[L811fsX]}$, incapables de produire des siARN fonctionnels, voient leur taux de production de piARN augmenter par rapport à la condition sauvage (*w*^{1118}), en carcasses. Les substrats d'ARN des ET ne pouvant être pris en charge par les siARN, deviennent potentiellement disponibles pour la voie des piARN. Dans un contexte infectieux, des résultats opposés aux conditions non infectées ont été identifiés, incluant une augmentation du taux de transcrits d'ET, dépendant d'une diminution de la production des piARN ciblant ces ET et incluant une forte diminution de l'expression du gène *piwi*, chez les mutants *Dcr2*$^{[L811fsX]}$ par rapport aux drosophiles sauvages. Lors de l'infection, la diminution de la production de piARN observée entraîne une augmentation du taux des transcrits d'ET des mutants *Dcr2*$^{[L811fsX]}$, ce qui est l'opposé des modulations observées chez *D. melanogaster w*^{1118}. En considérant l'ensemble de nos résultats dans les carcasses, nous proposons un modèle impliquant des précurseurs d'ARN communs entre les voies d'ARNi (siARN et piARN). Ces résultats, d'une importance fondamentale pour la génomique de l'évolution, indiquent que les infections virales, telles que les infections chroniques ou les épidémies, pourraient impacter la fréquence de mutation d'un organisme à travers la modulation de l'activité des ET et ainsi impacter la vitesse de diversification génétique des organismes.

Au niveau des ovaires de *D. simulans* infectées par SINV, aucune particule infectieuse de SINV, et aucune modulation de l'abondance des transcrits d'ET n'ont été détectées. Cependant, une modulation importante des petits ARN interférents a été identifiée, incluant une diminution de la production des piARN et une augmentation de la production des siARN. Les résultats, dans cet organe, sont différents chez *D. melanogaster*, montrant une légère diminution du taux de transcrits d'ET et des petits ARN interférents. Ces résultats suggèrent une forte pression sélective sur le maintien de la stabilité du transcriptome des ET dans les ovaires, toutefois résultant en des profils distincts de modulation des petits ARN interférents lors de l'infection virale chez *D. simulans* et *D. melanogaster*. De plus, de futures investigations sont nécessaires pour explorer les impacts de la modulation des piARN ovariens transmis à la descendance chez *D. simulans*.

Dans une troisième partie, nous avons analysé l'impact de l'infection de *D. simulans* par DMelSV, sur l'abondance des transcrits d'ET et la production des petits ARN interférents, dans les ovaires et dans les carcasses de drosophiles. Dans les ovaires de *D. simulans* infectés par DMelSV, les transcrits de DMelSV ont été identifiés, suggérant une réplication virale au

niveau de cet organe. En parallèle, une légère diminution de la quantité de transcrits d'ET et une légère augmentation de la production de piARN ont été révélées. Ces résultats confirment la démonstration de modulation de l'activité des ET lors d'une infection virale. Nous avons également comparé ces résultats à ceux obtenus lors de l'analyse du transcriptome de *D. melanogaster* infectées par DCV, à partir de mouches entières ou du corps gras des drosophiles. En effet, le corps gras représente un organe cible de DCV (Deddouche et al., 2008; Merkling et al., 2015). Nous avons identifié une augmentation du taux de transcrits d'ET pour les carcasses infectées par DMelSV, et pour les drosophiles entières et le corps gras lors de l'infection par DCV. Ces résultats suggèrent une modulation des ET commune lors de l'infection par des virus de drosophile, mais diffère de celle identifiée lors de l'infection par un arbovirus, tel que SINV.

Cette approche assez novatrice contribue à l'amélioration des connaissances concernant les interactions hôte-virus chez les insectes. Les modulations de l'activité des ET lors de l'infection virale, impliquant la machinerie d'ARNi, peuvent avoir un impact sur la diversité génétique des insectes et l'évolution des génomes des insectes. L'ensemble de ces résultats présente une preuve de concept des interactions entre les siARN et les piARN lors d'une infection virale et peuvent ainsi permettre de nouvelles découvertes mettant en perspective les ET et les virus chez les insectes.

SYNTHESE BIBLIOGRAPHIQUE

SYNTHESE BIBLIOGRAPHIQUE

Mon projet de thèse se situe à l'interface entre les domaines de la virologie et de la génomique. Le modèle drosophile a été utilisé pour l'ensemble de mes travaux de thèse et sera détaillé dans la partie I (La drosophile), incluant la comparaison entre les deux espèces de drosophiles étudiées, *Drosophila melanogaster* et *Drosophila simulans*. Ces dernières présentent des différences en terme d'abondance de séquences d'éléments transposables (ET) au sein de leur génome et également au niveau de la prévalence des virus qui les infectent naturellement. Dans une deuxième partie, l'ensemble des virus à ARN sera présenté, intégrant les virus infectant naturellement la drosophile et ceux n'infectant pas naturellement la drosophile. Ces derniers incluent les arbovirus, se répliquant naturellement chez leur vecteurs, les arthropodes hématophages et pouvant ainsi transmettre ces virus à des hôtes vertébrés. Nous verrons ensuite dans une troisième partie les lignes de défense antivirale mises en place chez les insectes, en dehors de la machinerie d'ARN interférence (ARNi). Les ET, présentés au sein de la quatrième partie, incluent les rétrovirus endogènes, présentant des similarités de structure avec certains virus. Dans la cinquième partie, nous aborderons précisément la machinerie d'ARNi, impliquée dans la régulation des virus (siARN) et des ET (piARN). Enfin, dans une sixième partie, l'ensemble de cette bibliographie nous permettra de mettre en perspective les voies des siARN et des piARN, présentant des implications croisées dans les défenses génomique et antivirale chez les insectes.

I. La drosophile

1. Le modèle

La drosophile est depuis de nombreuses années un modèle de génétique de choix (Yamaguchi and Yoshida, 2018), incluant un temps court de génération, ainsi qu'un élevage simple et peu coûteux. Dans le cadre de cette étude, ce modèle est intéressant du fait de l'importante connaissance scientifique à son sujet, notamment la connaissance des séquences complètes des génomes de drosophiles et la disponibilité de divers outils génétiques. Son génome est constitué d'environ 20% d'éléments transposables (ET), incluant au moins 30% d'éléments non tronqués, supposés être actifs (Barrón et al., 2014; Kaminker et al., 2002; Quesneville et al., 2005). Le génome de *D. melanogaster* est un des génomes eucaryotes les plus étudiés et représente un modèle puissant pour l'études des ET eucaryotes. De plus, contrairement aux insectes hématophages, tels que les moustiques, les drosophiles sont facilement manipulables. Le modèle drosophile est très utilisé dans d'autres domaines que

celui de la génétique ou encore du développement, incluant les neurosciences et l'immunité. Les récepteurs de la voie Toll (TLRs, *Toll-like receptors*) indispensables à l'immunité innée ont notamment été découverts chez la drosophile (Lemaitre et al., 1996; Medzhitov et al., 1997). Le modèle drosophile peut être infecté naturellement ou expérimentalement par un large panel de virus. Il est devenu également un modèle très important pour l'étude des interactions hôte-pathogène incluant les virus et pouvant aider à améliorer la connaissance scientifique des mécanismes d'interaction entre les arbovirus et les vecteurs insectes, tels que les moustiques. Par ailleurs, certaines divergences ont été identifiées au niveau de la machinerie d'ARNi entre la drosophile et le moustique, particulièrement pour la voie des piARN, comprenant une expansion des protéines PIWI chez le moustique (Campbell et al., 2008). Il est important de noter que les voies d'ARNi des réponses immunitaires antivirales incluent la voie des siARN chez la drosophile et le moustique, et la voie des piARN uniquement chez le moustique (Léger et al., 2013; Miesen et al., 2015; Morazzani et al., 2012; Schnettler et al., 2013a; Vodovar et al., 2012).

2. *Drosophila simulans versus melanogaster*

On s'intéresse plus précisément aux espèces *D. melanogaster* et *D. simulans*, qui seront par la suite les modèles d'insectes utilisés dans mes travaux de recherche. Malgré de nombreuses similarités entre ces deux espèces, notamment une distribution géographique semblable, une taille de génome très proche et une distance phylogénétique très faible (environ 2 à 3 millions d'années), leur quantité d'ET est très variable. On sait depuis longtemps que le génome de *D. melanogaster* contient trois fois plus de séquences d'ET que chez *D. simulans* (Dowsett and Young, 1982). Il est intéressant de noter que cette teneur plus faible en ET chez *D. simulans*, par rapport à *D. melanogaster*, est notamment corrélée à des effets épigénétiques des ET plus importants, permettant de favoriser la répression épigénétique (Lee and Karpen, 2017). L'enrichissement des marques épigénétiques répressives H3K9me2 au niveau des séquences encadrant les ET est notamment plus élevé chez *D. simulans* par rapport à *D. melanogaster*, conduisant à une répression transcriptionnelle des ET par des modifications de la structure de la chromatine (Lee and Karpen, 2017). Il a également été montré que *D. melanogaster* contient des copies entières ainsi potentiellement actives, alors que *D. simulans* aurait plus de sites d'insertion d'ET mais présente plutôt des séquences anciennes et dégradées (Lerat et al., 2011). De plus, *D. simulans* présente une variabilité intraspécifique importante, notamment au niveau du rétrovirus endogène *tirant* (Akkouche et al., 2013, 2012; Fablet et al., 2019).

Ces deux espèces, dans la nature, partagent certains virus (Webster et al., 2015). Cependant, leur prévalence est plus importante chez *D. melanogaster* pour un certain nombre de virus, tels que le virus A de drosophile (DAV), le virus Kallithea ou encore le virus Motts Mill (**Figure 1.**). Ces mêmes auteurs ont montré que le virus sigma n'infecte pas naturellement *D. simulans*.

L'ensemble de ces résultats indique que ces deux espèces ont accumulé rapidement de nombreuses différences depuis la différenciation à partir de leur ancêtre commun (Capy and Gibert, 2004). Pour approfondir, l'analyse des différences entre *D. melanogaster* et *D. simulans* au niveau des gènes impliqués dans l'immunité antivirale pourrait être un axe de recherche très intéressant.

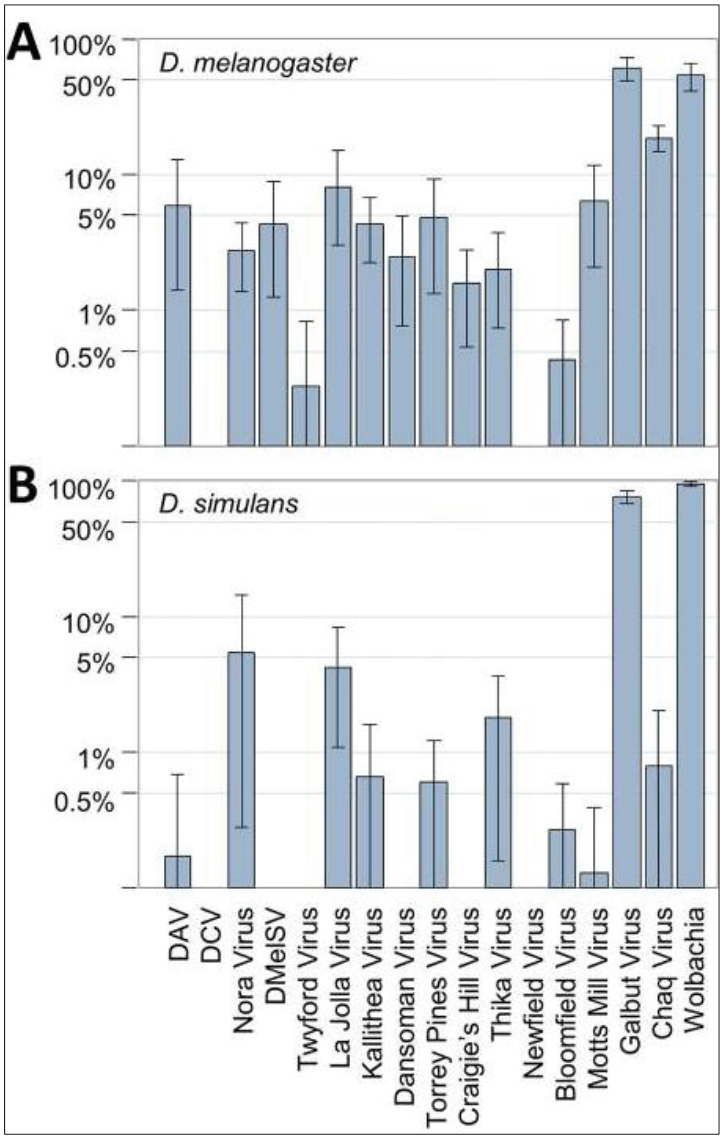


Figure 1. *D. melanogaster* et *D. simulans* présentent des différences de distribution et de prévalence des virus qui les infectent naturellement. Détection des virus par RT-PCR à partir de 1653 *D. melanogaster* et 658 *D. simulans* adultes récoltées dans 17 lieux dans le monde. Les graphes représentent la prévalence virale moyenne (de toutes le populations) chez *D. melanogaster* (A), et chez *D. simulans* (B), sur une échelle logarithmique (barres d'erreurs à 95% d'intervalle de confiance) (Webster et al., 2015).

II. Les virus à ARN

Les virus comportant un génome ARN sont très représentés chez les insectes et particulièrement chez la drosophile (**Figure 2., virus en violet**). Les interactions entre les virus spécifiques de *Drosophila* (décrits dans la partie II.1) et les réponses immunitaires sont particulièrement étudiées (Dostert et al., 2005; Ferreira et al., 2014; Lopez et al., 2018; Tsai et al., 2008). Par ailleurs, la drosophile peut être infectée expérimentalement par de nombreux virus (**Figure 2., virus en vert** ; décrits dans la partie II.2), permettant notamment de mettre en lumière de nouveaux mécanismes d'interaction hôte-pathogène (Goic et al., 2013; Lamiable et al., 2016b, 2016a; Tassetto et al., 2017). Ces virus incluent de nombreux arbovirus, détaillés plus précisément dans la partie II.3.

Comme tous les virus, ce sont des parasites intracellulaires obligatoires qui nécessitent une cellule hôte pour se répliquer. Le cycle de réplication virale engendre la production de nouvelles particules virales qui, une fois libérées, peuvent infecter de nouvelles cellules. La classification de Baltimore permet de distinguer les virus selon la nature de leur génome (ADN ou ARN, simple brin ou double brin) et la polarité positive ou négative du génome. Les virus peuvent aussi être classés selon l'ICTV (*International Committee on Taxonomy of Viruses* ou comité international de taxonomie des virus), permettant de distinguer l'ordre, la famille, le genre et l'espèce du virus. Le génome des virus peut également être segmenté comme par exemple pour le virus X de drosophile (DXV) ou le virus de la fièvre de la Vallée du Rift (RVFV). Les virus à ARN codent leur machinerie de réplication, et notamment l'enzyme RdRp (*RNA-dependent RNA-polymerase* ou ARN-polymérase ARN-dépendante) impliquée dans la réplication du génome viral. Malgré des divergences de séquences, les principales caractéristiques structurelles de RdRp sont conservées chez les virus à ARN. Cette enzyme n'a pas d'activité de relecture et entraîne un taux d'erreur élevé (10^{-6} à 10^{-4} substitutions par nucléotide par infection cellulaire) lors du processus de copie (Duffy, 2018; Sanjuán et al., 2010; Venkataraman et al., 2018). L'augmentation du taux de mutation entraîne une variabilité du génome viral et est corrélée à une diversité et une capacité d'évolution accrue. Dans la majorité des cas reportés, l'infection virale est asymptomatique chez l'hôte insecte (Fleuriet, 1981; Habayeb et al., 2009; Myles et al., 2008). De nombreux virus à ARN sont capables de contourner les défenses de l'hôte permettant la réplication et la propagation du virus (Berry et al., 2009; Kakumani et al., 2013; Schnettler et al., 2012; van Rij et al., 2006). Cependant, selon le tropisme viral, certains virus peuvent toutefois engendrer des symptômes ou la mort de l'insecte, c'est notamment le cas du virus C de drosophile (DCV), responsable d'occlusion intestinale entraînant la mort de la drosophile (Chtarbanova et al., 2014).

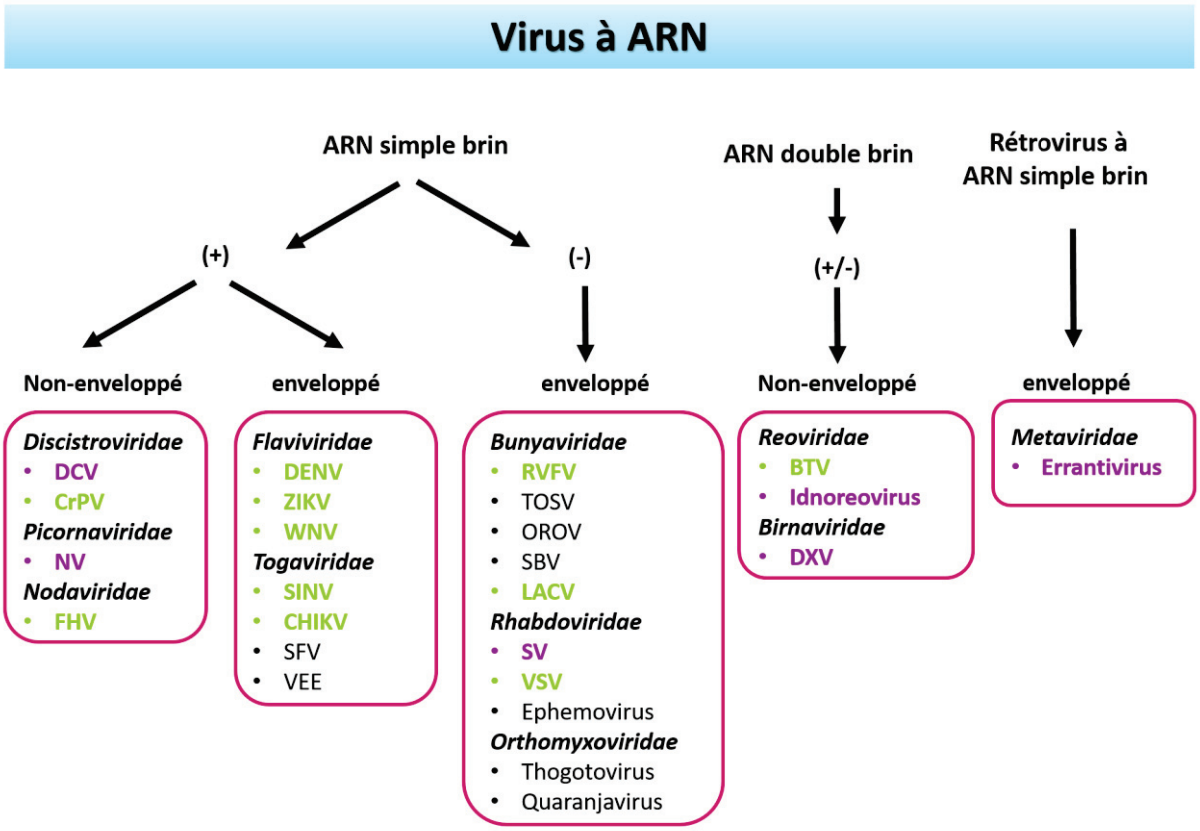


Figure 2. Une grande diversité de virus à ARN infecte les insectes. Les virus sont classés selon la nature de leur génome (classification de Baltimore) : ARN simple brin de polarité positive ou négative, ARN double brin. Ces virus ne présentent pas d'étape sous forme d'ADN, à l'exception des rétrovirus à ARN simple brin qui ont un intermédiaire ADN dans leur cycle de réplication. On les distingue aussi par la présence ou l'absence d'une enveloppe virale. De nombreuses familles virales infectent naturellement la drosophile (virus en violet) : *Discistroviridae* (*Drosophila C virus*, DCV), *Picornaviridae* (*Nora virus*, NV), *Rhabdoviridae* (*Sigma virus*, SV), *Reoviridae* (Idnoreovirus), *Birnaviridae* (*Drosophila X virus*, DXV), et *Metaviridae* (Errantivirus). Le modèle drosophile a été utilisé pour l'étude des infections expérimentales suivantes (virus en vert) : *Cricket Paralysis virus* (CrPV) (Nayak et al., 2010), *Flock House Virus* (FHV) (Goic et al., 2013), *Dengue Virus* (DENV) (Rancès et al., 2012), *Zika virus* (ZIKV) (Delorme-Axford and Klionsky, 2019), *West Nile Virus* (WNV) (Chotkowski et al., 2008), *Sindbis Virus* (SINV) (Galiana-Arnoux et al., 2006), *Chikungunya virus* (CHIKV) (Avadhanula et al., 2009), *Rift Valley Fever Virus* (RFV) (Dietrich et al., 2017), *La Crosse Virus* (LACV) (Glaser and Meola, 2010), *Vesicular Stomatitis Virus* (VSV) (Mueller et al., 2010), *Bluetongue virus* (BTV-1) (Shaw et al., 2012). Chez la drosophile, aucune étude n'a été reportée avec les virus suivants : *Venezuelan Equine Encephalitis virus* (VEE), *Toscana virus* (TOSV), *Oropouche virus* (OROV), *Schmallenberg virus* (SBV), Ephemovirus, Thogotovirus et Quaranjavirus. D'après (Hughes et al., 2012; Webster et al., 2015; Xu and Cherry, 2014). La liste des virus n'est pas exhaustive.

1. Les virus infectant naturellement la drosophile

Comme mentionné ci-dessus, les virus de drosophiles sont quasiment exclusivement représentés par des virus à ARN. Un seul virus à ADN a été identifié chez *D. melanogaster*. En effet, récemment, en utilisant une approche métagénomique, le virus Kallithea de la famille

Nudiviridae a été découvert (Webster et al., 2015). Par ailleurs, la drosophile peut être naturellement infectée par de nombreux virus à ARN, tels que DCV (Jousset et al., 1977), DAV (Ambrose et al., 2009), DXV (Teninges et al., 1979), le virus Nora (NV) (Habayeb et al., 2006) et le virus sigma (SV) (Carpenter et al., 2007) (**Figure 2., virus en violet**). Ces virus sont responsables d'infections chroniques définies par une longue durée et caractérisées par une réponse immunitaire généralement modérée. La prévalence de ces virus est importante puisque 40 % des populations sauvages de drosophile sont infectées par un ou plusieurs de ces virus (Plus et al., 1975). La communauté virale est très différente entre les drosophiles de laboratoire et celles de la nature avec une prévalence virale très importante pour les lignées de laboratoire (Webster et al., 2015). L'analyse du transcriptome à haut débit (RNA-seq) à partir de 9656 données publiques de drosophile de laboratoire et de culture cellulaire a notamment révélé environ 40 % d'échantillons positifs, incluant un seuil de détection virale supérieur à 100 lectures par million (Webster et al., 2015). Comme nous le détaillerons par la suite, certains de ces virus codent des suppresseurs de la voie d'ARNi. Il est indispensable de prendre en compte ces infections virales chroniques lors de l'étude des interactions hôte-pathogène et des voies d'ARNi (Merkling and van Rij, 2015).

A. Le virus sigma

Le virus sigma constitue un modèle d'étude pour mes travaux de thèse et sera introduit de façon plus conséquente que les autres virus spécifiques de drosophile. Le virus sigma de *Drosophila melanogaster* (DMelSV) est un virus enveloppé, avec une forme caractéristique en « balle de fusil », appartenant à l'ordre des *Mononegavirales* et à la famille des *Rhabdoviridae*. Contrairement à *D. simulans* qui n'est jamais infectée naturellement par DMelSV (Webster et al., 2015), d'autres espèces de drosophiles sont fréquemment infectées par le virus sigma, telles que *D. affinis*, *D. athabasca*, *D. obscura*, *D. tristis*, *D. immigrans* et *D. ananassae* (Longdon et al., 2012). Un niveau élevé de divergence entre ces Rhabdovirus permet de distinguer un virus sigma spécifique pour chacune de ces espèces (Longdon et al., 2010). Leur génome ARN simple brin, de polarité négative et d'environ 12 kb contient six gènes codant six protéines structurales (**Figure 3.**), nommées la nucléoprotéine (N), la protéine associée à la polymérase (P), la protéine X, la protéine de matrice (M), la glycoprotéine (G), et la polymérase (L) (Longdon et al., 2012; Teninges et al., 1993). La fonction du sixième gène, appelé X ou PP3 n'est pas connue mis à part son implication au niveau de l'immunité innée antivirale (Longdon et al., 2010; Tsai et al., 2008).

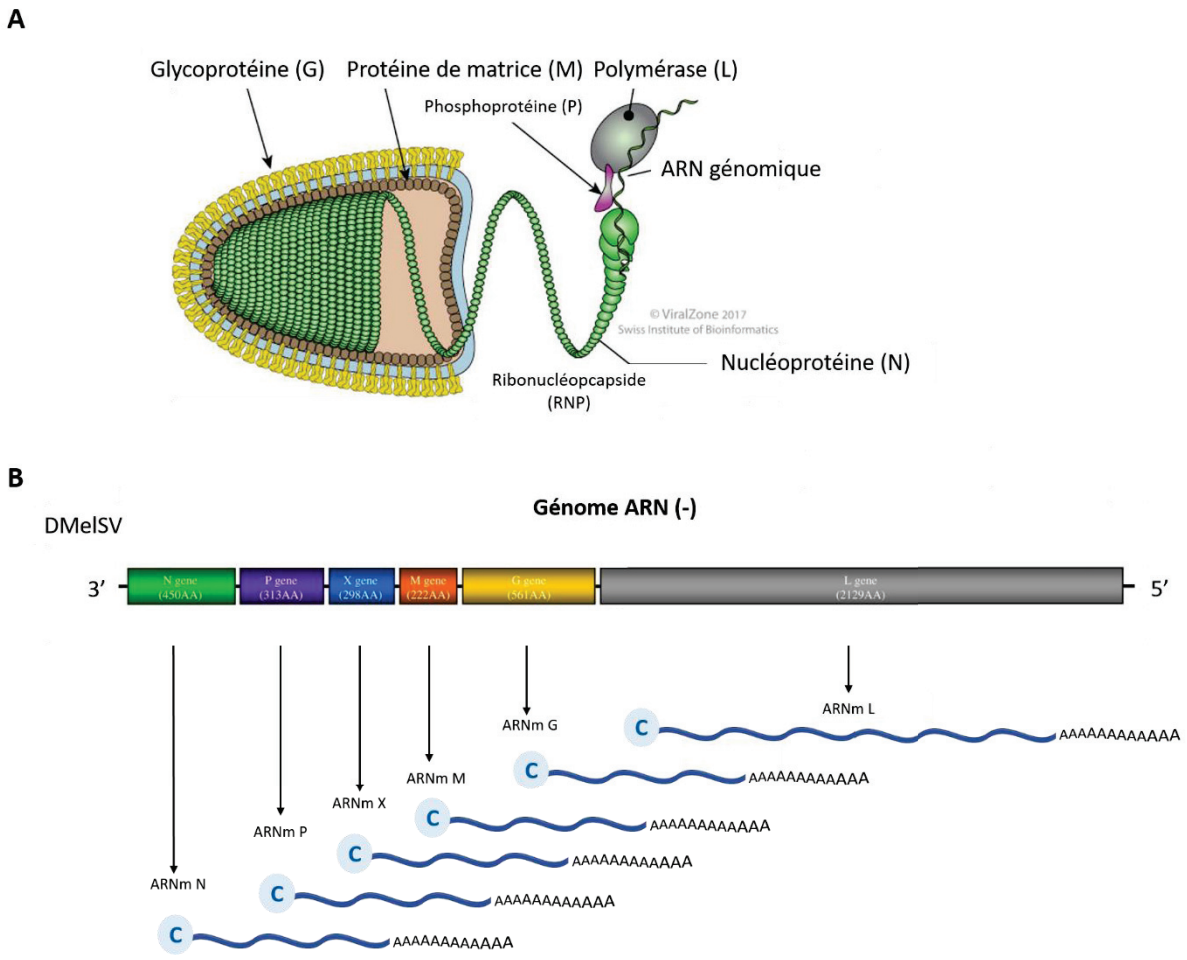


Figure 3. Schéma du virion et du génome du virus sigma. Le virion est une particule virale infectieuse libérée par la cellule infectée dont la machinerie génétique a été détournée pour sa production. **(A)** L'enveloppe du virion du virus sigma contient la glycoprotéine de surface (G). La protéine de matrice (M) participe à l'assemblage de la particule virale infectieuse, à la réplication du génome et module les voies de transcription, traduction et d'apoptose de la cellule hôte au bénéfice du virus (Graham et al., 2008). La nucléocapside est constituée de la nucléoprotéine (N), la phosphoprotéine (P), la protéine PP3 (X), et de la polymérase (L) et englobe le génome viral, constituant le complexe ribonucléoprotéique viral indispensable à la réplication du génome viral. **(B)** Le génome ARN du virus sigma, de polarité négative code six protéines structurales (N, P, X, M, G et L). La transcription s'effectue à partir de 6 cadres de lecture ouverts (ORF) permettant la transcription de 6 ARN messager (ARNm) et la traduction de 6 protéines. Chaque ARNm viral possède une coiffe en 3' et est polyadénylé en 5'. D'après (Longdon et al., 2010; Piontkivska et al., 2016). Images issues de Viral Zone 2017: https://viralzone.expasy.org/4656).

La caractérisation précise du cycle de réplication de DMelSV, incluant le nombre de virions transmis entre générations chez la drosophile, n'est pas clairement définie. Cependant, le mécanisme de réplication des Rhabdovirus est assez conservé au sein de cette famille et est documenté dans la revue de Dietzgen et collaborateurs (Dietzgen et al., 2017). Lors du cycle de réplication d'un Rhabdovirus, incluant DMelSV (**Figure 4.**), la glycoprotéine d'enveloppe permet l'attachement du virion à la cellule, aboutissant à l'endocytose de la particule virale

(Lyles, D.S., and Rupprecht, C.E., 2007). L'endosome présente une forte acidité permettant un changement de conformation de la protéine G, devenant une protéine de fusion (Roche et al., 2007, 2006). La fusion entre la membrane de l'endosome et la membrane virale induit la libération de la nucléocapside dans le cytoplasme (Luyet et al., 2008). Le complexe ribonucléoprotéique, incluant l'ARN génomique de polarité négative, encapsidé par la protéine N est ensuite libéré dans le cytoplasme et sert de matrice pour la transcription en ARN messager (ARNm). L'ARN-polymérase ARN-dépendante L permet la transcription du premier gène (c'est-à-dire du gène codant la protéine N) puis des 5 autres gènes (P-X-M-G-L). Lors de cette première transcription, l'ensemble des transcrits viraux est polyadénylé et présente une coiffe (Rose, 1980). Les ARNm viraux peuvent être pris en charge par les ribosomes cellulaires, permettant leur traduction en protéines virales. La protéine L permet ensuite la réplication du génome, aboutissant à la formation de copies antigénomiques de polarité positive. La protéine N est également impliquée dans l'initiation de la réplication du génome viral (Arnheiter et al., 1985). En effet, lorsqu'il y a assez de protéine N pour permettre l'encapsidation de l'ARN nouvellement synthétisé, la polymérase enclenche la production d'antigénome sur toute la longueur du génome, servant de matrice pour la réplication du génome viral (Fearns et al., 1997). La deuxième transcription du génome viral permet la synthèse d'ARNm à partir des 6 cadres de lecture ouverts. En parallèle, certains de ces ARN génomiques nouvellement synthétisés sont pris en charge par les protéines virales, permettant la production de nouveaux virions, pouvant infecter d'autres cellules après bourgeonnement à partir de la première cellule (Harty et al., 2001).

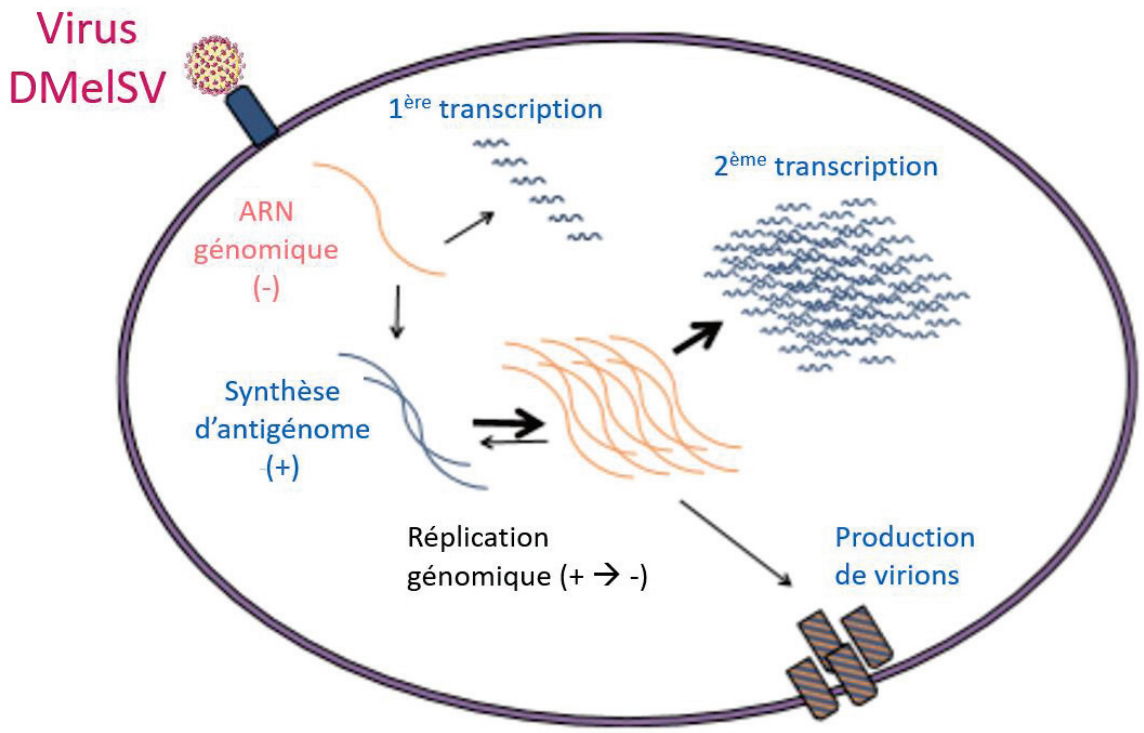


Figure 4. Schéma du cylce de réplication de DMelSV. D'après (Piontkivska et al., 2016).

Le virus DMelSV a été découvert en 1937 chez *D. melanogaster* (L'Heritier, 1948). Les drosophiles infectées par DMelSV présentent une hypersensibilité au dioxyde de carbone (CO_2), entraînant une paralysie puis la mort (L'Heritier, 1948). A l'inverse, les drosophiles non infectées par DMelSV ne présentent pas de sensibilité au CO_2. La paralysie est la conséquence des dommages causés aux tissus nerveux, en particulier aux ganglions thoraciques lors de la réplication de DMelSV et suite à l'exposition au CO_2 (Tsai et al., 2008). En effet, DMelSV se lie aux récepteurs à acétylcholine du tissu nerveux pouvant potentiellement être impliqués dans le symptôme de paralysie (Brun and Plus, 1980). Cependant, le mécanisme exact reste inconnu. DMelSV se réplique faiblement dans des cellules embryonnaires de drosophile mais ne présente pas d'effet cytopathique (Brun and Plus, 1980; Ohanessian and Echalier, 1967). Lors de l'infection par DMelSV, l'hypersensibilité des drosophiles au CO_2 permet d'utiliser le test CO_2 pour déterminer si les drosophiles sont infectées ou non. Les drosophiles sont soumises à 100% de CO_2 puis sont posées sur la glace pendant 15 minutes. Les drosophiles sont ensuite exposées à l'air ambiant. L'observation des drosophiles 24h après le test permet de visualiser les drosophiles infectées, correspondant à celles paralysées ou déjà mortes. Le virus DMelSV est non létal chez la drosophile sans exposition au CO_2 mais il engendre une diminution de la viabilité des œufs (Wayne et al., 2011), une augmentation du temps de développement (Fleuriet, 1996), et une réduction d'aptitude ou « fitness » des drosophiles dans des conditions de surpeuplement (Yampolsky et al., 1999).

Une caractéristique très intéressante de DMelSV est sa transmission verticale par l'intermédiaire du sperme et des œufs de drosophile (Brun and Plus, 1980; Fleuriet, 1988; L'heritier, 1958; Longdon et al., 2010) alors que la transmission horizontale est plutôt rare ou inexistante. Cette transmission par les mâles et les femelles permet au DMelSV de persister et de se propager dans les populations de drosophile (L'heritier, 1958; Longdon et al., 2017).

L'ensemble des virus infectant naturellement la drosophile est important à condidérer lors de l'étude d'une infection virale chez *Drosophila*. De plus, au cours du travail expérimental, les voies d'ARNi seront précisément étudiées, ce qui implique une connaissance précise des éventuels suppresseurs d'ARNi codés par ces virus.

B. Le virus C de drosophile (DCV)

DCV fera également l'objet d'une étude bioinfomatique au cours de mes travaux de recherche. DCV appartient à la famille des *Dicistroviridae*. Il n'est pas enveloppé et comporte un génome ARN simple brin, de polarité positive. C'est un modèle viral très utilisé en laboratoire pour étudier les interactions hôte-virus (Kemp et al., 2013; Mierlo et al., 2012; van Rij et al., 2006). La protéine 1A de DCV est un suppresseur de la voie d'ARNi, empêchant le clivage des ARN double brin par Dicer-2 (Dcr-2), et inhibant la production de siARN (van Rij et

al., 2006). DCV est très pathogène pour la drosophile, incluant un stress nutritionnel qui provient d'une obstruction intestinale et pouvant conduire à la mort des drosophiles infectées (Chtarbanova et al., 2014). Aucune donnée n'a révélé l'infection de la lignée germinale et la transmission verticale de DCV (Gomariz-Zilber et al., 1995; Plus et al., 1975). De plus, malgré une large prévalence dans les lignées de laboratoire, DCV n'a pas été détecté au niveau d'un large panel de drosophiles de la nature (*D. melanogaster* et *D. simulans*), provenant de 17 zones dans le monde, suggérant une faible représentation dans la nature (Merkling and van Rij, 2015; Webster et al., 2015).

C. Le virus X de drosophile (DXV)

DXV est un virus non-enveloppé qui appartient à la famille des *Birnaviridae* et comporte un génome ARN double brin bi-segmenté (Zambon et al., 2005). Il code un suppresseur d'ARN interférence, la protéine VP3 (Viral protein 3) empêchant la protéine Dcr-2 d'induire la production de siARN (van Cleef et al., 2014). Les drosophiles infectées par DXV, comme pour DMelSV, développent une sensibilité aiguë à l'anoxie, induisant la mort des drosophiles (Zambon et al., 2006).

D. Le virus A de drosophile (DAV)

DAV est un virus ARN simple brin, de polarité positive. Il est très peu étudié et la caractérisation de son génome n'a pas été reportée. N'étant pas spécifiquement classé, il est relié à la famille des *Permutotetraviridae* (Ambrose et al., 2009). Les drosophiles de la nature et de laboratoire sont fréquemment infectées de façon persistante par DAV (Merkling and van Rij, 2015; Plus et al., 1975).

E. Le virus Nora (NV)

NV est un virus non-enveloppé contenant un génome ARN simple brin, de polarité positive, et appartenant à l'ordre des *Picornavirales*. NV est responsable d'infections persistantes, chez les drosophiles de laboratoire et les drosophiles sauvages (Habayeb et al., 2006; Webster et al., 2015). Lors d'une infection persistante, le virus n'est pas éliminé mais se réplique au niveau de cellules spécifiques, entraînant peu d'effet sur les traits d'histoire de vie de l'hôte (Roekring et al., 2006). NV se réplique au niveau des intestins des drosophiles et est excrété dans les fèces, permettant une transmission horizontale, mais n'est pas capable de se transmettre verticalement (Habayeb et al., 2006). Il a été suggéré que la voie des siARN serait stimulée pendant l'infection persistante, par l'intermédiaire de la production de siARN viraux distincts

et plus abondants (Goic et al., 2013). Par ailleurs, la protéine VP1 (Viral protein 1) du NV est un suppresseur des voies d'ARNi qui bloque l'action de la protéine Argonaute 2 (Ago2) responsable du clivage de l'ARNm viral cible (Mierlo et al., 2014). Cette persistance virale suggère un équilibre entre les voies de défense antivirale et la réplication du virus aboutissant à une réplication atténuée du virus.

2. Les virus n'infectant pas naturellement la drosophile

Les populations de drosophile de laboratoire peuvent être expérimentalement infectées par de nombreux virus à ARN (**Tableau 1.**), permettant d'élargir les études des interactions hôte-virus en utilisant le modèle drosophile. Parmi ces virus, on trouve des virus d'insecte tels que *Cricket paralysis virus* (CrPV) (Nayak et al., 2010) ; *Flock house virus* (FHV) (Dasgupta et al., 2007) et de nombreux arbovirus (décrits dans la partie II.3), notamment *Sindbis virus* (SINV) ; *Vesicular stomatitis virus* (VSV) ; *Chikungunya virus* (CHIKV) ; *Rift Valley Fever Virus* (RVFV) ; *La Crosse Virus* (LACV) ; *West Nile Virus* (WNV) et *Dengue Virus* (DENV).

CrPV est un virus à ARN simple brin de polarité positive, infectant naturellement le criquet (Reinganum, 1975) et pouvant expérimentalement infecter la drosophile après injection intra-thoracique (Galiana-Arnoux et al., 2006; Manousis and Moore, 1987; van Rij et al., 2006). De la même façon que pour DCV, CrPV appartient à la famille des *Dicistroviridae* et code un suppresseur de la voie d'ARNi, nommé CrPV-1A (Nayak et al., 2010). FHV est un autre virus non spécifique de drosophile, appartant à la famille des *Nodaviridae* qui a été isolé à partir d'insectes de l'ordre des Coléoptères (Scotti et al., 1983). FHV peut infecter les cellules de drosophile et les mouches adultes après injection intrathoracique et code également un suppresseur d'ARNi (protéine B2) (Dasgupta et al., 2007; Galiana-Arnoux et al., 2006; Goic et al., 2013). De plus, le virus *Invertebrate Iridescent Virus* 6 (IIV-6) est le seul virus à ADN qui se réplique expérimentalement chez *Drosophila*, induisant la mort des drosophiles infectées (Teixeira et al., 2008). IIV-6, appartenant à la famille des *Iridoviridae*, infecte de nombreux insectes tels que les insectes de l'ordre des Lépidoptères (Jakob et al., 2002).

Parmi les arbovirus décrits pour leur capacité à infecter expérimentalement *Drosophila*, on trouve notamment des membres de la famille des *Togaviridae* et du genre Alphavirus tel que SINV (Galiana-Arnoux et al., 2006; Petit et al., 2016; Tassetto et al., 2017), qui sera le modèle viral principal de mes travaux de thèse (décrit dans la partie II.3.A). Après infection exérimentale par SINV, de nombreux tissus sont infectés, incluant le corps gras et les cellules musculaires, bien qu'aucune mortalité n'ait été décrite chez la drosophile (Galiana-Arnoux et al., 2006). De plus, aucun suppresseur d'ARNi n'a été décrit pour SINV. Par ailleurs, un autre alphavirus, CHIKV, se réplique chez la drosophile après inoculation virale (Avadhanula et al., 2009; Phillips et al., 2010) dont les protéines non structurales nsp2 et nsP3 présentent une activité suppressive d'ARNi (Mathur et al., 2016).

D'autres arbovirus se répliquent chez *Drosophila* après inoculation expérimentale, incluant des membres de la famille des *Flaviviridae*, tels que les virus WNV (Chotkowski et al., 2008), DENV (Rancès et al., 2012) et *Zika virus* (ZIKV) (Delorme-Axford and Klionsky, 2019; Harsh et al., 2018), un virus de la famille des *Phenuiviridae*, RVFV (Dietrich et al., 2017; Filone et al., 2010) et un virus de la famille des *Bunyaviridae*, LACV (Glaser and Meola, 2010). De plus, VSV est un virus appartenant à la même famille que le virus sigma, soit la famille des *Rhabodoviridae*, et il se réplique dans l'ensemble des cellules et des hôtes testés, impliquant une glycoprotéine d'enveloppe ubiquitaire. Cette caractéristique fait du VSV un outil très intéressant pour l'étude des interactions hôte-virus (Mueller et al., 2010). Un seul arbovirus à ARN double brin capable d'infecter *Drosophila* a été décrit à ce jour. Il s'agit de *Bluetongue virus* (BTV), un virus de la famille des *Reoviridae* (Shaw et al., 2012).

Virus	Infection naturelle	Famille	Genre	Génome	Pathogénicité chez *Drosophila*	Suppresseur d'ARNi
CrPV	Virus d'insecte	*Dicistroviridae*	Cripavirus	ARN + simple brin	mortalité	Oui
FHV	Virus d'insecte	*Nodaviridae*	Alphanodavirus	ARN + simple brin	mortalité	Oui
SINV	arbovirus	*Togaviridae*	Alphavirus	ARN + simple brin	asymptomatique	Non documenté
CHIKV	arbovirus	*Togaviridae*	Alphavirus	ARN + simple brin	asymptomatique	Oui
WNV	arbovirus	*Flaviviridae*	Flavivirus	ARN + simple brin	asymptomatique	Oui
DENV	arbovirus	*Flaviviridae*	Flavivirus	ARN + simple brin	asymptomatique	Oui
ZIKV	arbovirus	*Flaviviridae*	Flavivirus	ARN + simple brin	asymptomatique	Non documenté
RVFV	arbovirus	*Phenuiviridae*	Phlebovirus	ARN - simple brin	Non documenté	Non trouvé
LACV	arbovirus	*Bunyaviridae*	Orthobunyavirus	ARN - simple brin	asymptomatique	Non documenté
VSV	arbovirus	*Rhabdoviridae*	Vesiculovirus	ARN - simple brin	asymptomatique	Non documenté
BTV	arbovirus	*Reoviridae*	Orbivirus	ARN double brin	asymptomatique	Non documenté

Tableau 1. Virus étudiés expérimentalement chez *Drosophila*.

3. Arbovirus

Arbovirus est l'acronyme d'ARthropode-BOrne VIRUS, ce qui indique que ces virus sont véhiculés et transmis par des arthropodes hématophages, principalement des moustiques, des tiques et des culicoïdes (Gubler, 2001). Ces virus infectent naturellement leur vecteur, incluant une réplication virale et permettant la transmission virale à des hôtes vertébrés lors

d'un repas sanguin. Les arbovirus sont presque exclusivement représentés par des virus contenant un génome ARN. Par ailleurs, il existe une grande diversité d'arbovirus, représentés par des familles distinctes, incluant une variabilité de séquences et de structures virales. Ces virus sont très étudiés pour leur impact sur la santé humaine, l'agriculture et en recherche fondamentale utilisant notamment le modèle drosophile (décrits dans la partie II.2). Les arbovirus représentent une importante menace pour la santé humaine et animale, incluant notamment les Alphavirus (tels que les virus CHIKV et SINV), les Flavivirus (par exemple les virus ZIKV et DENV) ou encore les Phlébovirus (tel que RVFV). Ils sont responsables d'épidémies dans le monde, incluant des taux élevés de mortalité et de morbidité chez l'Homme. Ils impactent également directement l'économie, notamment en provoquant des pertes au sein des élevages et des cultures, pouvant aboutir à des embargos commerciaux (Chengula et al., 2013).

A. Le virus Sindbis

Nous détaillons ici particulièrement le cas de SINV, qui sera utilisé par la suite dans le travail expérimental. SINV a été isolé pour la première fois à partir de moustiques ornithophiles des espèces *Culex pipiens* et *Culex univittatus*, collectés en 1952 en Egypte (Taylor et al., 1955). Il est largement distribué dans le monde, notamment en Afrique, Australie, Allemagne, Russie, Norvège, Finlande et Suède (Adouchief et al., 2016; Lundström and Pfeffer, 2010; Lwande et al., 2015).

Il appartient à la famille des *Togaviridae* et au genre Alphavirus. Les alphavirus sont classiquement décrits en deux groupes selon leur distribution géographique, les virus du Vieux Monde et les virus du Nouveau Monde. Les virus du Vieux Monde, incluant SINV, provoquent généralement de la fièvre, des éruptions cutanées, et des arthralgies, tels que des troubles musculo-squelettiques chez l'Homme (Adouchief et al., 2016; Laine et al., 2004) alors que les virus du Nouveau Monde peuvent engendrer des encéphalites (Zacks and Paessler, 2010). En Suède, la maladie causée par SINV est nommée maladie d'Ockelbo (Niklasson et al., 1984). SINV, comme d'autres alphavirus, peut engendrer des arthralgies qui persistent sur plusieurs années (Adouchief et al., 2016; Kurkela et al., 2008a).

La distribution géographique de SINV est contrainte par différents facteurs écologiques impliquant le réservoir hôte, le vecteur et le virus. Les oiseaux migrateurs, résidents et sauvages sont des hôtes amplificateurs importants de SINV, notamment en Europe du Nord (Kurkela et al., 2008b). Les moustiques ornithophiles des genres *Culex* et *Culiseta* ont été décrits comme les principaux vecteurs, incluant un taux d'infection très élevé chez le moustique *Culex torrentium* (Hesson et al., 2015; Lwande et al., 2019) (**Figure 5.**). SINV a également été détecté chez les moustiques *Aedes* adultes et au stade larvaire (Francy et al., 1989; L'vov et al., 1985; Sang et al., 2017; Tingström et al., 2016). Ces résultats indiquent que

les genres *Aedes* et *Culiseta* peuvent également être des vecteurs de SINV et qu'une transmission verticale à partir des œufs (transovarienne) est possible (Dhileepan et al., 1996; Lwande et al., 2019).

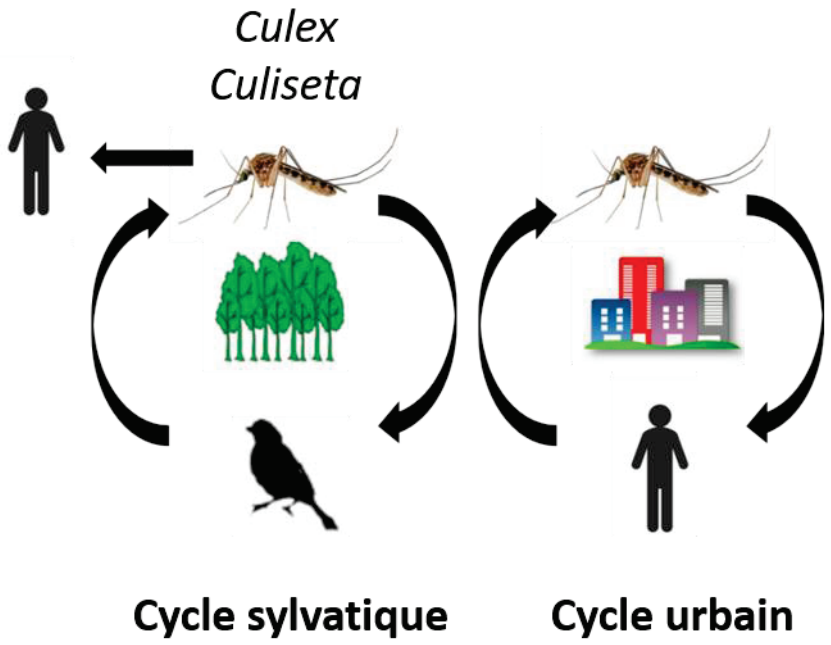


Figure 5. Transmission virale de SINV. On distingue deux cycles de transmission virale selon l'environnement. Le cycle sylvatique permet la circulation de SINV entre les vecteurs moustiques, principalement des genres *Culex* et *Culiseta* (Francy et al., 1989; Lwande et al., 2019) et la faune aviaire, un réservoir important de SINV, permettant l'amplification et la distribution de SINV (Lundström et al., 2001). Au cours d'un événement accidentel, l'Homme peut être infecté. A l'inverse, dans le cycle urbain de transmission, SINV circule entre les vecteurs moustiques et l'hôte humain. Aucune transmission interhumaine n'a été rapportée (Lim et al., 2018).

Les particules virales de SINV sont enveloppées, avec une structure icosaédrique et présentent une petite taille, comprise entre 50 et 70 nm de diamètre. Le génome de SINV est composé d'un ARN simple brin non ségmenté, de polarité positive, et d'une longueur d'environ 11,5 kb, incluant une coiffe en 5' et une queue poly(A) (Strauss et al., 1984). Le génome de SINV contient deux cadres de lecture ouverts, correspondant à l'ARN génomique (ARN 49S) codant les protéines non-structurales (nsP1, nsP2, nsP3 et nsP4) nécessaires à la machinerie de réplication virale, et à l'ARN sous-génomique (ARN 26S) codant les protéines structurales permettant la formation des particules virales et incluant la capside (C), les glycoprotéines d'enveloppe E1, E2 et E3 et un petit peptide 6K (Rupp et al., 2015) (**Figure 6.**). La réplication de l'ARN génomique révèle le promoteur sous-génomique (SGP), se situant sur le brin négatif de l'ARN génomique (Levis et al., 1990).

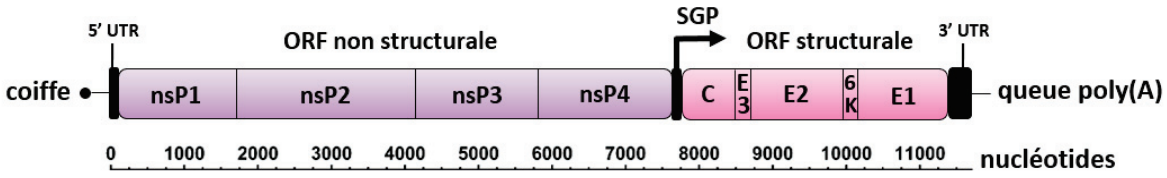


Figure 6. Schéma de la structure du génome de SINV. C'est un génome ARN simple brin de polarité positive codant deux cadres de lecture ouverts. Le premier (en violet) code les protéines non structurales (nsp1-4) et le deuxième code les protéines structurales (capside (C), glycoprotéines d'enveloppe (E1-3) et un petit peptide de 6 kDa (6K)). SGP (Subgenomic promoter), pour promoteur sous-génomique. D'après (Pietilä et al., 2017).

Lors du cycle de réplication du virus SINV (**Figure 7.**), l'attachement de la particule virale à la cellule implique la glycoprotéine d'enveloppe E2 se liant à des récepteurs protéiques cellulaires et l'entrée se fait par endocytose médiée par des clathrines (Brown et al., 2018; Kielian et al., 2010). Le récepteur cellulaire varie selon l'espèce d'alphavirus. Il a été décrit que la protéine NRAMP2, pour *Natural Resistance-Associated Macrophage Protein 2* est le récepteur spécifique de SINV chez la drosophile (Panda et al., 2013; Rose et al., 2011). Cependant, l'homologue de NRAMP2 n'a pas été identifié chez le moustique *Aedes* ce qui suggère que NRAMP2 n'est pas le récepteur de SINV au niveau de son vecteur, tel que *Ae. aegypti* ou d'autres moustiques de la famille des *Culicidae* (Tsujimoto et al., 2018). L'attachement peut également être facilité par des liaisons de type héparane sulfate. La particule virale est ensuite internalisée au sein d'une vésicule d'endocytose et la maturation de cet endosome va engendrer une augmentation de l'acidité. A faible pH, la conformation des dimères E2/E1 est modifiée permettant ensuite la fusion entre les membranes virales et cellulaires et engendre la libération de la nucléocapside dans le cytoplasme. La nucléocapside est ensuite dissociée et l'ARN génomique libéré dans le cytoplasme.

Nous allons décrire plus précisément la réplication du génome viral (**Figures 7 et 8.**). L'ARN génomique est traduit en une polyprotéine nommé P1234. Au début de l'infection, la protéase nsP2 clive en *cis* cette polyprotéine entre nsP3 et nsP4, induisant la production d'un complexe de réplication initial, composé de la polyprotéine P123 et de nsP4 (de Groot et al., 1990; Hahn et al., 1989; Hardy and Strauss, 1989). Ce complexe permet la synthèse du brin négatif d'ARN, qui sert de matrice pour la synthèse de l'ARN génomique et sous-génomique (Strauss and Strauss, 1994). Lorsque la polyprotéine P123 atteint une concentration élevée, le clivage par nsP2 en *trans* permettant la libération de nsP1 et la formation de P23, induit la synthèse du brin d'ARN positif incluant la production des ARN génomique et sous-génomique (Lemm et al., 1994; Shirako and Strauss, 1990). La polyprotéine P23 a une demi-vie très courte (quelques secondes) (Hardy et al., 1990). Par conséquent, le clivage de P23 en nsP2 et nsP3 permet la production de chaque nsP individuellement et forme le complexe principal de réplication permettant la synthèse de l'ARN viral, et incluant les ARN génomique et sous-génomique avec

une production en excès de l'ARN sous-génomique par rapport au génome viral (Lemm et al., 1994; Shirako and Strauss, 1994).

nsP1 est l'enzyme permettant le coiffage de l'ARN viral par l'intermédiaire de son domaine N-terminal contenant un motif méthyltransférase (LaPointe et al., 2018; Rozanov et al., 1992; Schluckebier et al., 1995). nsP1 est également associée aux membranes de l'hôte servant probablement de membrane d'ancrage du complexe de réplication (Salonen et al., 2003). nsP2 est une ARN hélicase et la protéase responsable de la maturation de la poly-protéine P1234 (Rupp et al., 2015). nsP3 interagit avec de nombreuses protéines cellulaires et module l'ADP-ribosylation de protéines (Lark et al., 2018), et nsP4 est l'ARN-polymérase ARN-dépendante virale, nécessaire à la synthèse d'ARN (Rubach et al., 2009).

Le complexe de réplication est assemblé à partir de ces protéines non structurales permettant la réplication de l'ARN génomique incluant la synthèse de l'intermédiaire d'ARN de polarité négative. La synthèse d'ARN viral se produit au niveau de la membrane d'inclusions intracellulaires ou de vacuoles cytoplasmiques de type I (CPVI) (Pietilä et al., 2017). Le site du promoteur sous-génomique est sur le brin négatif de l'ARN génomique permettant la production d'un ARN sous-génomique, correspondant à environ un tiers du génome qui est ensuite traduit en une polyprotéine structurale (Levis et al., 1990). Des étapes de clivage permettent la libération de cinq protéines structurales correspondant à la protéine de capside, le précurseur de la glycoprotéine (pE2), la glycoprotéine E1 et deux petits peptides, 6K et TF, pour *transframe*. A l'extrémité N-terminal de cette polyprotéine structurale, la protéine de capside est une autoprotéase qui se clive elle-même de façon co-traductionnelle (Melancon and Garoff, 1987). La protéine de capside, traduite la première à partir de la poly-protéine, permet l'encapsidation de l'ARN génomique au sein de la nucléocapside (Lee et al., 1996; Weiss et al., 1989). La polyprotéine restante contenant les protéines d'enveloppe est insérée dans le réticulum endoplasmique, impliquant notamment la séquence signal d'E3 (Bonatti et al., 1984). Des signaux peptidase de l'hôte libèrent pE2, 6K et E1 (Lobigs et al., 1990; Melancon and Garoff, 1987). La protéine 6K agit également comme une séquence signal pour le traitement de la protéine E1 (Sanz et al., 2003). pE2 et E1 interagissent et forment des hétérodimères (Strauss and Strauss, 1994; von Bonsdorff and Harrison, 1975). Ces complexes hétérodimères sont ensuite transportés depuis le réticulum endoplasmique jusqu'à la surface cellulaire par l'intermédiaire du complexe de Golgi (de Curtis and Simons, 1988; Green et al., 1981; Sariola et al., 1995). Au temps tardif du transport, le précurseur E2 est clivé par une protéase cellulaire de type furine et engendre la production des protéines matures E2 et E3 (Zhang et al., 2003). Suite à ce clivage, l'interaction entre E1 et E2 est affaiblie, initiant l'activation du peptide fusion lors d'une exposition à un faible pH (Wahlberg et al., 1989). Les protéines structurales et l'ARN génomique s'assemblent et forment des particules virales

enveloppées qui après bourgeonnement sont libérées de la cellule et pourront ainsi infecter d'autres cellules (Brown et al., 2018; Fuller, 1987).

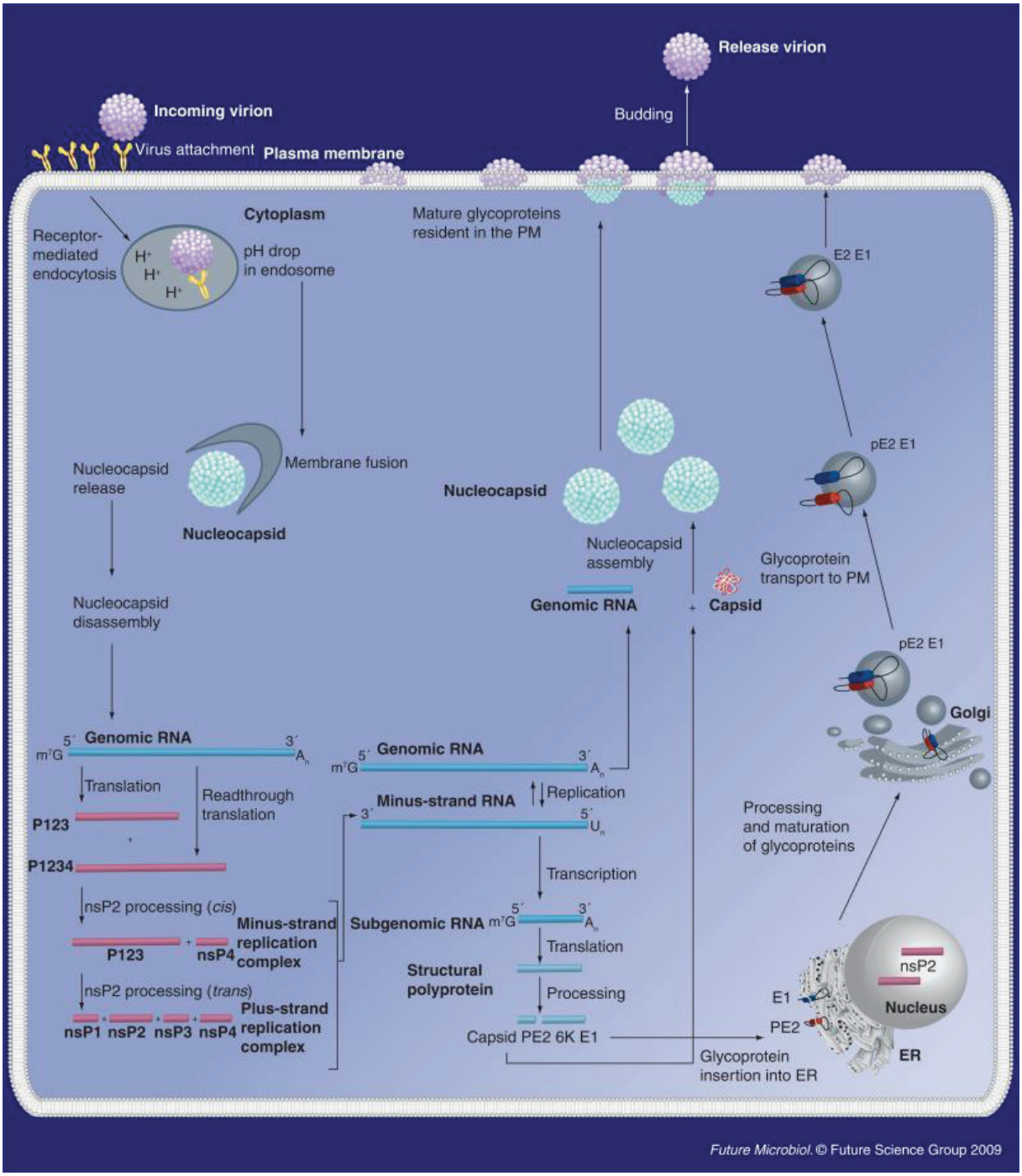


Figure 7. Cycle de réplication de SINV. La particule virale s'attache aux récepteurs cellulaires par l'intermédiaire de la glycoprotéine d'enveloppe E2. Les récepteurs de type clathrine permettent ensuite l'endocytose de la particule virale au sein de la cellule. Après une augmentation de l'acidité de l'endosome, la fusion entre l'enveloppe virale et la membrane cellulaire s'effectue et aboutit à la libération de la nucléocapside virale puis de l'ARN génomique dans le cytoplasme. Les protéines non structurales sont ensuite traduites et maturées dans le cytoplasme. Elles permettent l'activation de la réplication de l'ARN génomique ainsi que la traduction de

l'ARNm sous-génomique en protéines structurales (capside, pE2, 6K et E1). Les glycoprotéines traversent le réticulum endoplasmique et sont transportées à travers le Golgi jusqu'à la membrane plasmique. L'assemblage cytoplasmique de l'ARN génomique et de la capside constitue le corps de la nucléocapside qui s'associe ensuite aux glycoprotéines matures, permettant le bourgeonnement et la libération de la particule virale. ER: réticulum endoplasmique; nsP: protéine non structurale; PM: membrane plasmique. D'après (Jose et al., 2009).

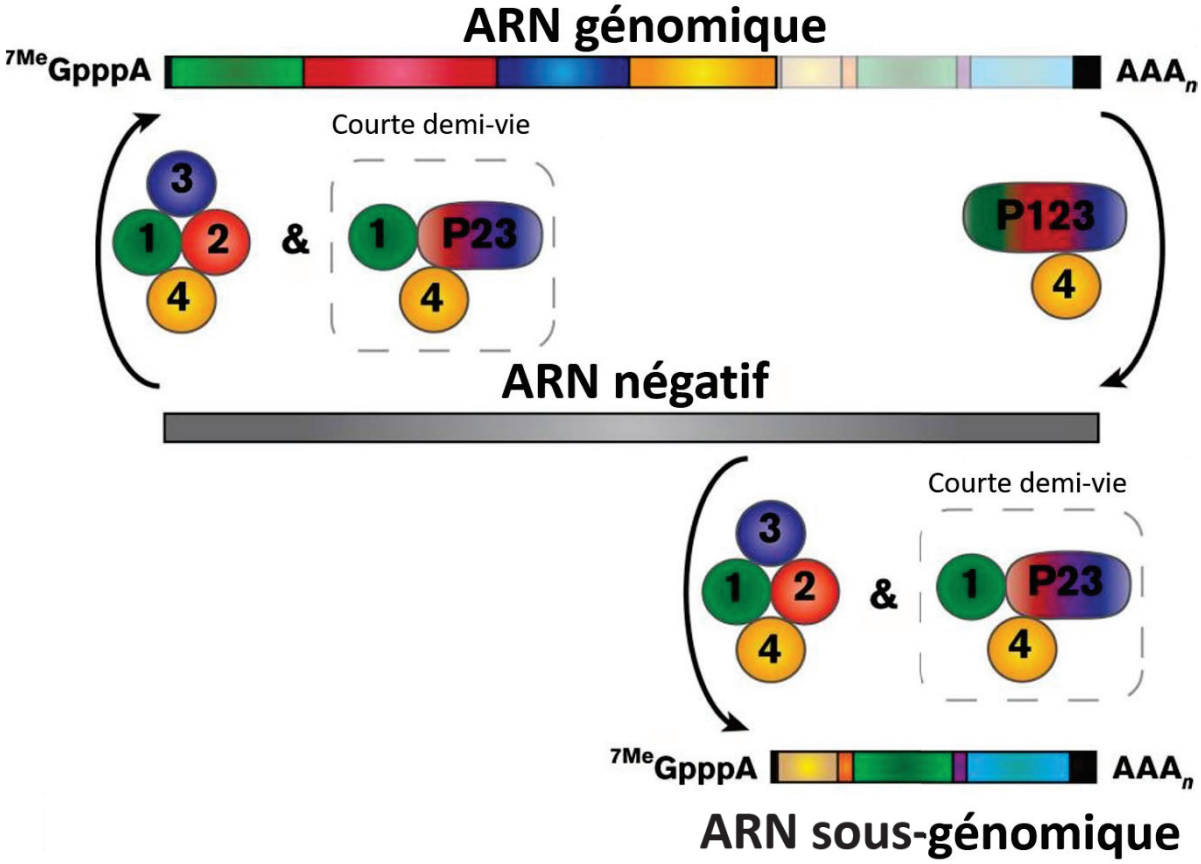


Figure 8. Réplication du génome de SINV. L'ARN génomique simple brin de SINV contient une coiffe de type O 5' 7-methylguanylate (^{7Me}GpppA) (Hefti et al., 1975) et est polyadénylé à l'extrémité 3'. Après l'infection d'une cellule, la traduction de l'ARN génomique permet la production de la polyprotéine P1234. La protéase nsP2 clive cette polyprotéine et forme le complexe de réplication initial, composé de P123 et nsP4 (en haut à gauche). Ce complexe initie la synthèse du brin d'ARN négatif, utilisé comme matrice pour la synthèse des ARN génomique et sous-génomique. La polyprotéine P123 est clivée et libère nsP1 et P23, présentant une demi-vie courte et permettant la synthèse du brin positif des ARN génomique et sous-génomique. Le clivage entre nsP2 et nsP3 permet la production du complexe principal de réplication (nsP1, nsP2, nsP3 et nsP4) impliqué dans la synthèse des ARN génomiques et antigénomiques. D'après (Rupp et al., 2015).

B. Interaction entre le virus et le vecteur insecte

Le cycle de transmission d'un arbovirus implique un premier repas sanguin au cours duquel le vecteur arthropode ingère du sang contaminé d'un hôte vertébré. Le virus se réplique dans l'organisme du vecteur et surmonte la première barrière intestinale, correspondant à la barrière d'infection de l'intestin moyen (MIB, *midgut infection barrier*), permettant son infection. La dissémination virale au sein de l'épithélium de l'intestin moyen aboutit à l'amplification du virus, qui va ensuite surmonter la barrière d'échappement de l'intestin moyen (MEB, *midgut escape barrier*) (pour revue : (Franz et al., 2015)). A l'extérieur de l'intestin moyen, le virus se propage et se réplique au niveau des tissus secondaires, pouvant ensuite induire l'infection des glandes salivaires (Myles et al., 2004). Le vecteur est à ce moment-là capable de transmettre le virus après une piqûre (deuxième repas sanguin) à d'autres hôtes vertébrés. La compétence vectorielle d'un vecteur arthropode inclut la capacité d'infection du vecteur, la réplication et propagation du virus dans les tissus du vecteur, ainsi que la transmission virale à un nouvel hôte vertébré (Beerntsen et al., 2000; Bennett et al., 2005). Seuls les arthropodes femelles sont hématophages car l'ingestion de sang est impliquée dans le processus d'ovogénèse (Jason Pitts, 2014). Le vecteur contrôle ces infections virales, permettant d'améliorer sa durée de vie et de diminuer au minimum les coûts pour son organisme. L'équilibre entre les réponses antivirales et la réplication virale permet la persistance du virus. Les arbovirus sont en très grande majorité non pathogènes chez l'hôte vecteur, impliquant notamment un système immunitaire inné très performant et permettant une transmission et une dissémination importante à travers le monde (Kim et al., 2004; Myles et al., 2008).

De nombreuses études s'intéressent aux interactions entre les arbovirus et leurs hôtes vertébrés, permettant l'identification des facteurs importants pour la réplication virale chez l'hôte infecté. Par ailleurs, les interactions entre les arbovirus et leurs vecteurs arthropodes, bien que très peu caractérisées, sont de plus en plus étudiées et améliorent les connaissances à propos de la transmission virale et des maladies endémiques et épidémiques sous-jacentes. Il s'agit d'une importante opportunité pour lutter contre ces infections virales. La compétence vectorielle représente un des principaux mécanismes étudiés (Agarwal et al., 2017). Par ailleurs, Tabachnick a démontré une hétérogénéité entre les individus et les populations d'une espèce vectorielle concernant des traits directement impliqués dans la compétence vectorielle (Tabachnick, 1994). Les variants génétiques d'arthropodes sont associés à des variations de compétence vectorielle (Tabachnick, 1994). En effet, cet auteur a montré que des éléments génétiques, incluant l'identification de loci, ou encore de marqueurs polymorphes, sont impliqués dans la compétence vectorielle d'un arthropode (Tabachnick, 1994). Les mécanismes moléculaires impliqués dans les interférences entre le génotype du vecteur insecte et la compétence vectorielle restent à élucider. De plus, le virus doit surmonter

les voies de l'immunité innée (développées dans la partie III et IV), aboutissant à l'infection de l'intestin moyen et des glandes salivaires (Franz et al., 2015). Nous allons maintenant nous intéresser aux premières voies de défenses anti-microbiennes de l'immunité innée et comprendre leur rôle lors d'une infection virale.

III. Lignes de défense antivirale chez les insectes

La première ligne de défense chez les insectes est non-spécifique du pathogène et représente les barrières physiques impliquant la cuticule. Les épithéliums intestinaux constituent également une barrière de protection face aux infections (Davis and Engström, 2012). Par ailleurs, les réponses de l'immunité innée représentent le pilier de la défense immunitaire chez les insectes. Le principal mécanisme de défense antivirale spécifique du virus est assuré par la voie des siARN chez les insectes. Cependant, d'autres voies antimicrobiennes innées moins spécifiques sont également importantes pour la défense antivirale, telles que les voies *Immune Deficiency* (Imd), Toll et *Janus kinase/signal transducer and activator of transcription* (Jak/STAT).

La réponse immunitaire est initiée dès la reconnaissance du pathogène par des récepteurs de type PRR (*pattern recognition receptors*) qui reconnaissent des motifs conservés de différents agents pathogènes, appelés PAMPs (*pathogen-associated molecular patterns*), à la surface du pathogène (Kawai and Akira, 2009). Le signal d'alerte est ensuite propagé et stimule les réponses effectrices de l'organisme. Le système interféron n'a pas été identifié chez les insectes. Ce sont des gènes antibactériens et antifongiques codant des peptides antimicrobiens (AMP, *antimicrobial peptides*) qui ont été décrits, ainsi que l'expression de gènes antiviraux. On distingue principalement la voie Jak/STAT, la voie Imd et la voie Toll.

1. La voie Jak/STAT

La voie Jak/STAT est un mécanisme largement conservé chez les insectes, caractérisé pour la première fois chez la drosophile. Cette voie a été décrite initialement pour son rôle dans le développement et la prolifération des hémocytes (Chao et al., 2004; Zettervall et al., 2004). L'importance conséquente de la voie Jak/STAT dans la réponse aux infections bactériennes et virales a été mise en évidence, incluant la production de molécules effectrices tels que les AMP (Agaisse and Perrimon, 2004). La voie de signalisation comprend le facteur de transcription STAT, une kinase Janus Hopscotch (hop), un récepteur Domeless (dome), et la protéine *Unpaired-related ligands* (upd3), pour ligands liés non appariés (Brown et al., 2001; Harrison et al., 1998; Hombría et al., 2005) (**Figure 9.**). Suite à la reconnaissance virale par la

cellule impliquant les interactions de type PAMP-PRR, la liaison de udp3 induit la dimérisation du récepteur dome, permettant la trans-phosphorylation des kinases hop. Le facteur de transcription est ensuite recruté et trans-phosphorylé au niveau du résidu tyrosine qui est conservé, permettant la dimérisation et le transport de STAT dans le noyau où il régulera l'expression des gènes effecteurs en aval. Parmi eux, on distingue notamment les membres de la famille tot (torandot), incluant les gènes *totA*, *totC* et *totM*, impliqués dans la réponse au stress (Agaisse and Perrimon, 2004). De plus, un autre ligand de la voie Jak/STAT a été identifié, correspondant à la protéine Vago qui suite à sa sécrétion active la voie Jak/STAT, lors de l'infection de cellules de *Culex quinquefasciatus* par WNV (Paradkar et al., 2012). Par ailleurs, l'expression de *vago* est stimulée lors de l'infection par SINV et DCV (Deddouche et al., 2008), deux virus étudiés au cours de mes travaux de thèse. Vago ne se lie pas au récepteur dome mais à un autre récepteur qui n'a pas été clairement décrit (Deddouche et al., 2008). Ces mêmes auteurs ont également démontré que l'expression du gène *vago* est dépendante de Dicer-2, excluant l'impact d'autres composants de la voie des siARN.

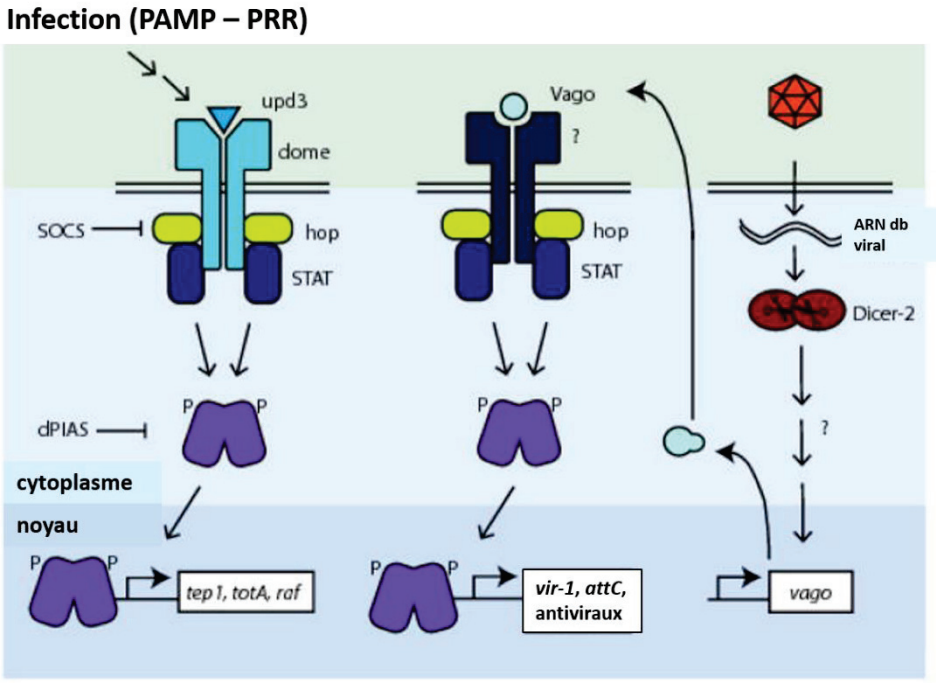


Figure 9. La voie Jak/STAT chez les insectes. Les ligands udp ou vago se lient à leur récepteur (dome ou inconnu, respectivement). Le récepteur se dimérise et induit la trans-phosphorylation des kinases hop. STAT est ensuite recruté et phosphorylé, permettant son transport dans le noyau et l'expression des gènes effecteurs, tels que *tep1*, *totA* ou *raf*. Les protéines SOCS et dPIAS ont été décrites comme inhibiteurs de hop et STAT, respectivement. Le gène antiviral *vago* est induit lors d'infection de virus à ARN, pouvant activer la voie Jak/STAT et permettre la transcription des gènes *vir-1*, *attC*, etc. D'après (Kingsolver et al., 2013).

2. Les voies immunitaires Imd et Toll

Les voies Imd et toll, relatives à la voie NF-κB, sont activées lors d'infections bactériennes et fongiques chez la drosophile (Lemaitre and Hoffmann, 2007). Relish est le facteur de transcription de la voie Imd (Leulier et al., 2002, 2000; Naitza et al., 2002; Silverman et al., 2000) (**Figure 10 A.**) alors que Cactus et Dif (*Dorsal-related immune factor*) sont impliqués dans la voie Toll (Arnot et al., 2010; Meng et al., 1999) (**Figure 10 B.**). Les récepteurs PGRP (*peptidoglycan recognition proteins*) activent la voie Toll, incluant l'activation du ligand Späztle puis sa liaison au récepteur Toll (**Figure 10 B.**). La cascade protéolytique aboutit à la libération de Dif, régulant la transcription des gènes codant les effecteurs de la voie Toll tels que les AMP (**Figure 10 B.**).

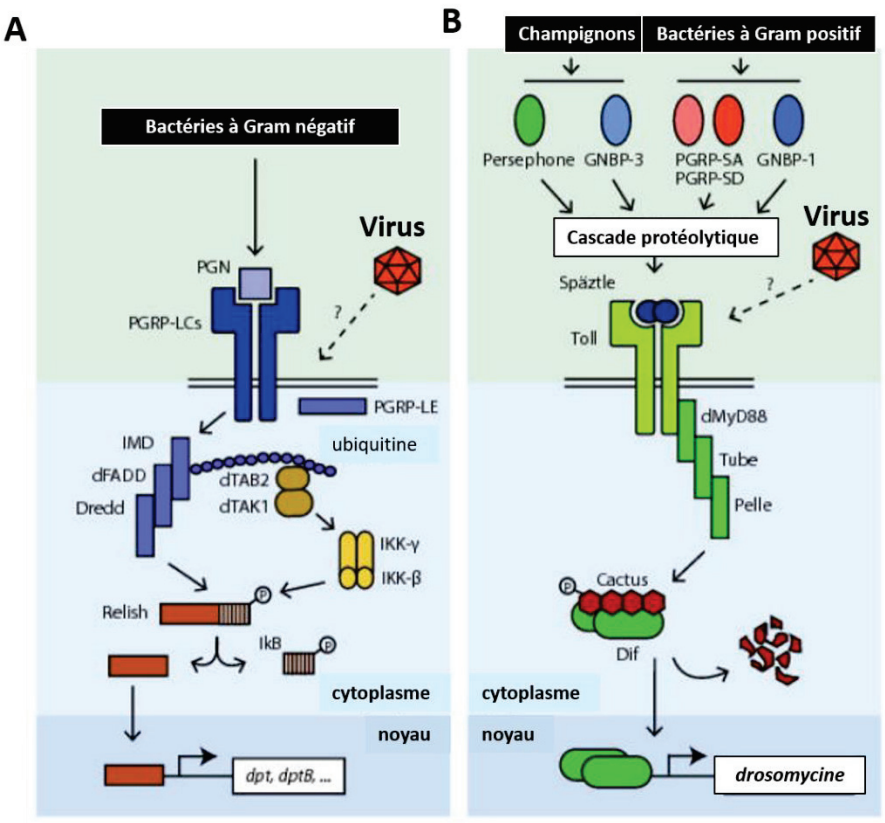


Figure 10. Voies immunitaires relatives à NF-κB chez les insectes. (A) La voie Imd. Le peptidoclycane (PGN) des bactéries Gram négatives et des *Bacillus* se fixe au récepteur transmembranaire PGRP-LCs et au récepteur intracellulaire PGRP-LE. Le signal est ensuite transmis à la molécule adaptatrice IMD qui recrute et active dFADD, puis Dredd, permettant la poly-ubiquitination d'IMD. L'assemblage et l'activation du complexe IKK implique la protéine TAK1 et induit la phosphorylation de Relish. Après clivage par la protéine Dredd, le domaine de liaison à l'ADN, correspondant au N-terminal de Relish, est transporté dans le noyau, régulant la transcription de différentes molécules effectrices, telles que *diptéricine* et *diptéricineB*. **(B)** La voie Toll. L'activation de la voie Toll se fait généralement lors d'infections bactériennes (Gram positifs) ou fongiques. Suite à la reconnaissance des

motifs des pathogènes par les récepteurs PGRP-SA, PGRP-SD et GNBP-1 etc, la cascade protéolytique est enclenchée et permet l'activation du ligand Spätzle. La liaison de Spätzle sur son récepteur Toll induit un signal pour l'activation de Cactus et Dif et implique les protéines MyD88, Tube et Pelle. La phosphorylation de Cactus entraîne sa dégradation et libère Dif qui est exporté dans le noyau et régule la transcription des gènes effecteurs, tel que celui codant la *drosomycine*. Les étapes d'interaction des virus dans ces voies de l'immunité ne sont pas décrites. dFADD, dTAB2, dTAK1 et dMyD88 sont les homologues des protéines de mammifère FADD, TAB2, TAK1 et MyD88 chez la drosophile (d). D'après (Kingsolver et al., 2013).

3. Les voies Jak/STAT, Imd et Toll lors d'infections virales

L'implication de la voie Jak/STAT et de la voie Imd dans le contrôle de la réplication virale est virus spécifique et peut également dépendre du mode d'administration du virus, oral ou par piqûre (Mondotte and Saleh, 2018). La voie Jak/STAT est notamment impliquée dans la réponse antivirale lors de l'infection par les virus DCV (Dostert et al., 2005) et CrPV (Kemp et al., 2013) chez la drosophile. Pour d'autres virus pathogènes, tels que FHV ou IIV-6, la voie Jak/STAT ne semble pas être impliquée dans les réponses antivirales.

En utilisant des drosophiles transgéniques qui codent le réplicon ARN de SINV, la diminution de l'expression de la protéine STAT ou la perte de fonction de la protéine Relish (impliquée dans la voie Imd) ont révélé une augmentation de la production du réplicon SINV permettant de souligner l'importance des voies Jak/STAT et Imd dans le contrôle de la réplication de SINV chez la drosophile (Avadhanula et al., 2009). Le réplicon SINV code les protéines non structurales du virus, permettant la réplication du génome viral de façon autonome mais au vu de l'absence de protéines structurales le réplicon ne permet pas la formation des particules virales infectieuses au sein de la drosophile transgénique. Ces mêmes auteurs ont montré l'importance de la voie Imd dans le contrôle du titre viral de SINV sauvage. L'invalidation (KO, knockout) des gènes *attC* et *dptB*, codant les AMP, Attacin-C et Diptericin B, respectivement, augmente la réplication de SINV chez la drosophile (Z. Huang et al., 2013). Les protéines AttC et DptB sont régulées respectivement par la voie Jak/STAT et la voie Imd et sont surexprimées lors de l'infection par SINV (Z. Huang et al., 2013). De plus, lors de l'infection par le virus sigma, une augmentation des niveaux de transcription a été observée pour les composants de la voie Imd, tels que PGRP-SB1 et PGRP-SD, et en aval, au niveau des AMP (Tsai et al., 2008). Cependant, ces mêmes observations ne sont pas retrouvées en utilisant une autre lignée de drosophile (Carpenter et al., 2009). L'importance de la voie Toll a notamment été détectée lors de l'infection par DXV chez *Drosophila* (Zambon et al., 2005) et lors de l'infection par DENV chez le moustique (Ramirez and Dimopoulos, 2010). L'inactivation de la voie Toll provoque rapidement la mort des drosophiles mutées pour *Dif*, par un processus d'anoxie, et augmente les titres viraux du DXV par rapport aux drosophiles sauvages (Zambon et al., 2005). Bien que l'ensemble de ces voies permette la tolérance virale en diminuant la charge virale, elles doivent être régulées pour éviter une sur-activation, néfaste pour l'organisme hôte (Zambon et al., 2005). En effet, la voie Jak/STAT est régulée par l'histone

méthyltransférase G9a, conduisant à la tolérance virale lors de l'infection par DCV (Merkling et al., 2015). Ces auteurs ont identifié une hyperactivation de la voie Jak/STAT pour les mutants déficients pour la protéine G9a, associée à une mortalité des drosophiles infectées expérimentalement. Lors de la dérégulation des autres voies de défense antivirale, telles que la voie Imd, la tolérance de la drosophile est également impactée. La protéine Diedel est notamment impliquée dans la tolérance virale chez la drosophile lors de l'infection par SINV (Lamiable et al., 2016b). Le gène *diedel* (*die*) codant cette protéine est induit après infection expérimentale par injection de drosophiles avec SINV et permet de réguler l'activation de la voie Imd et de la signalisation NF-κB, empêchant une sur-activation de ces voies immunitaires, délétères pour l'organisme de la drosophile. L'ensemble de ces élements met en évidence le rôle crucial de la régulation de ces voies de défense antivirale dans les mécanismes de tolérance virale. Ajouté à ces voies de défense antivirale, le mécanisme d'ARNi (détaillé dans la partie V) joue un rôle primordial dans la protection de l'organisme hôte.

Des virus aux éléments transposables

L'ensemble des génomes des insectes contient des éléments mobiles de l'ADN, appelés éléments transposables (ET), et qui sont considérés comme des parasites du génome. Certains ET sont des vestiges d'infections rétrovirales, qui ont surmonté les barrières antivirales et atteint les cellules germinales. Le cycle de réplication des rétrovirus classiques inclut la transcription inverse du génome viral à ARN en ADN viral qui sera ensuite incorporé dans le génome de la cellule hôte. Selon certaines conditions, incluant le passage dans la lignée germinale, ces infections virales ont été transmises au génome du zygote, aboutissant à l'infection de l'ensemble des cellules de l'organisme. La dérive génétique peut ensuite entraîner la fixation de ces insertions rétrovirales, qui deviennent des insertions de rétrovirus endogènes (ERV), partagées par tous les individus d'une espèce (Dewannieux and Heidmann, 2013). Ce phénomène n'est pas exclusif aux rétrovirus, mais dans une moindre mesure un large panel de virus eucaryotes peut être endogénisé dans le génome d'une grande variété d'espèces hôtes, telles que des vertébrés ou des arthropodes (Feschotte and Gilbert, 2012). Une relation évolutive de longue durée entre un virus à ARN et son hôte pourrait aboutir à l'endogénisation de ces séquences virales. En plus de cette origine virale commune entre les ERV et les virus, il existe des similitudes au niveau de leur voie de contrôle, incluant la machinerie de l'ARNi, primordiale à leurs régulations respectives. Chez les insectes, les ET sont contrôlés par la voie des piARN (*Piwi-interacting RNA*) dans les gonades, tandis que la voie des siARN (*small interfering RNA*) contrôle les ET dans le soma et a un rôle primordial dans le contrôle systémique des infections virales. L'ensemble des ces informations suggère une interaction entre les virus et les ET, qui pourrait impliquer ces voies d'ARNi. La caractérisation de ces interactions potentielles est d'une grande importance pour l'étude des interactions entre les insectes et les virus.

IV. Les éléments transposables

1. Introduction générale

Les éléments transposables (ET) ont été découverts pour la première fois dans le maïs par Barbara McClintock dans les années 1950 (McClintock, 1950). Ses travaux lui ont valu le prix Nobel en 1983. Les ET ont d'abord été considérés comme de l'ADN poubelle ou des parasites génomiques, se multipliant dans le génome aux dépens de celui-ci. Ils sont aujourd'hui reconnus pour être des acteurs majeurs du génome, pouvant avoir un rôle dans le contrôle et le remodelage des gènes, et sont impliqués dans la fonctionnalité et l'évolution des génomes (Biémont, 2010).

Les ET sont ubiquitaires et largement représentés au sein des génomes eucaryotes et procaryotes, en persistant par l'intermédiaire d'une réplication de leurs séquences. Ils sont aussi appelés « éléments génétiques mobiles » car ce sont des séquences d'ADN répétées pouvant transposer (c'est-à-dire se mobiliser) et se multiplier tout au long des chromosomes, par l'intermédiaire de la machinerie enzymatique qu'ils codent. On distingue des séquences composées d'éléments actifs, pouvant transposer, et des séquences dégénérées, immobiles. Le nombre de copies de ces ET peut être multiplié par l'intermédiaire de ces événements de transposition, induisant une source potentielle d'instabilité génétique.

Les mécanismes de transposition, ainsi que la structure générale des ET permettent de les diviser en deux grandes classes. La première classe correspond aux rétrotransposons, incluant les rétrotransposons à LTR (*Long Terminal Repeats*) et sans LTR. Parmi les rétrotransposons à LTR, on retrouve les rétrovirus endogènes (ERV) codant une glycoprotéine d'enveloppe qui leur permet d'avoir un pouvoir infectieux. Les rétrotransposons sans LTR incluent les LINEs, pour *Long Interspersed Nuclear Elements* qui se déplacent de manière autonome dans le génome à l'inverse des SINEs, pour *Short Interspersed Elements* qui sont des rétrotransposons non autonomes et très courts (inférieur à 500 pb). L'ensemble de ces rétrotransposons transposent en utilisant un mécanisme de « copier-coller » impliquant un intermédiaire ARN à une étape du cycle de réplication. La deuxième classe correspond aux transposons à ADN qui utilisent un mécanisme de « couper-coller » pour transposer. Au sein de ces deux grandes classes, les ET partageant des similarités de structure et de séquence vont constituer les familles d'ET. La caractérisation de leur structure inclut notamment la présence ou l'absence de domaine ou de gène particulier, l'ordre des domaines et la signature des acides aminés. Un dernier niveau de classification permet de distinguer plusieurs superfamilles, en fonction des motifs et des signatures des acides aminés.

2. Rétrotransposons, classe I

Les rétrotransposons transposent par la transcription inverse d'un intermédiaire ARN et l'insertion de sa copie ADN au nouveau site d'intégration génomique. Ils utilisent un mécanisme de transposition de type « copier-coller » qui est par conséquent réplicatif, générant ainsi un nombre exponentiel de copies (**Figure 11.**).

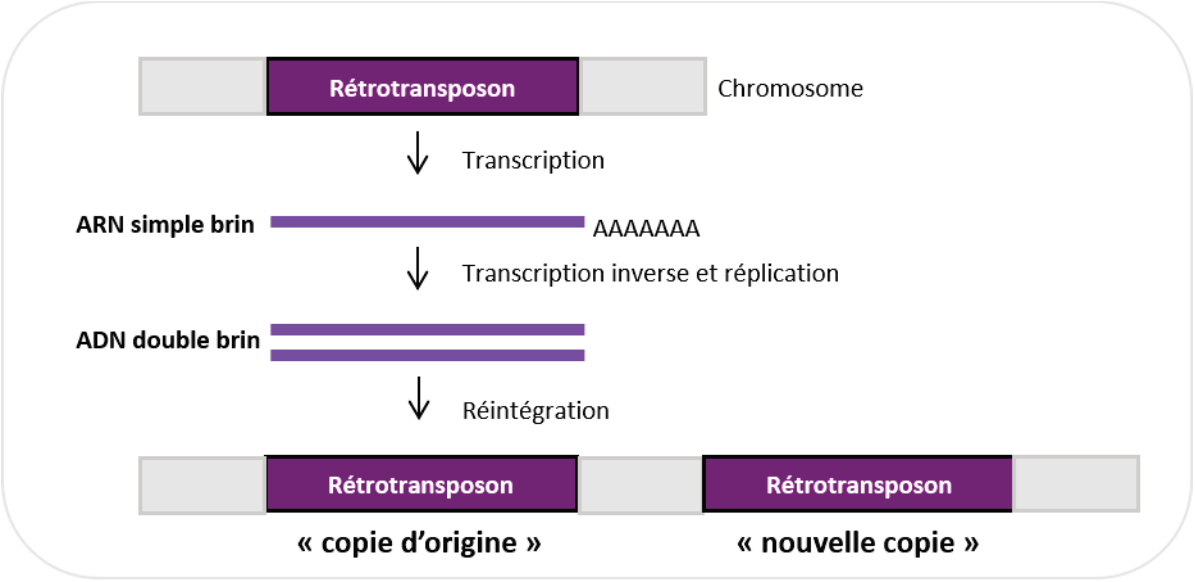


Figure 11. Schéma de la transposition d'un rétrotransposon impliquant un mécanisme de type « copier-coller ». Le cycle du rétrotransposon débute par la transcription du rétrotransposon en ARN simple brin, puis ce dernier est rétro-transcrit en ADN double brin qui va se réintégrer dans un nouveau site génomique. Le premier rétrotransposon reste intact.

A. Rétrotransposons à LTR

Les rétrotransposons à LTR sont les ET les plus abondants chez la drosophile et constituent la plus grande proportion de l'euchromatine (2,65%) par rapport aux autres classes d'éléments transposables (Kaminker et al., 2002). Ils incluent de nombreuses familles d'ET, dont la plus abondante, nommée *roo*, contient 146 copies (Kaminker et al., 2002). Ils sont également présents en grande quantité dans l'hétérochromatine des chromosomes de drosophiles. Plusieurs études montrent que ces copies hétérochromatiques ont été intégrées chez *D. melanogaster* il y a au moins 100 000 ans, avant l'intégration des rétrotransposons sans LTR (Bergman and Bensasson, 2007; Mugnier et al., 2008).

a) Classification

On distingue trois super-familles de rétrotransposons à LTR : *Ty1/Copia*, *BEL/Pao* et *Ty3/Gypsy*, comprenant huit clades et 35 familles (de la Chaux and Wagner, 2011; Mugnier et al., 2008). Le groupe Gypsy est le plus représenté, avec 27 familles, réparties en cinq sous-groupes : *gypsy*, *ZAM*, *Idefix*, *412* et *blastopia*. Les groupes Copia et BEL/Pao présentent seulement quatre et cinq familles respectivement. Le groupe Copia ne possède pas le même arrangement de cadre de lecture du gène *pol* par rapport aux deux autres groupes (Bowen and McDonald, 2001). Les mécanismes de transposition peuvent varier légèrement entre ces groupes mais ils contiennent tous des terminaisons LTR, tout comme les rétrovirus.

b) Structure

Les rétrotransposons à LTR présentent des similarités de structure avec les rétrovirus des vertébrés et ainsi, ils mettent en jeu un mécanisme de réplication reproduisant un cycle viral, infectieux ou non (Capy et al., 1997). Ils sont bordés par des répétions longues de type LTR, qui incluent des séquences régulatrices. Ils se composent de deux ou trois cadres de lecture ouverts : *gag* code les protéines de la capside et *pol* code les enzymes indispensables au cycle de réplication du rétroélément (**Figure 12.**). Ces enzymes sont la transcriptase inverse (RT) incluant un domaine RNAse-H et permettant la synthèse d'ADN complémentaire (ADNc), l'intégrase (IN) permettant l'intégration de l'ADN proviral au niveau d'un nouveau site chromosomique dans le génome, et la protéase virale nécessaire au clivage de la polyprotéine Gag-Pol. Les ERV possèdent un génome similaire à celui des rétrotransposons à LTR excepté l'ajout d'un troisième cadre de lecture ouvert, *env* qui code la protéine d'enveloppe (Env) (**Figure 12.**). Ces derniers peuvent infecter d'autres cellules, par l'intermédiaire de la protéine Env, lors de leur cycle de transposition, ou rester dans la même cellule (**Figure 13.**).

c) Cycle de transposition

Le cycle de transposition d'un rétrotransposon à LTR (**Figure 13.**) commence par la transcription de l'élément à partir de son domaine LTR, permettant la synthèse d'un ARN viral polyadénylé. Ce dernier est ensuite exporté dans le cytoplasme de la cellule et est traduit en protéines virales (Gag, Pol). L'encapsidation au sein de la capside de l'ARN génomique viral ainsi que des enzymes issues de Pol (RT, RH et INT) implique la séquence PSI (ψ). Cela permet la formation de la VLP (*virus like particle*), suivie de la transcription inverse de l'ARN génomique. L'ADN double brin est ré-intégré dans le génome par l'intermédiaire de l'intégrase et pourra recommencer un cycle de transposition réplicatif.

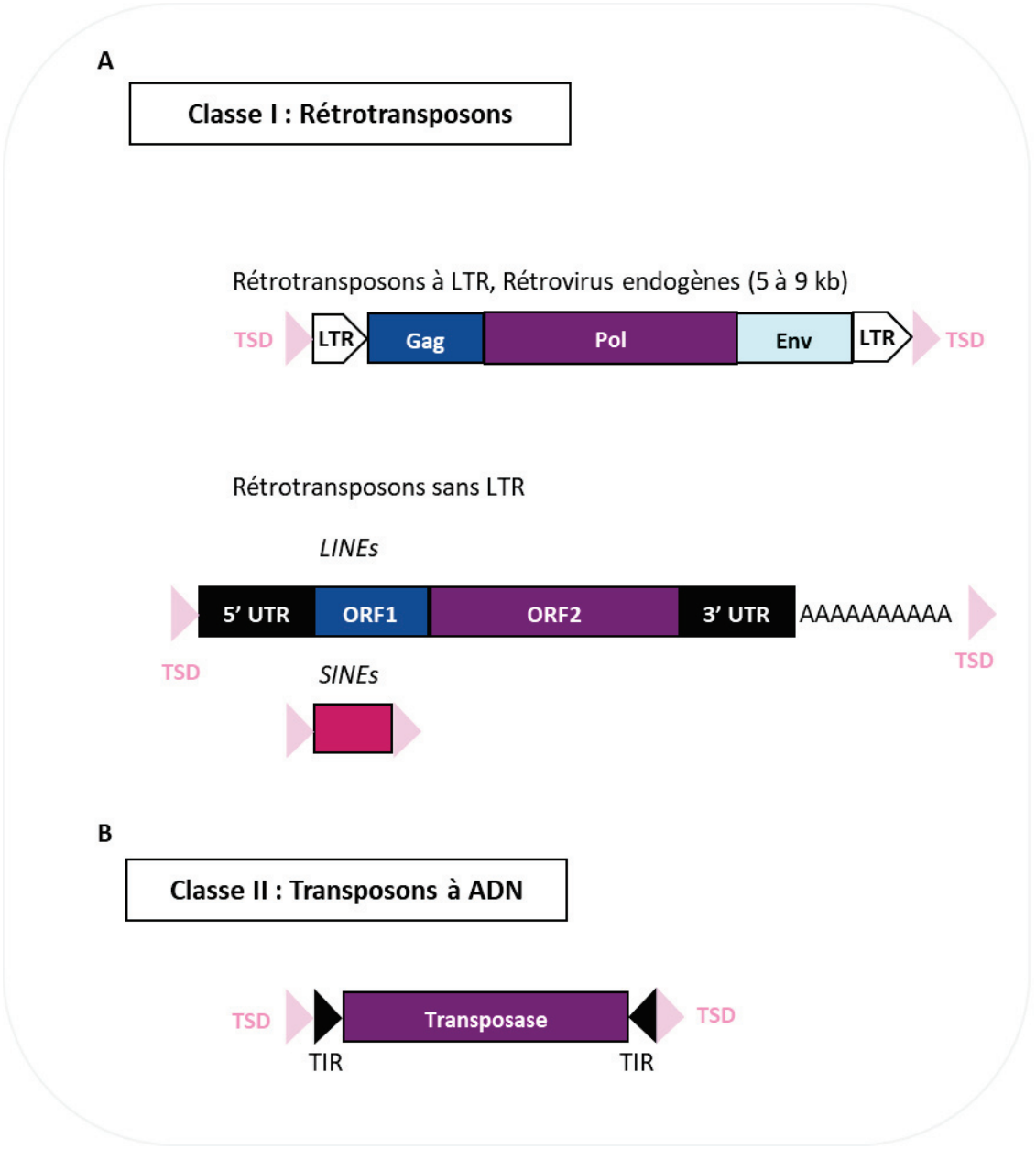


Figure 12. Schéma de la structure des différentes classes et sous-classes des ET. (A) Rétrotransposons à LTR (classe I). Ils contiennent deux cadres ouverts de lecture (*open reading frame*, ORF) : *gag* (codant les protéines de structure pour la formation de la capside), *pol* (codant les protéines de réplication et de rétrotransposition). Les rétrovirus endogènes possèdent un troisième cadre ouvert de lecture qui code la protéine Env (protéine d'enveloppe). Ils sont encadrés à chaque extrémité par deux longues répétitions terminales (LTR). **Rétrotransposons sans LTR (classe I).** Les longs éléments nucléaires intercalés (LINE) sont composés d'un premier cadre ouvert de lecture codant une protéine de structure Gag et un second codant des protéines impliquées dans la réplication. Ils sont flanqués par des extrémités UTR (*untranslated region*) non traduites ainsi que des séquences TSD, pour *target site duplications* (triangles roses), et se terminent par une queue poly(A). Les petits éléments nucléaires intercalés (SINE) sont des séquences non codantes, dérivées de transcrits d'ARN de transfert ou d'ARN impliqués dans le splicéosome et ils sont pris en charge par l'ARN polymérase III. (LINEs : *Long interspersed nuclear element*. SINEs : *Short interspersed nuclear elements*.). **(B) Transposons à ADN (classe II).** Ils contiennent une ORF qui code une transposase délimitée par deux répétitions inversées terminales (TIR).

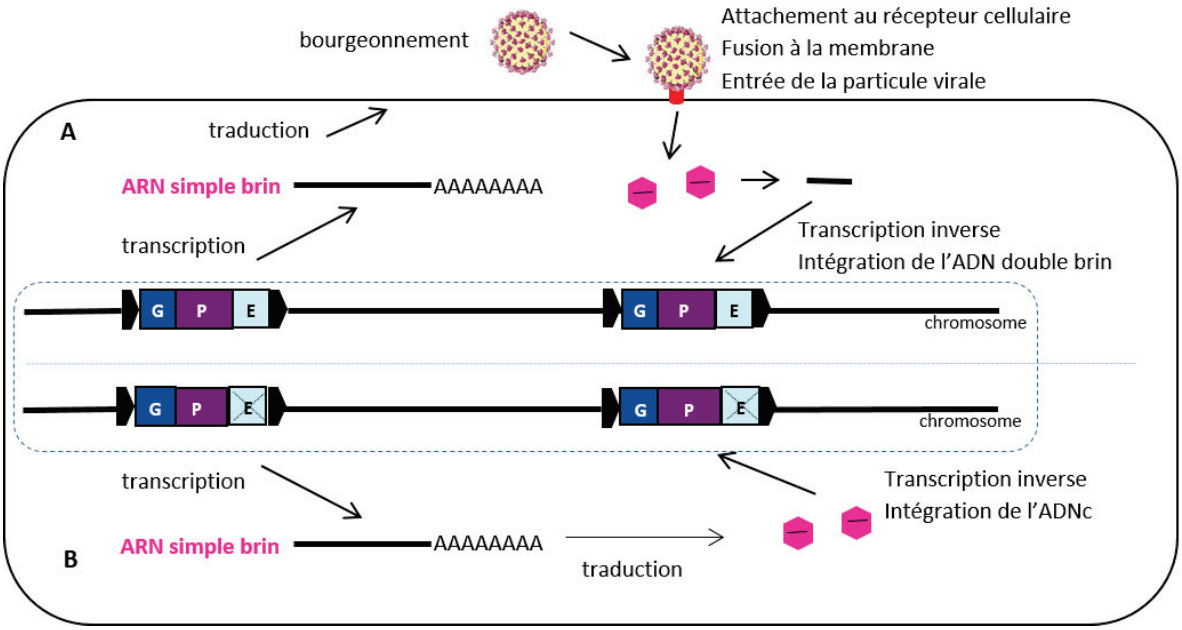


Figure 13. Schéma de la transposition d'un rétrotransposon à LTR. (A) Cycle de transposition réplicatif d'un rétrovirus endogène. L'amplification du ERV peut impliquer une étape extracellulaire. Les ERV sont rétro-transcrits en ARN simple brin puis traduits en protéines virales qui vont s'assembler au niveau de la membrane plasmique cellulaire pour former des particules virales infectieuses après bourgeonnement. Ces dernières peuvent infecter d'autres cellules lors de la reconnaissance du récepteur cellulaire (représenté en rouge) par la glycoprotéine d'enveloppe. La capside virale et ensuite l'ARN viral sont libérés dans le cytoplasme où se produit la transcription inverse. L'ADN double brin (provirus) peut se ré-intégrer au sein du génome. **(B) Cycle de transposition réplicatif d'un rétrotransposon à LTR.** Ces rétrotransposons à LTR ont un cycle réplicatif strictement intracellulaire. Ils peuvent présenter notamment des mutations au niveau des protéines Gag ou Env, empêchant l'expression ou la fonctionnalité de la protéine Env. Ces rétrotransposons à LTR sont également transcrits et traduits en protéines virales conduisant à la formation de particules intracellulaires non infectieuses. La transcription inverse a lieu à l'intérieur de ces particules, permettant la formation d'ADNc ou provirus qui peut après réplication se ré-insérer dans le génome de la même cellule. D'après (Dewannieux and Heidmann, 2013).

d) Le cas des rétrovirus endogènes

Parmi les rétrotransposons à LTR, on compte les rétrovirus endogènes (ERV), qui sont des séquences rétrovirales largement dispersées et intégrées dans le génome de l'hôte. Les ERV présentent la même structure que les rétrotransposons à LTR, avec un gène d'enveloppe. Selon l'ICTV, les ERV chez les insectes sont classés en fonction de leur similarité de structure, dans la famille des *Metaviridae*, tandis que pour les vertébrés, ils sont classés dans la famille des *Retroviridae*. Chez les insectes, les *Metaviridae* présentent un troisième cadre de lecture ouvert (*open reading frame*, ORF) codant la protéine Env, permettant de les classer dans le groupe monophylétique des Errantivirus.

Par ailleurs, les *Retroviridae* sont spécifiques des vertébrés et contiennent un troisième ORF, codant une enveloppe qui leur donne leur pouvoir infectieux. Par conséquent, ils sont verticalement transmis, suivant les lois de Mendel, des parents à la descendance sous forme

intégrée dans le génome (ADN proviral). Ils proviennent généralement d'anciennes infections de rétrovirus exogènes ayant franchi les barrières antivirales de l'hôte afin d'infecter les cellules germinales puis l'ensemble de l'organisme (Belshaw et al., 2004; Dewannieux and Heidmann, 2013). Ainsi, certaines séquences de rétrovirus peuvent se fixer et être partagées par tous les individus d'une espèce, on parle d'endogénisation. Elles représentent 8% du génome humain (Lander et al., 2001), principalement sous forme de copies inactives ou dégénérées.

Chez les insectes, les ERV, ont été décrits depuis de nombreuses années et incluent *gypsy*, *ZAM*, *Idefix*, *tirant*, *17.6*, *nomad*, *297*, *Tom*, et *rover* chez *D. melanogaster* (Vidal et al., 2009). Chez *D. melanogaster*, l'euchromatine contient environ 2% d'ERV (Kaminker et al., 2002). Contrairement aux ERV chez les vertébrés, le mécanisme d'acquisition du gène *env* ne semble pas provenir d'une ancienne infection rétrovirale. Plusieurs travaux apportent des précisions, avec notamment la détection du gène *Id130* codant une glycoprotéine d'enveloppe issue d'ERV du groupe Gypsy, et possédant des similarités avec la glycoprotéine d'enveloppe F des Baculovirus chez la drosophile (Malik et al., 2000; Pearson and Rohrmann, 2002; Terzian et al., 2001). Ces derniers sont des virus à ADN double brin infectant les invertébrés, notamment les arthropodes. L'acquisition du gène *env* chez les insectes impliquerait une recombinaison entre certains rétrotransposons à LTR et le gène *env* des Baculovirus.

B. Rétrotransposons sans LTR

a) LINEs

Les LINEs (*Long Interspersed Nuclear Element*) sont des rétrotransposons sans LTR, transposant de manière autonome dans le génome puisqu'ils codent leur propre transcriptase inverse. Ils possèdent deux cadres de lecture ouverts, partageant des similarités avec les gènes *gag* et *pol* et se terminent par une queue poly(A) (Koito and Ikeda, 2013) (**Figure 12.**). Ils présentent un cycle de transposition réplicatif intracellulaire (**Figure 14.**) nettement différent de celui des rétrotransposons à LTR.

Plus précisément, ils contiennent un promoteur interne d'ARN polymérase II dans la région 5' UTR non-traduite (UTR, pour *UnTranslated Region*) suivie par deux cadres de lecture ouverts. Le premier (ORF1) code une protéine de liaison à l'ARN nécessaire à la formation des particules de ribonucléoprotéine (RNP) dans le cytoplasme. Le second (ORF2) code une protéine ayant des activités d'endonucléase et de reverse-transcriptase. Enfin, une région courte 3' UTR est suivie d'une queue polyadénylée et l'élément est flanqué par des séquences TSD, pour *target site duplications*.

La transcription est initiée par le promoteur interne de ces éléments au niveau de la région 5'UTR. L'ARNm est ensuite traduit et les protéines résultantes s'associent à l'ARN du rétro-

élément, permettant la formation du complexe ribonucléoprotéique dans le cytoplasme. Ces intermédiaires de rétrotransposition sont nécessaires au mécanisme de transposition « copier-coller ». Les copies d'ARN de ces éléments sont rétro-transcrits dans le noyau des cellules. L'extrémité polyadénylée de l'ARN s'apparie à l'ADN chromosomique au niveau d'une région riche en A/T. Les activités d'endonucléase et de reverse transcriptase de la protéine codée par l'ORF2 (ORF2p) permet la cassure simple brin au niveau de l'ADN chromosomique, ce qui permet de réaliser une transcription inverse ciblée, appelé TPRT, pour *target-primed reverse transcription*. L'ARN sert de matrice pour la synthèse d'ADN complémentaire en utilisant le groupement hydroxyle 3' (3'-OH) de l'ADN comme amorce. L'endonucléase induit une cassure au niveau du second brin d'ADN pour la synthèse de l'autre brin d'ADN, en utilisant le deuxième groupement 3'-OH. L'intégration se termine par la réparation de l'ADN, ce qui génère des séquences TSD dupliquées autour de l'élément.

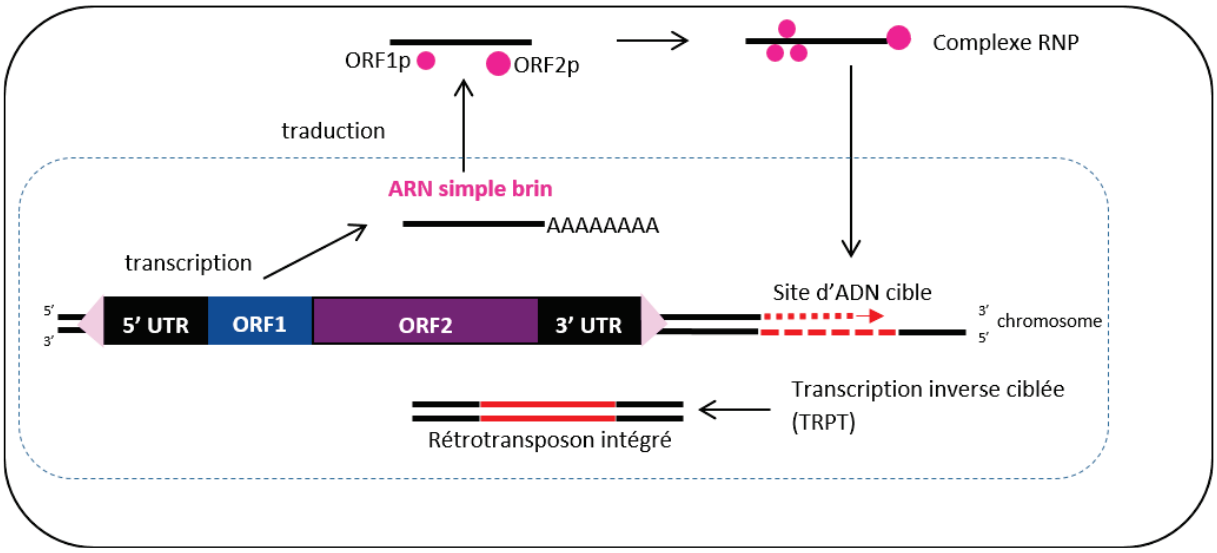


Figure 14. Schéma de la transposition d'un rétrotransposon sans LTR. Un promoteur interne dans la région 5' de l'élément initie la transcription. Les deux cadres de lecture ouverts codent les protéines nécessaires à la rétrotransposition en *cis*. Ces protéines s'associent préférentiellement avec leurs propres ARN codants et forment un complexe ribonuléoprotéique dans le cytoplasme, essentiel au mécanisme de rétrotransposition. Dans le noyau, la synthèse et l'intégration de l'ADN dans le génome reposent sur la transcription inverse ciblée du transcrit (*target-primed reverse transcription*, TRPT) au cours de laquelle l'ARN sert de matrice. D'après (Koito and Ikeda, 2013).

b) SINEs

Les SINEs (*Short Interspersed Nuclear Element*) sont des rétroéléments non-autonomes, souvent inférieurs à 500 pb. Ce sont des séquences non-codantes, issues de transpositions accidentelles de transcrits de polymérase III, tels que les ARN de transfert (Wicker et al., 2007). Ils contiennent également une séquence similaire à celle des LINEs leur permettant d'être pris

en charge dans leur cycle de transposition (**Figure 12.**). Ainsi, ne codant pas leur propre machinerie de transposition, ils vont utiliser celle des LINEs pour se mobiliser. On retrouve par exemple l'élément Alu de *Homo sapiens* dans cette sous-classe (Dewannieux and Heidmann, 2013).

3. Transposons à ADN, classe II

La deuxième classe correspond aux transposons à ADN qui utilisent un mécanisme de « couper-coller » pour transposer, sans intermédiaire ARN. Ils codent une enzyme appelée transposase qui est encadrée par des répétitions inversées TIR, pour *Terminal Inverted Repeats* (**Figure 12.**). Lors de sa transposition, le transposon à ADN est excisé depuis son site d'origine puis intégré dans un site cible du génome. Ce mécanisme implique la reconnaissance des TIR par la transposase qui permet la formation d'un complexe, *Single End Complex*. Les extrémités du transposon se rejoignent par l'intermédiaire des deux monomères de transposase qui clivent l'extrémité 5' des séquences TIR par hydrolyse de la liaison phosphodiester et forment un dimère permettant la génération du deuxième complexe, *Paired-End Complex*. Ainsi, cela permet l'excision du corps du transposon à ADN. Ensuite, le dimère de transposase reconnaît et s'attache à un dinucléotide TA et forme le troisième complexe, *Target Capture Complex*, correspondant à l'insertion du transposon au nouveau site génomique (Muñoz-López and García-Pérez, 2010). Lors de l'insertion, le site cible TA est dupliqué et forme les TSD à chaque extrémité du transposon. Ainsi, la majorité des transposons à ADN transposent sans mécanisme de réplication (**Figure 15.**).

Selon les séquences TIR et TSD, on observe deux sous-classes de transposons à ADN. La première sous-classe inclut *Tc1/mariner*, *PIF/Harbinger*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *piggyBac* et *CACTA*. Chez *D. melanogaster*, les principaux incluent la famille *P*, constituée notamment de l'élément *P* (Majumdar and Rio, 2015; Robertson and Engels, 1989), la famille *Hat* incluant *hobo* (Rubin et al., 2001) et la famille Tc1/mariner avec l'élément *mariner* (Maruyama and Hartl, 1991). Plus récemment découverts, les transposons à ADN pouvant se répliquer appartiennent à la deuxième sous-classe comme *Helitron* et *Maverick*. La transposition des éléments du groupe des Helitrons implique un mécanisme de réplication circulaire de l'ADN (rolling-circle replication) (Kapitonov and Jurka, 2001) tandis que le mécanisme de transposition réplicative des Mavericks n'est pas clairement décrit mais impliquerait leur propre ADN polymérase (Pritham et al., 2007). Par ailleurs, de la même façon que les SINEs, certains transposons à ADN ne sont pas autonomes et vont utiliser la machinerie d'un transposon à ADN codant la transposase pour se mobiliser. C'est notamment le cas pour *Miniature Inverted-repeat Transposable Elements* (MITEs), un élément de type transposon à ADN, souvent retrouvé chez les plantes (Feschotte et al., 2002).

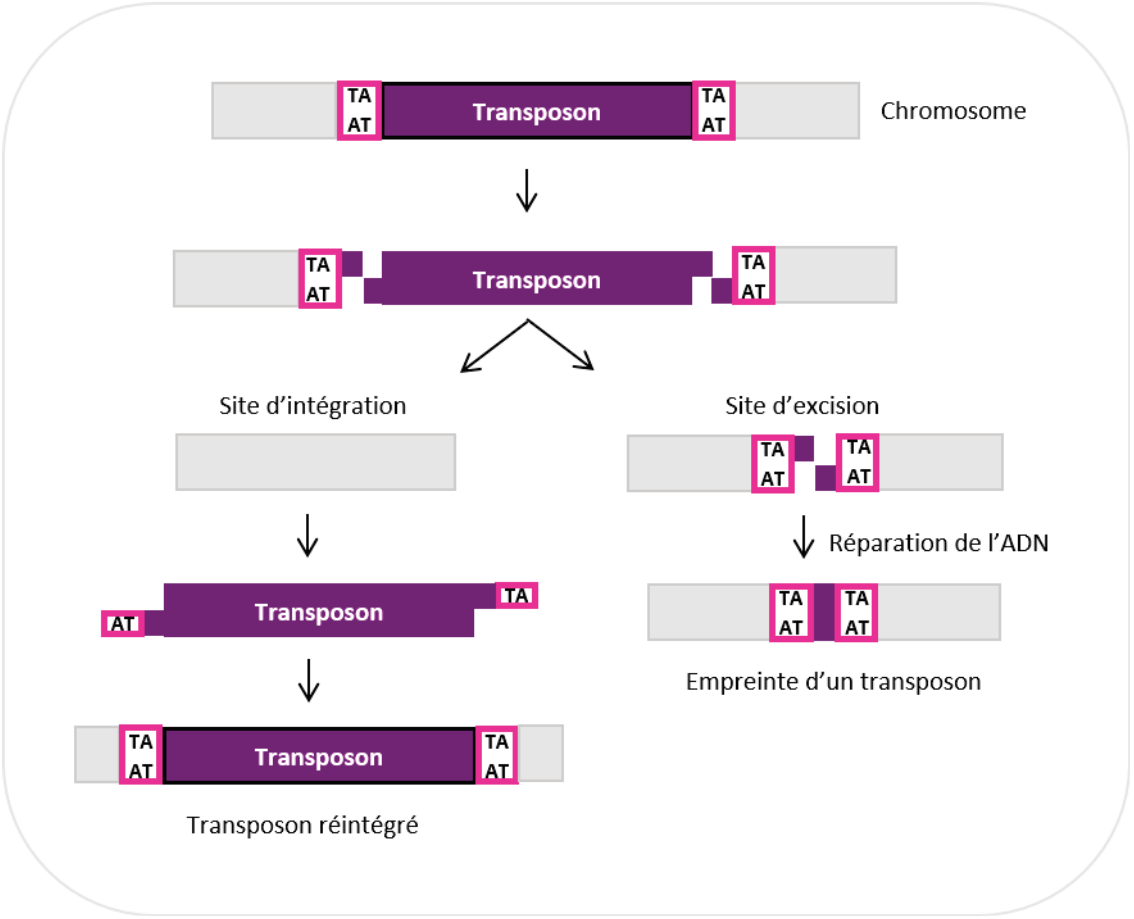


Figure 15. Schéma de la transposition d'un transposon à ADN impliquant un mécanisme de type « couper-coller ». Le transposon est excisé à partir de son site d'origine sur le génome et réintégré au niveau d'un site cible incluant un dinucléotide TA. Lors de la réintégration, le dinucléotide TA est dupliqué et forme le TSD, pour *Target Site Duplication*. Au niveau du site excisé, l'hôte met en place un mécanisme de réparation de l'ADN en formant des jonctions d'extrémités non homologues (NHEJ, *nonhomologous end-joining*). D'après (Muñoz-López and García-Pérez, 2010).

4. Paysage des ET selon l'hôte

Les organismes eucaryotes présentent des abondances de séquences d'ET très variables (**Figure 16.**). La structure des ET, leur système de transposition ainsi que le nombre et la localisation des copies lors de la transposition sont des facteurs pouvant être impliqués dans cette diversité. On retrouve des variations entre espèces et au sein d'une même espèce (Bosco et al., 2007; Vieira et al., 1999). C'est notamment le cas si on compare *D. melanogaster* et *D. simulans*. Les ET représentent environ 20 % du génome chez *D. melanogaster* et sont responsables d'une grande partie des mutations spontanées observées (Guerreiro, 2012; Harada et al., 1990). Chez *D. simulans*, les ET représentent seulement 5 % du génome (Dowsett and Young, 1982; Vieira and Biémont, 2004). Chez l'Homme, ils représentent environ 45 % du génome mais incluent beaucoup de séquences tronquées ou dégénérées, ce qui induirait

seulement 1 à 2 % des mutations (Huang et al., 2012; Mills et al., 2007). La proportion d'ET est très importante chez les moustiques, notamment au niveau du génome séquencé d'*Anopheles gambiae*, le principal vecteur du paludisme, où ils représentent 16 % (Holt et al., 2002) et encore davantage dans le génome séquencé *d'Aedes aegypti* qui en contient 47% (Nene et al., 2007), incluant 10,5% de rétrotransposons à LTR (Matthews et al., 2018), et atteignant 59% d'ET dans le génome séquencé d'*Aedes albopictus* (Chen et al., 2015). Au niveau du génome de *Culex quinquefasciatus,* les ET représentent 29 % du génome (Arensburger et al., 2010), incluant seulement 1 % de rétrotransposons à LTR (Marsano et al., 2012). Par ailleurs, plusieurs auteurs ont identifié l'influence de la proportion d'ET sur la taille des génomes (Biémont, 2008; Biémont and Vieira, 2006; Boulesteix et al., 2006; Kidwell, 2002; Sessegolo et al., 2016).

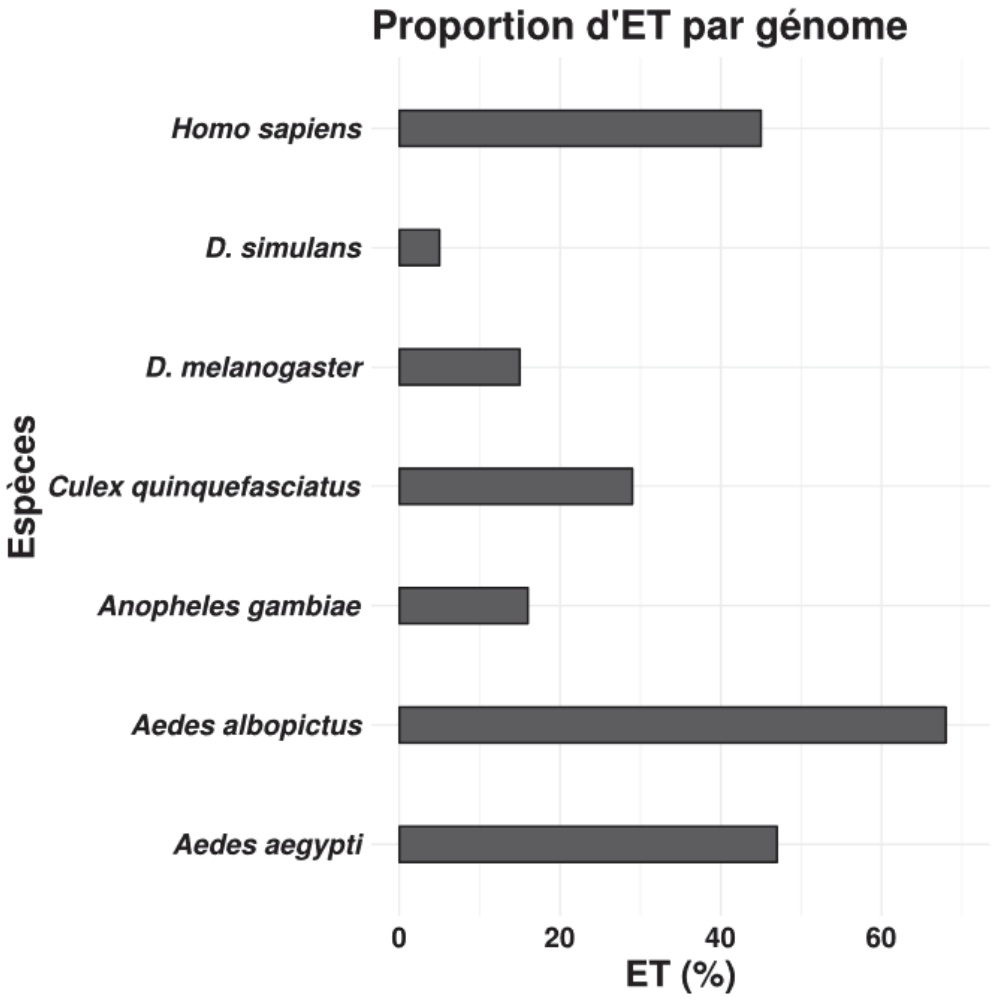


Figure 16. Proportion des ET dans différents génomes eucaryotes séquencés. *Homo sapiens* (Lander et al., 2001), *D. melanogaster* et *D. simulans* (Drosophila 12 Genomes Consortium et al., 2007), *Culex quinquefasciatus* (Arensburger et al., 2010)**,** *Anopheles gambiae* (Holt et al., 2002), *Aedes albopictus* (Chen et al., 2015), *Aedes aegypti* (Nene et al., 2007).

5. Impacts de la transposition des ET

On parle souvent de « parasites du génome » car on considère que les ET se multiplient au sein des génomes à leurs dépens. Il existe une grande variabilité d'insertion d'ET, ayant un impact sur l'expression des gènes de l'hôte. La transposition des ET peut conduire à des insertions dans une région de régulation des gènes ou directement dans un gène, pouvant entraîner une perte ou un gain de fonction de ce gène. La présence de ces ET dans le génome des organismes peut également avoir un impact sur la structure de la chromatine au niveau des régions où ils sont intégrés, car ils sont la cible privilégiée de marques hétérochromatiques qui peuvent diffuser au voisinage.

Leur activité a été démontrée comme pouvant être néfaste pour les génomes, notamment par l'intermédiaire de mutations et réversions spontanées (Eickbush and Furano, 2002; Kuzin et al., 1994). Chez l'Homme, la transposition des ET est responsable de mutations délétères et est aussi décrite pour son implication au niveau des premiers stades de maladies humaines, telles que le cancer (Hancks and Kazazian, 2016). Cependant, on comprend maintenant davantage leur rôle dans l'évolution de la structure des chromosomes et du contenu en gènes. Ces insertions, loin d'être toujours délétères pour le génome, peuvent conduire à des mutations permettant une meilleure adaptation chez l'hôte. Comme n'importe quelle mutation, la transposition des ET constitue également une source pour l'évolution des génomes. Une étude montre que l'insertion d'un rétrotransposon LTR augmente la viabilité de *D. melanogaster* dans des conditions normales ou de stress (Merenciano et al., 2016). Une augmentation globale de la quantité d'ET a été observée dans des lignées dérivées de *D. melanogaster* par rapport aux souches ancestrales, permettant probablement l'adaptation de ces drosophiles aux nouveaux environnements (Vieira et al., 1999). L'origine de l'espèce étant localisée en Afrique, les populations d'Afrique correspondent à l'aire ancestrale tandis que les populations ailleurs dans le monde correspondent à l'aire dérivée. De plus, les routes de colonisation du monde de *D. melanogaster* sont les mêmes que chez *Homo sapiens*.

Les insertions d'ET sont reconnues pour être une source de séquences régulatrices des génomes eucaryotes. Ils sont notamment impliqués dans la régulation de processus biologiques importants, tels que le développement et la physiologie (Chuong et al., 2017; Gerdes et al., 2016). Les insertions en amont des régions codantes du génome et directement au niveau de gènes conservés sont majoritairement impliquées dans la survie cellulaire (transduction du signal, morphogénèse, réponses aux stimuli extérieurs et comportement) (Ganko et al., 2006). Le génome de l'hôte est continuellement exposé aux ET et peut conduire au recrutement de séquences d'ET, devenant des nouveaux gènes après leur domestication (Volff, 2006). L'exemple caractéristique de domestication chez la drosophile étant les deux rétrotransposons sans LTR, *TART* et *HeT-A* permettant le maintien des télomères (Biessmann et al., 1990; Levis et al., 1993).

Ces transpositions peuvent également être responsables de variations phénotypiques, sans effets délétères, comme par exemple une diversité de la taille des chiens (Sutter et al., 2007) ou lors de la découverte historique des ET par Barbara McClintock, une variation de pigmentation des grains de maïs est associée à l'activité des transposons *Ac* et *Ds* (McClintock, 1950). La transposition des ET a également été décrite au niveau des neurones liés à la mémoire dans le cerveau de la drosophile, pouvant apporter une hétérogénéité génomique associée à une variabilité de comportement chez les individus (Perrat et al., 2013).

6. Impact d'un stress sur la transposition des ET

Chaque organisme est exposé à différents stress, incluant les stress abiotiques et les stress biotiques. Les stress abiotiques, impliquant des facteurs environnementaux, tels que les changements climatiques, ont été étudiés chez la drosophile. Les ET sont des facteurs clés à l'adaptation environnementale (Casacuberta and González, 2013; González et al., 2008). Il a notamment été identifié lors d'une analyse à l'échelle du génome, que parmi 763 ET annotés, 32 ET présentent une fréquence plus importante dans les populations de *D. melanogaster* provenant d'Amérique du Nord par rapport aux populations d'Afrique (González et al., 2010). Parmis ces ET, il a été proposé que certains sont impliqués dans l'adaptation aux environnements tempérés, alors que la fréquence de certaines autres insertions d'ET est associée à des variables environnementales, telles que la température ou les précipitations chez *D. melanogaster* (González et al., 2010). Ces résultats révèlent l'impact du climat sur la fréquence des ET, permettant l'adaptation des populations de drosophiles à de nouveaux environnements. Le nombre de copies de certains ET de plantes augmente également lors d'un stress. En effet, il a notamment été montré que le nombre de copies des rétrotransposons du tabac *Tto1* et *Tto2*, ou encore des rétrotransposons du riz, *Tos10*, *Tos17* et *Tos19* augmente dans les cellules en culture et au niveau de plantes régénérées à partir de cellules en culture (pour revue : (Wessler, 1996)). De plus, les transcrits de ces ET ont été détectés dans les cellules en culture et non à partir des cellules extraites directement des plantes (Wessler, 1996). Par ailleurs, les insertions d'ET peuvent également être impliquées dans la résistance aux insecticides. C'est notamment le cas de l'insertion de l'élément *accord* au niveau de l'extrémité 5' du gène *Cyp6g1*, permettant la résistance à certains insecticides chez *D. melanogaster* (Daborn et al., 2002). La transposition des ET peut être une réponse de l'hôte permettant son adaptation face au stress en modifiant l'expression de certains gènes.

L'impact des stress biotiques, causés par un organisme étranger, a été beaucoup moins étudié. L'injection de particules du virus RAV-2 (*avian Rous Associated Virus type 2*) dans des embryons de drosophiles a induit la transposition somatique de l'élément *mdg1* mais pas celle de *copia* (Jouan-Dufournel et al., 1996). En ce qui concerne l'impact de la transposition sur un pathogène de drosophile, il a été montré que l'insertion de l'élément *Doc1420* dans le gène *CHKov1*, codant une protéine contenant un domaine kinase choline, est à l'origine d'une

résistance de *D. melanogaster* à l'infection expérimentale par DMelSV (Magwire et al., 2011). Deux transcrits, codant potentiellement des protéines tronquées, sont produits à partir de cette insertion, qui permet une résistance augmentée face aux pesticides organophosphorés (Aminetzach et al., 2005). Postérieurement, deux duplications respectives du gène *CHKov1* et de l'ET inséré, induisent une résistance accrue de *D. melanogaste*r face au DMelSV (**Figure 17.**) (Magwire et al., 2011). Ces duplications impliquent des séquences partielles des gènes codant les protéines CHKov1 et CHKov2, qui sont des protéines parologues, situés dans une région voisine. La première duplication inclut une importante portion de l'extrémité 5' du gène *CHKov1* et environ deux tiers de l'extrémité 3' du gène *CHKov2*. La seconde duplication est en orientation inverse et inclut le gène *CHKov2* et l'extrémité 5' du gène *CHKov1*. Ce réarrangement génétique induit une surexpression du gène *CHKov2*, conférant probablement la résistance au DMelSV. Néanmoins, plusieurs éléments suggèrent que ces mutations ont été sélectionnées et stabilisées lors d'infections par des Rhabdovirus (Magwire et al., 2011). En effet, le nombre de modifications de séquences accumulées dans l'élément *Doc* suggère que l'insertion a eu lieu il y a environ 90000 ans, ce qui est bien antérieur à l'utilisation d'insecticides (Aminetzach et al., 2005). De plus, il y a environ 2000 ans, l'ancêtre commun le plus récent des isolats du virus sigma actuel existait déjà (Carpenter et al., 2007). Par conséquent, l'élément *Doc* aurait initialement joué un rôle dans la défense antivirale des drosophiles, ce qui a apporté un avantage après l'utilisation des traitements aux insecticides organophosphorés (Magwire et al., 2011).

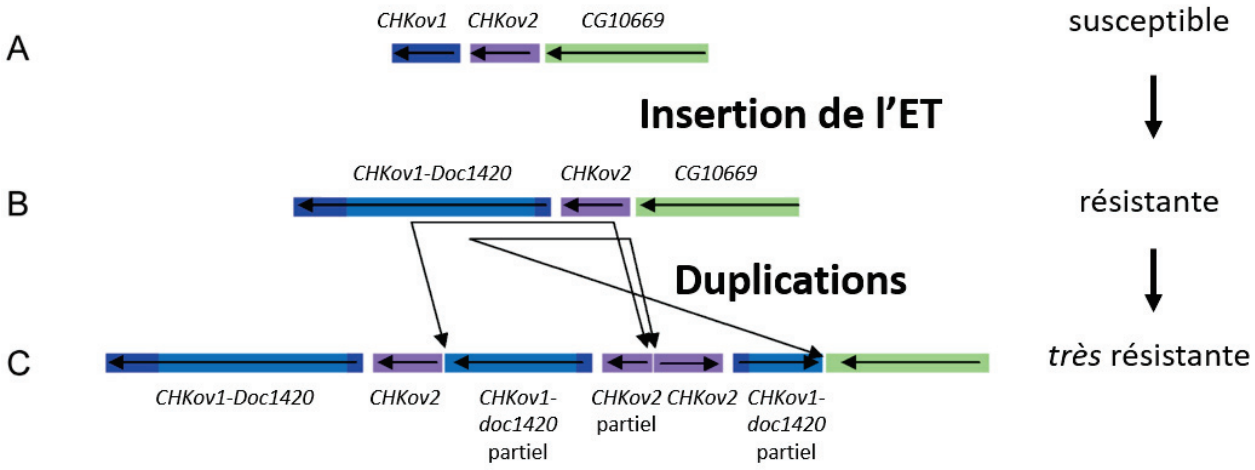


Figure 17. Mécanisme d'acquisition de la résistance au DMelSV. (A) Les drosophiles dont les séquences ne présentent ni d'insertion de *Doc1420* ni de réarrangement génétique sont susceptibles à l'infection DMelSV. **(B)** L'insertion de l'ET *Doc1420* dans le gène *CHKov1* tronque la séquence codante de ce gène, induisant une résistance des drosophiles face au DMelSV. **(C)** Le gène tronqué *CHKov1* a été dupliqué deux fois, incluant deux copies complètes et une copie partielle du gène *CHKov2*. D'après (Magwire et al., 2011)**.**

Le mécanisme de résistance aux insecticides et aux Rhabdovirus n'est pas clairement décrit dans la littérature. Le domaine kinase choline codé par le gène *CHKov1*, pourrait être impliqué dans la résistance aux organophosphorés, impliquant une dérégulation du métabolisme de type choline ou en ciblant plus spécifiquement l'acétylcholine estérase (Aminetzach et al., 2005). De plus, le gène *CHKov2* code aussi une protéine contenant un domaine kinase choline. Il est proposé que la résistance aux Rhabdovirus impliquerait également le métabolisme de type choline, appuyé par le fait que des récepteurs de type acétylcholine sont impliqués dans l'entrée des Rhabdovirus dans la cellule (Lentz et al., 1983). Ces deux propositions restent à élucider.

Chez les plantes, le rétrotransposon *Tnt1* est également activé lors de l'utilisation d'un extrait de champignons au cours de la préparation de protoplastes de feuilles (Pouteau et al., 1994). Il a également été montré que l'amplification du rétrotransposon *Tnt1A*, distribué au niveau des plantes de la famille *Solanaceae* a été stimulée par les extraits fongiques indiquant que l'amplification est corrélée avec l'activité de transcription. De plus, les infections bactériennes et virales induisant une réponse hypersensible de défense des plantes engendrent également la transcription de *Tnt1* (Pouteau et al., 1994). Ces résultats montrent que les facteurs microbiens peuvent générer de la diversité génétique dans les génomes des plantes, à travers l'activation de rétrotransposons (Grandbastien et al., 2005).

V. Le mécanisme d'ARN interférence

Face aux différents perturbateurs, tels que les virus ou les ET, la sélection naturelle a permis la mise au point de mécanismes de défense. Les réponses de l'immunité innée antivirale chez l'insecte incluent le mécanisme d'ARN interférence (ARNi). Ce mécanisme est basé sur la production de petites molécules d'ARN, d'une taille entre 20 nucléotides et 30 nucléotides, pouvant cibler la chromatine et les transcrits et permettant la protection du génome et du transcriptome. Trois classes de petites molécules d'ARN non codant sont caractérisées chez les eucaryotes, incluant les siARN (*small interfering RNA*), les miARN (*microRNA*) et les piARN (*piwi-interacting RNA*), générées par des précurseurs de tailles variables et présentant des origines génomiques et des machineries de production différentes. Ces voies d'ARNi impliquent des effecteurs moléculaires distincts et sont décrites de façon indépendante (**Tableau 2.**).

Propriétés	siARN	miARN	piARN
Taille (nt)	21	20-22	23-30
Modifications chimiques	extrémité 3', méthylation 2'-O	/	extrémité 3', méthylation 2'-O
Biais de séquence	Pas de préférence à l'uracile	Commence préférentiellement par un uracile	1U (lié à Piwi et Aub) ou 10A (lié à Ago3)
Orientation	Proportion identique de brin sens et anti-sens	Sens (relatif au locus miARN)	Anti-sens (lié à Piwi et Aub) plus abondant que les sens (lié à Ago3)
Précurseurs	ARN double brin longs	pri-miARN simple brin, organisé en épingle à cheveux pour la formation des pre-miARN	Simple brin, transcrits à partir des clusters
Effecteurs de biogénèse	Dcr-2	Dcr-1	Voie primaire : Zuc Voie secondaire : Ago3 et Aub
Effecteur de répression	Ago2	Ago1	Voie primaire : Piwi Voie secondaire : Ago3 et Aub

Tableau 2. Propriétés des petits ARN interférents. A partir de données issues d'études chez la drosophile. Dcr- 2 : Dicer-2 ; Ago2 : Argonaute 2 ; Dcr-1 : Dicer-1 ; Ago1 : Argonaute 1 ; Piwi *: P-element induced wimpy testis* ; Aub : Aubergine ; Ago3 : Argonaute 3 ; Zuc : Zucchini ; pri-miRNA : *MicroRNA primary-miRNA* ; pre-miRNA : *MicroRNA precursor-miRNA*. D'après (Fablet, 2014).

1. siARN

Les siARN (*Small interfering RNA*) sont des petits ARN interférents double brin d'une taille de 21 nucléotides, chez les insectes, et permettant le clivage spécifique d'un ARNm. La voie des siARN est médiée par des ARN double brin longs, de source variable, exogène ou endogène. Les sources exogènes proviennent de l'environnement, tel que l'ARN double brin viral qui est libéré ou généré dans le cytoplasme lors d'une infection virale. La voie des siARN, exprimée de façon ubiquitaire, représente l'une des principales voies de l'immunité innée antivirale chez les insectes (Bronkhorst and van Rij, 2014; Gammon and Mello, 2015; Kingsolver et al., 2013; Wang et al., 2006; Zambon et al., 2006). De l'ARN double brin endogène est produit à partir d'une grande variabilité de sources, incluant les éléments transposables.

La voie des siARN est également dédiée au contrôle somatique des ET chez la drosophile, impliquant les endo-siRNA (*Endogenous small interfering RNAs*) (Chung et al., 2008; Czech et al., 2008; Okamura et al., 2008b), comme nous le développerons dans la partie V.1.C, dédiée aux endo-siARN.

A. siARN exogènes

Les sources exogènes d'ARN double brin proviennent de virus ARN ou ADN. Lors de la réplication virale, l'ARN double brin peut exister en tant qu'intermédiaire de réplication pour les virus à ARN simple brin. L'ARN double brin peut également provenir directement des génomes viraux dans le cas des virus à ARN double brin ou à partir de la transcription bidirectionnelle chevauchante des génomes de virus à ADN. La voie des siARN (**Figure 18.**) est initiée par la reconnaissance puis le clivage de molécules d'ARN double brin par une endonucléase de type RNase III, nommée Dicer-2 (Dcr-2) (Galiana-Arnoux et al., 2006; Wang et al., 2006). Cette étape de clivage permet la production de molécules siARN (double brin) d'une longueur de 21 nucléotides chez la drosophile. Ces derniers sont ensuite pris en charge par le complexe RISC, pour *RNA-induced silencing complex* ou complexe de répression d'ARN interférence incluant l'endonucléase Argonaute-2 (Ago2) (van Rij et al., 2006). Les cofacteurs de Dcr-2, *Loquacious* (LOQ-S) (siARN endogènes) et R2D2 (siARN exogènes) sont nécessaires à la formation du complexe RISC et au recrutement de siARN. Le brin passager est éliminé du complexe RISC alors que le brin restant, appelé guide est méthylé en 2'-O par la méthyltransférase Hen1, à l'extrémité 3', permettant la maturation du complexe RISC. Une fois le complexe RISC mature, l'ARN guide se lie spécifiquement à l'ARN viral par complémentarité totale entre chaque paire de base. L'activité RNase d'Ago2 permet le clivage de l'ARN viral et inhibe la réplication virale. De plus, les siARN viraux ont une proportion équivalente de brins sens et anti-sens lorsque Dcr-2 est fonctionnelle, incluant des points chauds et des points froids de production le long du génome viral (Dietrich et al., 2017; Morazzani et al., 2012; Schnettler et al., 2013b).

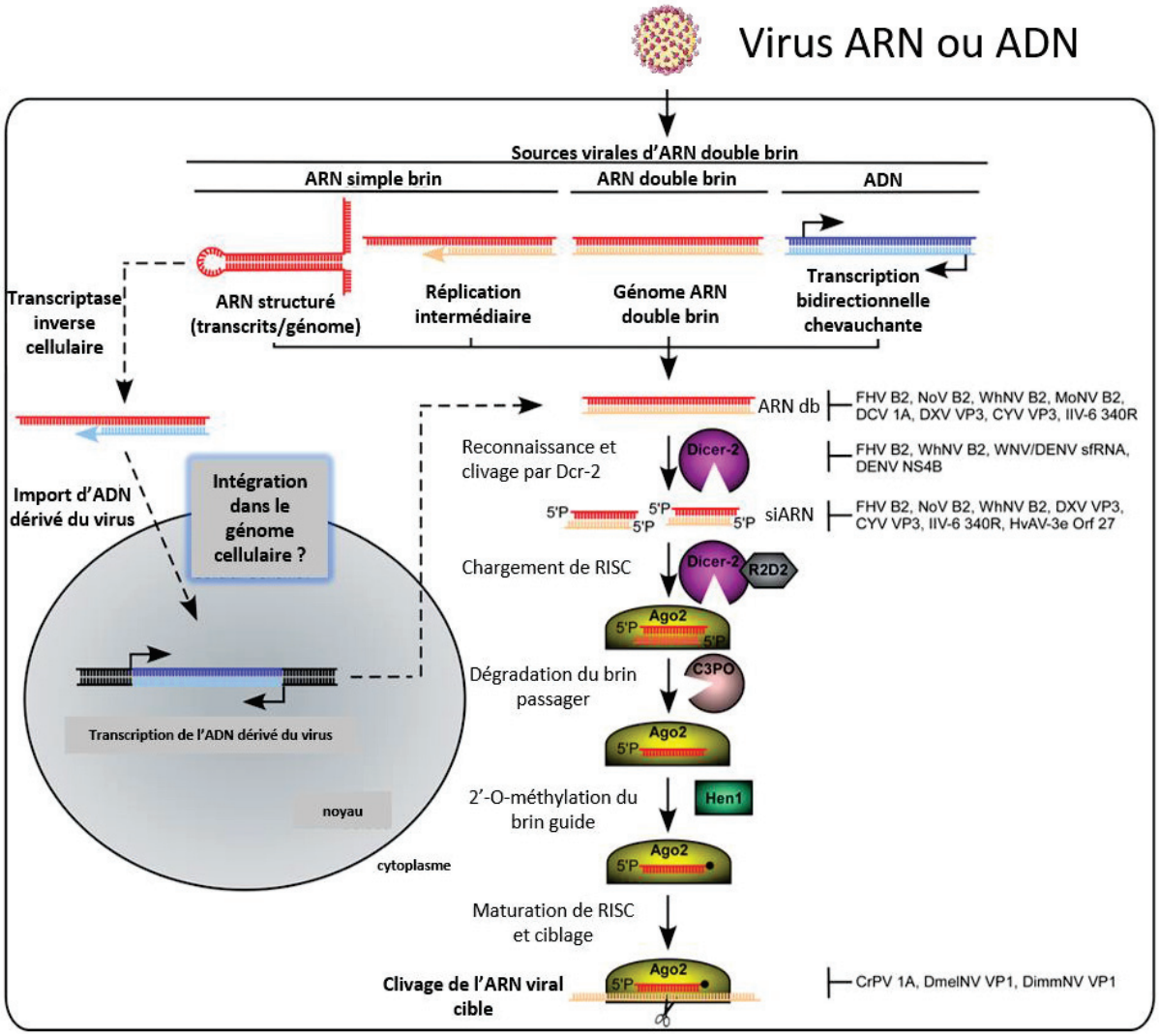


Figure 18. La voie de l'immunité antivirale des siARN chez les insectes. Les ARN double brin proviennent d'une source variable et sont reconnus par l'endonucléase Dcr-2. Cette dernière clive les ARN double brin, permettant la production des petites molécules d'ARN double brin, d'une taille de 21 nucléotides, appelés siARN. Ces siARN sont pris en charge par le complexe RISC, incluant la protéine Ago2. Cette endonucléase dégrade un des deux brin, appelé brin passager et la méthyltransférase Hen1 méthyle le carbone 2' du ribose (2'-O) au niveau de l'extrémité 3' du brin guide restant. Le complexe RISC mature cible spécifiquement l'ARN viral à dégrader, par complémentarité de séquence parfaite avec le brin guide permettant à Ago2 de le cliver. Goic et collaborateurs ont proposé l'implication d'activité de transcriptase inverse cellulaire (traits en pointillés), provenant de rétrotransposons à LTR, dans la formation d'ADN viral (Goic et al., 2013). Ce dernier est issu de la transcription inverse de l'ARN de différents virus et induit une nouvelle source d'ARN double brin, contenant des séquences virales (Goic et al., 2013). Ces ARN double brin entrent dans la voie des siARN et permettent la production de nouveaux siARN qui seraient impliqués dans la mise en place d'une infection persistante chez *Drosophila* (Goic et al., 2013) et de la tolérance virale chez le moustique (Goic et al., 2016). L'intégration de l'ADN viral dans le génome de l'hôte n'est pas clairement élucidée. De plus, de nombreux virus, à génome ARN ou ADN codent des suppresseurs de la voie d'ARNi. La protéine virale identifiée comme étant le suppresseur est indiquée dans la figure pour les virus suivants : FHV, *Nodamura virus* (NoV), *Wuhan Nodavirus* (WhNV) et le mosinovirus (MoNV) : protéine B2 ; DCV, protéine A1 ; DXV, *Culex Y virus* (CYV) : protéine VP3 ; IIV-6 protéine 340R ; WNV et DENV : sfRNA (*sub-genomic flavivirus RNA*) ; *Heliothis virescens ascovirus-3e* (HvAV-3e), ORF 27 (RNase III). D'après (Gammon and Mello, 2015).

B. Réponse ARN interférence systémique chez *Drosophila*

Il a été montré que la voie des siARN se met en place de façon systémique chez les insectes lors d'infections virales (Karlikow et al., 2014). Des études récentes ont notamment permis d'identifier une voie d'amplification de la réponse siARN lors de l'infection par un virus ARN chez *Drosophila* (Goic et al., 2013) et chez le moustique (Goic et al., 2016). La transcription inverse de l'ARN viral permet la synthèse d'ADN viral, impliquant une activité de transcriptase inverse cellulaire provenant probablement de rétrotransposons à LTR (Goic et al., 2013) et permettant la production de siARN secondaires. L'inhibition de la production d'ADN viral par un traitement à l'azidothymidine (AZT) chez la drosophile infectée par FHV ou chez le moustique infecté par CHIKV induit une perte de la persistance et de la tolérance virale (Goic et al., 2016, 2013).

Les mécanismes qui permettent le transfert intercellulaire des signaux induisant une voie ARNi systémique ne sont pas totalement élucidés. Cependant, il a été montré que lors d'une réponse virale systémique, ces signaux antiviraux sont disséminés depuis le site d'infection jusqu'aux tissus non infectés distants (Saleh et al., 2009). Il a notamment été décrit que les récepteurs *scavenger* sont impliqués dans la mise en place d'une voie d'ARNi systémique, en permettant l'entrée d'ARN double brin (injecté artificiellement) dans la cellule (Saleh et al., 2006). Les gènes *egghead* (*egh*), *NinaC* et *CG4572* ont été identifiés comme étant impliqués dans la voie de l'import de d'ARN double brin (Saleh et al., 2006). Ces gènes sont indispensables à la tolérance virale chez la drosophile (sans élimination préalable des infections chroniques) au temps précoce de l'infection par DCV ou SINV (Saleh et al., 2009). Le gène *egh* code une glycosyltransférase à sept domaines transmembranaires, le gène *NinaC* code une protéine impliquée dans le transport de vésicule ; en revanche, la fonction du gène *CG4572* n'a pas été identifiée. Les drosophiles mutées pour ces différents gènes succombent rapidement à l'infection et présentent des titres viraux 10^1 à 10^5 plus importants que pour les drosophiles sauvages (Saleh et al., 2009). La prise en charge de ces ARN double brin, induits lors de l'infection, semble indispensable à la mise en place d'une réponse de type ARNi systémique (Saleh et al., 2009). Ces mêmes auteurs suggèrent que l'ARN double brin serait libéré de la cellule après sa lyse, du fait des effets cytopathiques du virus, et serait pris en charge par l'hémolymphe pour une distribution systémique dans l'ensemble de l'organisme. D'autres mécanismes proposés impliquent des structures de type nanotube, identifiées au niveau de cellules de drosophiles infectées de façon persistante par DCV ou FHV, permettant le transport d'ARN double brin et de composants du complexe RISC (Karlikow et al., 2016). Des vésicules de type exosome, dérivées d'hémocytes sont également impliquées dans la diffusion systémique d'ARN double brin lors de l'infection par SINV (Tassetto et al., 2017). Ces auteurs ont proposé un modèle de diffusion systémique de siARN par les hémocytes (**Figure 19.**). La prise en charge d'ARN double brin par les hémocytes stimule la production d'ADN viral de façon dépendante d'Ago2. Ces derniers servent de matrice pour la production secondaire de siARN et engendrent une activité antivirale systémique après libération par les hémocytes

sous forme de vésicules de type exosome. L'ensemble de ces données met en évidence plusieurs mécanismes, permettant la diffusion systémique des siARN (**Figure 20.**).

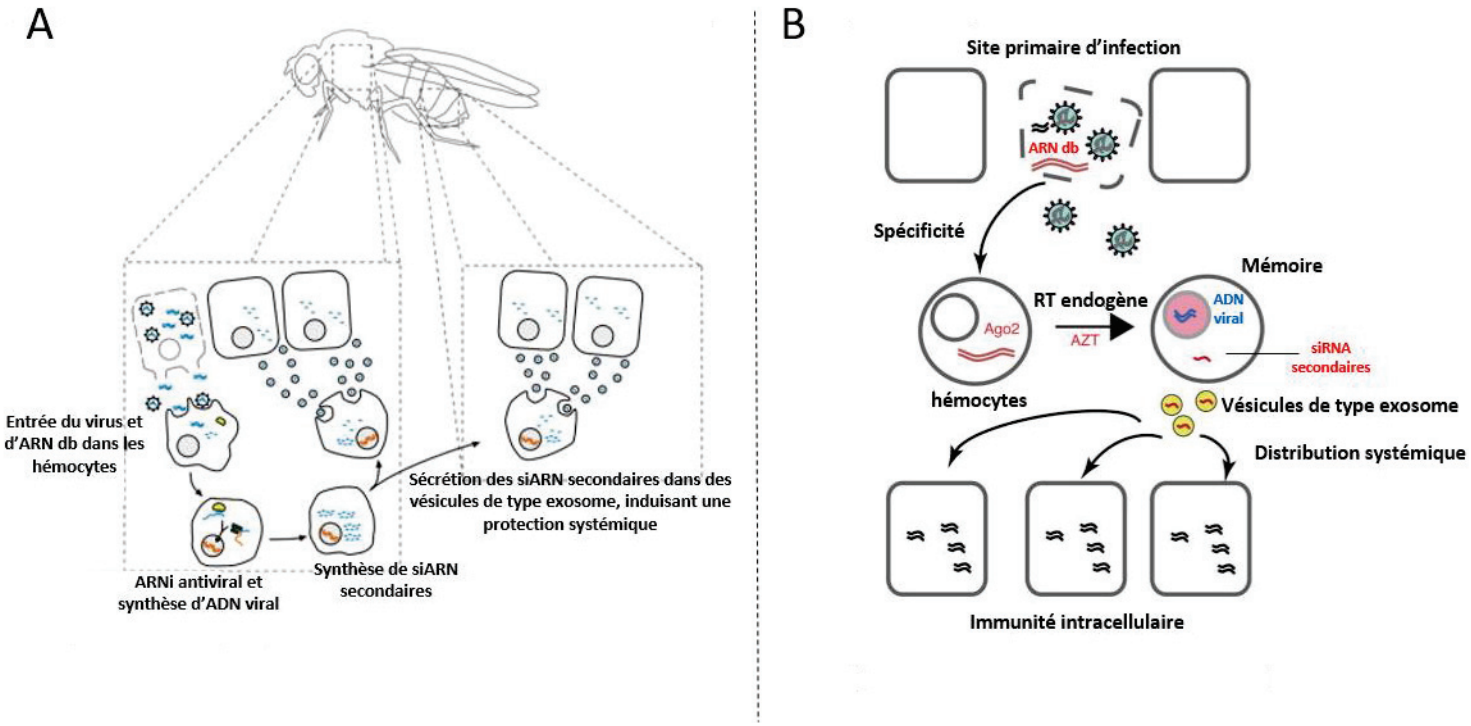


Figure 19. Diffusion systémique de siARN secondaires par les hémocytes chez la drosophile. (A) Lors de l'infection virale de SINV, les hémocytes prennent en charge le virus et l'ARN double brin libéré par les cellules infectées. La synthèse d'ADN viral double brin permet la production de siARN secondaires. Les vésicules de type exosome induisent la libération des siARN secondaires qui vont être pris en charge par de nouvelles cellules non infectées, induisant une protection antivirale. **(B)** L'ARN double brin de SINV pris en charge par les hémocytes induit une protection systémique spécifique de ce virus. Une activité de transcriptase inverse est nécessaire à la formation d'ADN viral à partir de l'ARN double brin. L'ADN viral, nécessaire à la production de siARN secondaires, constitue une mémoire antivirale pour l'organisme. D'après (Tassetto et al., 2017).

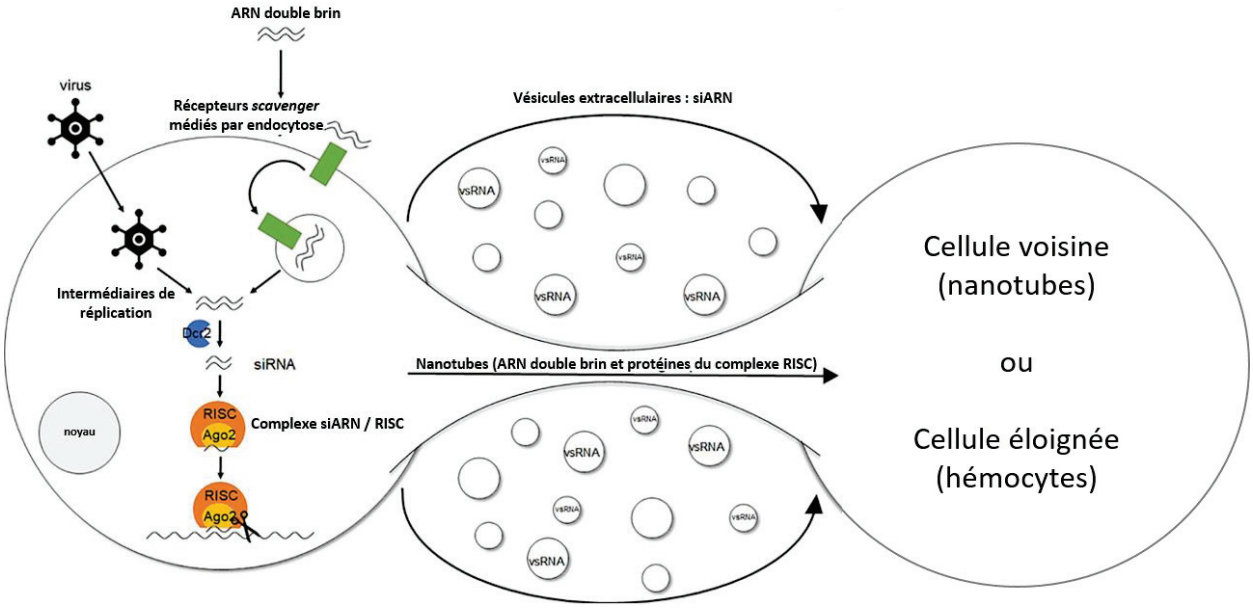


Figure 20. siARN et diffusion systémique chez les insectes. Le génome du virus est libéré dans le cytoplasme de la cellule apportant une source d'ARN double brin. Dans le cas d'un apport artificiel d'ARN double brin, leur entrée implique l'endocytose de récepteurs *scavenger*. Les ARN double brin sont ensuite pris en charge par la voie des siARN dans la cellule. Lors de l'infection, la formation des structures en nanotubes augmente et permet le transport, sur une courte distance, d'ARN double brin et de composants du complexe RISC, tel que Ago2 (Karlikow et al., 2016). La diffusion systémique sur l'ensemble de l'organisme impliquerait les hémocytes et permettrait la libération de vésicules de type exosome contenant les siARN secondaires viraux (vsRNA) (Tassetto et al., 2017). Schéma à partir de (Vogel et al., 2019).

Poirier et collaborateurs ont précisé la nature, la biogenèse et la régulation de l'ADN viral synthétisé (Poirier et al., 2018) permettant l'amplification de la production des siARN chez *Drosophila* (Goic et al., 2013; Tassetto et al., 2017) et chez le moustique (Goic et al., 2016). L'ADN viral produit lors de l'infection par un virus ARN se présente sous forme linéaire ou circulaire (Poirier et al., 2018). Ces mêmes auteurs ont montré que des génomes viraux défectueux du virus étudié (DVGs, *defective viral genomes*) servent de matrice pour la synthèse des différentes formes d'ADN viral. La synthèse d'ADN viral est régulée par le domaine hélicase DExD/H de Dcr-2 par l'intermédiaire d'un domaine distinct de celui impliqué dans la génération des siARN (Poirier et al., 2018).

C. endo-siARN

Les sources endogènes d'ARN double brin sont variables selon l'organisme et aboutissent à la production de petits ARN interférents double brin de 21 nucléotides, appelés endo-siARN. Le séquençage à haut débit des petits ARN interférents chez *D. melanogaster* a révélé une nouvelle source endogène de production d'endo-siARN : les éléments transposables (Czech et al., 2008; Kawamura et al., 2008). Ces endo-siARN sont impliqués dans la régulation des transposons dans les tissus somatiques (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008b), tandis que la voie des piARN cible principalement les éléments transposables au niveau des tissus germinaux chez la drosophile (Brennecke et al., 2007).

Chez *Drosophila*, les RdRp étant absentes, les ARN double brin longs peuvent provenir de la transcription bidirectionnelle d'un locus (**Figure 21 A.**), de l'interaction de transcrits complémentaires dérivés de locus séparés, comme les ET ou les pseudogènes (**Figure 21 B.**), ou de la formation de structures en épingle à cheveux, distinctes des miARN, comme celles provenant de séquences répétitives inversées (**Figure 21 C.**) (Claycomb, 2014). L'ARN double brin est pris en charge par Dcr-2 et son co-facteur Loqs-PD, permettant la production des endo-siARN de 21 nucléotides. De la même façon que pour les siARN exogènes, les duplex d'endo-siARN sont pris en charge par le complexe RISC et notamment par l'endonucléase Ago2. Cette protéine clive l'ARNm cible, après reconnaissance par complémentarité totale de séquence entre l'ARN guide des endo-siARN et l'ARN cible.

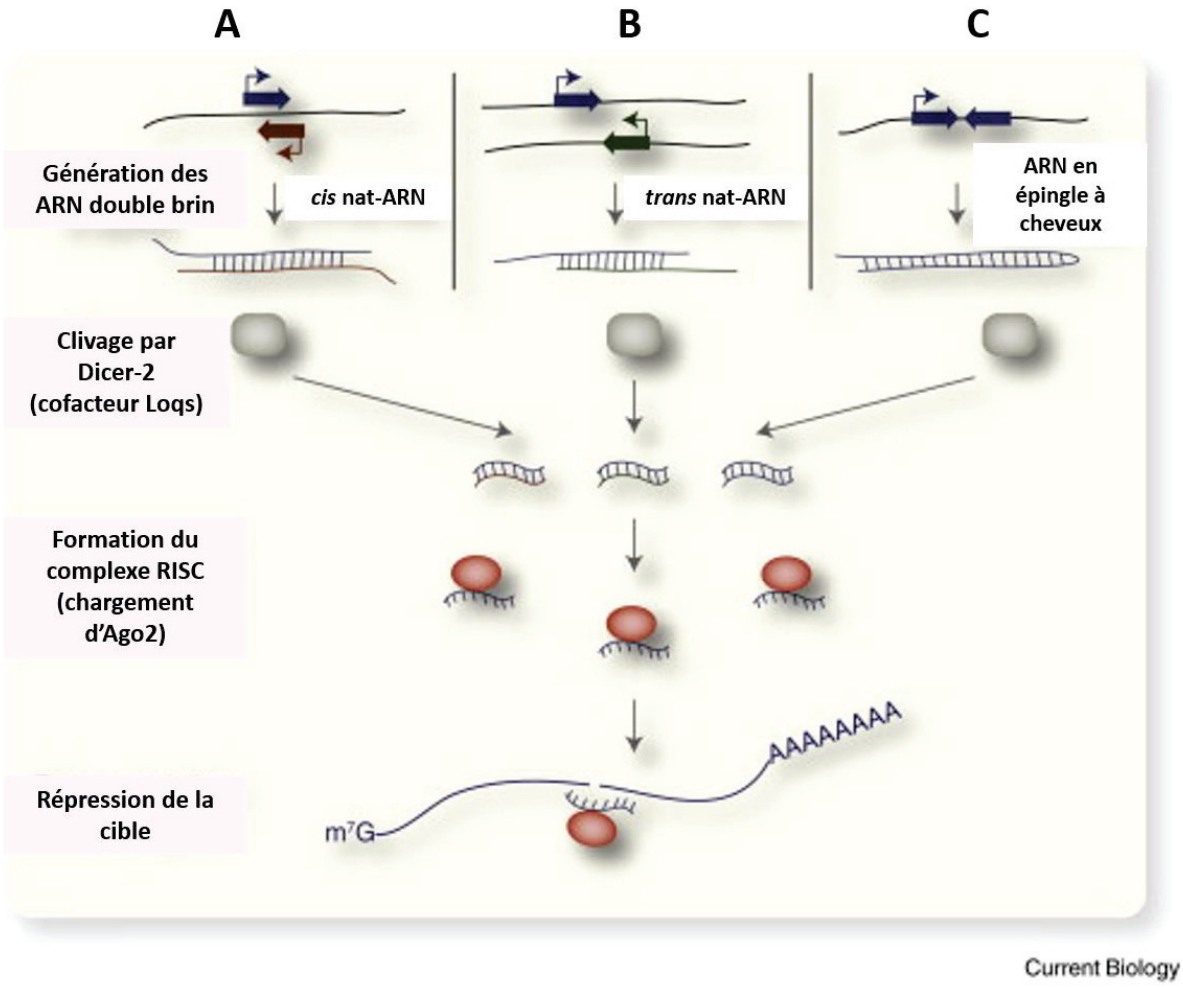


Figure 21. Production d'endo-siARN chez la drosophile. L'ARN double brin est produit à partir de différentes sources : **(A)** transcrits antisens *cis* qui se chevauchent au niveau de leur extrémités 3', aboutissant à un ARN double brin issu d'une transcription bidirectionnelle ; **(B)** transcrits antisens naturels *trans*, possédant une complémentarité (à partir de différentes régions du génome) ; **(C)** transcrits en épingle à cheveux. Dcr-2 et son cofacteur endogène Loqs prennent en charge l'ARN double brin, permettant la production des duplex siARN de 21 nucléotides. Le complexe RISC, incluant Ago2 prend en charge les endo-siARN et cible l'ARNm à réguler par complémentarité de séquence avec le brin guide des endo-siARN. D'après (Claycomb, 2014).

La voie des siARN endogènes ciblant les ET est particulièrement active dans différentes lignées cellulaires immortalisées (Chung et al., 2008; Lau et al., 2009) et implique des facteurs communs avec la voie des siARN exogènes, incluant une signature des séquences de 21 nucléotides chez la drosophile. Ces endo-siARN présentent également une modification 2'-O-méthyle à leur extrémité 3', catalysée par la méthylase Hen1 (Chung et al., 2008; Claycomb, 2014; Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Lau et al., 2009; Okamura et al., 2008b, 2008a).

L'invalidation des protéines Dcr-2 ou Ago2 augmente les niveaux de transcrits d'ET, au niveau de cellules somatiques de drosophiles (Chung et al., 2008; Czech et al., 2008; Ghildiyal et al.,

2008; Kawamura et al., 2008; Okamura et al., 2008a, 2008b). *In vivo*, les mutants Ago2 de drosophiles révèlent des niveaux d'expression des ET plus importants que chez les drosophiles sauvages, dans le cerveau, et particulièrement au niveau de leurs neurones. Cette baisse de contrôle des ET conduit à plusieurs défauts de comportement et à une durée de vie plus courte (Li et al., 2013). Ces résultats indiquent que la voie des siARN est impliquée dans le contrôle des ET, particulièrement au niveau des tissus somatiques chez la drosophile. Elle est complémentaire de la voie des piARN permettant de protéger l'intégrité du génome dans la lignée germinale. La perte des facteurs de la voie des siARN endogènes et exogènes, tels que Dcr-2 et Ago2 conduit également à une hausse de sensibilité aux infections virales (Galiana-Arnoux et al., 2006; Kemp et al., 2013; van Rij et al., 2006; Wang et al., 2006).

2. miARN

Les miARN (*microRNA*) sont des petites molécules d'ARN non codant, d'une taille entre 21 et 22 nucléotides de longueur, conservées et générées par tous les eucaryotes. La voie des miARN est active au niveau des tissus somatiques et germinaux. La fonction principale des miARN est de réguler l'expression des gènes de l'hôte en utilisant des mécanismes post-transcriptionnels de répression, appelés PTGS (*post-transcriptional gene silencing*), inhibant spécifiquement la traduction de l'ARNm cible ou en diminuant leur demi-vie dans le cytoplasme. La voie des miARN est impliquée dans la régulation de différents processus biologiques, tels que le développement des organes, la différenciation cellulaire et l'homéostasie cellulaire chez les insectes (Asgari, 2013) et chez l'Homme (Fu et al., 2013). Les altérations au niveau de la voie des miARN entraînent des maladies chez l'Homme, tels que des défauts de développement ou des cancers (Paul et al., 2018; Tüfekci et al., 2014).

La majorité des miARN proviennent de miARN primaires (pri-miARN), issus de la transcription de séquences d'ADN et permettant ensuite la production des miARN précurseurs (pre-miARN) et des miARN matures (**Figure 22.**). Les pri-miARN forment des structures d'ARN double brin en épingle à cheveux, activant la voie. Les pri-miARN sont pris en charge par le complexe ribonucléasique de type III, composé de Drosha et Pasha, permettant la production des pre-miARN, qui seront ensuite exportés dans le cytoplasme (Denli et al., 2004; Lee et al., 2003). Une fois dans le cytoplasme, la maturation des pre-miARN en miARN double brin d'environ 22 nucléotides, implique une seconde RNase de type III, Dicer-1 (Dcr-1) (Park et al., 2011) et son co-facteur Loqs, présentant plusieurs isoformes d'épissage (Loqs-PA, Loqs-PB) (Fukunaga et al., 2012; Saito et al., 2005). Notons que dans la voie des endo-siARN, l'interaction entre Loqs et Dcr-2 (présenté dans la partie V.1.C), n'implique pas les mêmes isoformes. Les miARN matures sont ensuite pris en charge par le complexe RISC, incluant Argonaute 1 (Ago1). Un des deux brins d'ARN est libéré dans le cytoplasme, et forme un complexe RISC mature (Miyoshi et al., 2005). La reconnaissance par le second brin ne nécessite pas une parfaite

homologie. Les ARNm codant les protéines sont les premières cibles, aboutissant à une inhibition de la traduction ou à la dégradation de l'ARNm par Ago1.

Dans la majorité des cas, les miARN interagissent avec l'extrémité 3' UTR de l'ARNm cible pour bloquer son expression (Ha and Kim, 2014). Les miARN peuvent également interagir avec les extrémités 5' UTR, les séquences codantes et les promoteurs des gènes (Broughton et al., 2016). Il a été reporté, selon certaines conditions, que les miARN peuvent stimuler l'expression des gènes de façon post-transcriptionnelle (Vasudevan, 2012). Des études récentes suggèrent que les miARN peuvent agir comme des molécules de signalisation entre différents compartiments cellulaires, qui contrôlent le taux de transcription et de traduction (Makarova et al., 2016; O'Brien et al., 2018).

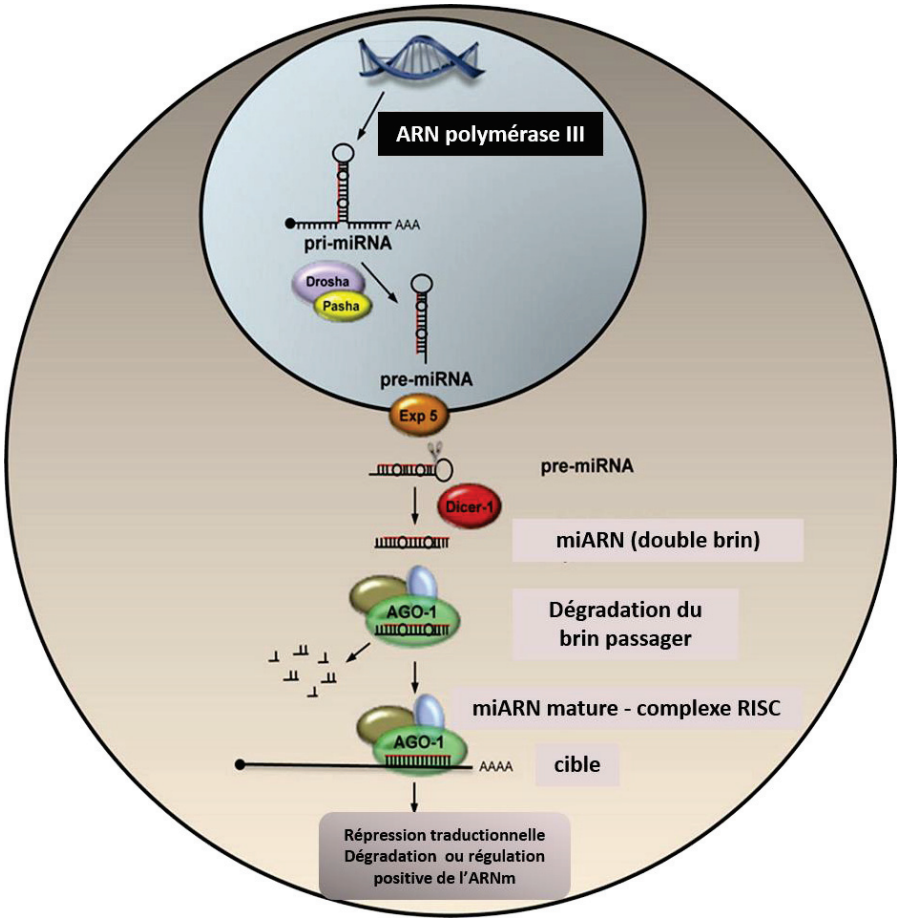


Figure 22. La voie des miARN chez les insectes. La voie est initiée par la transcription des gènes des miARN, permettant la production des miARN primaires (pri-miRNA) qui sont ensuite clivés par le complexe de deux ribonucléases de type III, Drosha et Pasha, en miARN précurseurs (pre-miRNA). Ces derniers sont transportés dans le cytoplasme, par l'intermédiaire de l'exportine 5 (Exp 5) et sont pris en charge par Dcr-1 et Loqs, permettant la production des miARN double brin. Le complexe RISC, incluant Ago1, prend en charge le duplex miARN et clive un des deux brins. Le second brin guide le complexe RISC-Ago1 jusqu'à la cible, pouvant aboutir à la répression traductionnelle, à la dégradation ou à la régulation positive de l'ARNm cible. D'après (Asgari, 2013).

3. piARN

A. Introduction générale

La sélection naturelle favorise le maintien de la stabilité génomique par l'intermédiaire de voies de régulation du contenu et de l'activité des ET, permettant d'éviter la survenue de mutations délétères (Aravin et al. 2007). Cette défense génomique est réalisée par la voie des piARN (*Piwi-interacting RNAs*), appelés initialement « rasiRNAs » pour *repeat-associated small interfering RNAs* (Klenov et al., 2007; Vagin et al., 2006). Il s'agit de la même façon que pour le contrôle de la réplication virale (siARN), de petits ARN interférents mais d'une classe différente.

La voie des piARN produit des petits ARN interférents plus longs que les autres classes de petits ARN interférents. Une des principales différences avec les siARN et les miARN est qu'ils proviennent de longs précurseurs d'ARN simple brin transcrits à partir de locus chromosomiques et que leur production ne dépend ni des protéines Dicer ni de Drosha (Brennecke et al., 2007). La production de ces piARN, d'une taille comprise entre 23 et 30 nucléotides chez *Drosophila*, repose sur l'association avec d'autres protéines spécifiques, de la sous-famille PIWI (Czech and Hannon, 2016; Iwasaki et al., 2015). La taille des piARN varie chez les autres organismes, allant de 21 à 35 nucléotides de long.

La voie des piARN a été initialement identifiée dans les gonades de *D. melanogaster*. L'activité des piARN a été largement décrite au niveau de la lignée germinale et des cellules environnantes et correspond notamment à des mécanismes de contrôle de l'expression des ET au niveau post-transcriptionnel (Yamashiro and Siomi, 2018). Cette régulation repose sur l'hybridation des piARN au niveau des transcrits d'ET, par complémentarité de séquence (non-totale) et implique leur clivage par l'intermédiaire de l'activité catalytique des protéines PIWI (Ghildiyal and Zamore, 2009; Senti and Brennecke, 2010; Siomi et al., 2011). Les piARN contrôlent également les ET par l'intermédiaire de mécanismes transcriptionnels (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012). En effet, les piARN peuvent cibler les transcrits d'ET naissants par l'intermédiaire de mécanismes de modifications de la chromatine (Iwasaki et al., 2016; Luteijn and Ketting, 2013). C'est notamment le cas dans les cellules somatiques ovariennes, où Piwi réprime les ET de façon transcriptionnelle.

B. piARN germinaux et somatiques dans les structures ovariennes

Les piARN protègent la lignée germinale par le contrôle des ET, chez les animaux. Les protéines PIWI et les piARN sont majoritairement exprimés au niveau de la lignée germinale chez la drosophile. Cependant, les piARN sont également efficaces dans les cellules

somatiques des gonades de la drosophile (C. Li et al., 2009; Malone et al., 2009). En dehors des gonades, des piARN somatiques fonctionnels ont été identifiés dans des lignées cellulaires spécifiques de *D. melanogaster*, dans le corps gras (Jones et al., 2016) ou le cerveau (Ghildiyal et al., 2008; Perrat et al., 2013). Les piARN somatiques, non gonadiques, sont particulièrement répandus chez d'autres insectes, comme chez les arthropodes hématophages, tels que les moustiques (Lewis et al., 2018; Miesen et al., 2015; Morazzani et al., 2012).

Chez la drosophile, on distingue 3 protéines PIWI : Ago3, Piwi et Aubergine (Aub). Contrairement à Aub et Ago3, qui sont des protéines cytoplasmiques, la protéine Piwi est aussi localisée dans le noyau des cellules. Ces protéines sont impliquées dans le contrôle des ET au niveau de la lignée germinale et sont par conséquent nécessaires au maintien de l'intégrité génomique et au développement des gonades (Iwasaki et al., 2015; Siomi et al., 2011; Vagin et al., 2006). Au niveau des cellules folliculaires des ovaires (**Figure 23.**), la production des piARN implique des mécanismes médiés par la voie de biogenèse primaire des piARN, incluant notamment la protéine Zucchini (Zuc) (Huang et al., 2017; Malone et al., 2009; Saito et al., 2009).

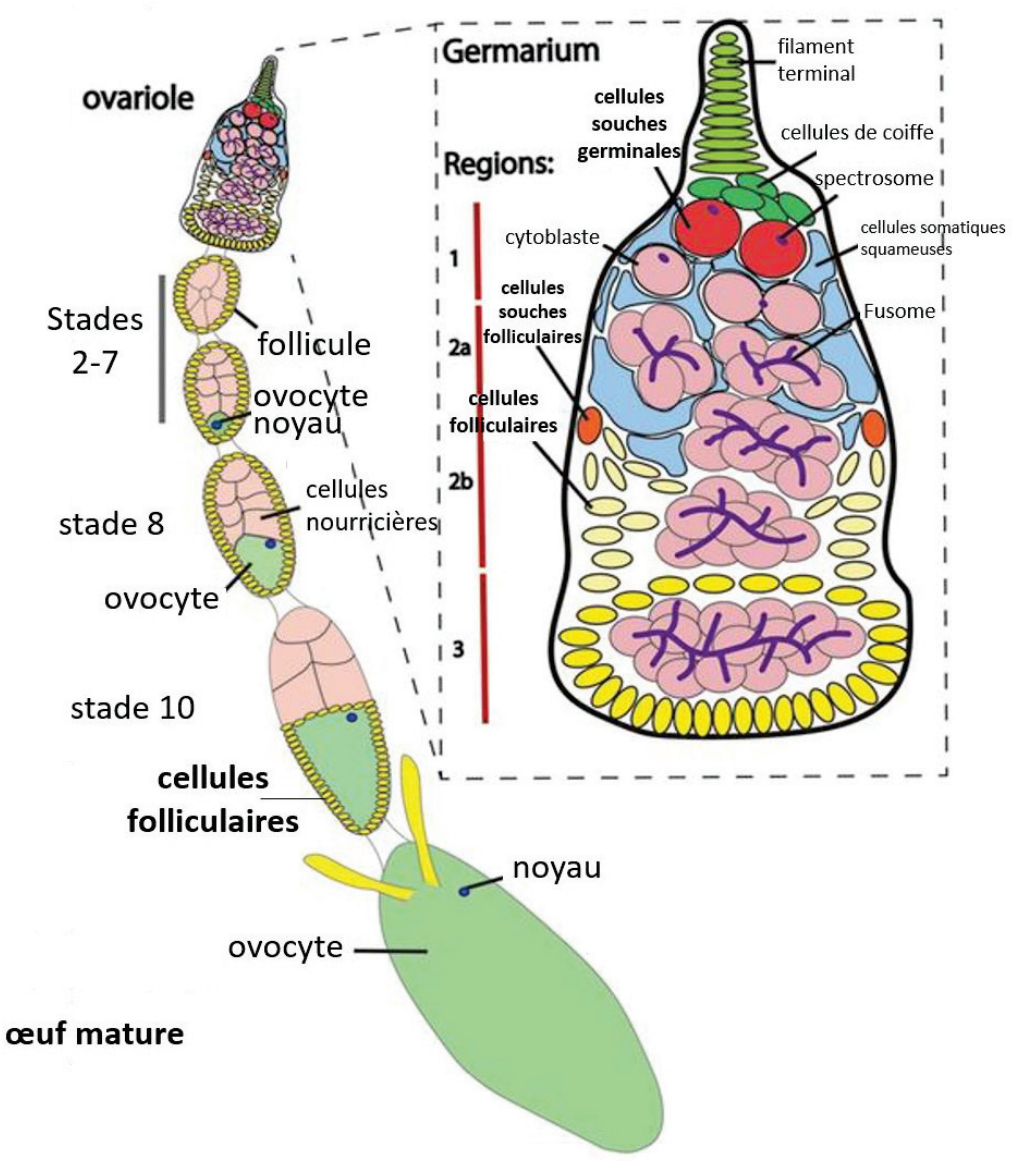


Figure 23. Schéma de l'ovogénèse et structure du germarium. Les ovaires de drosophile sont constitués de 16 à 20 ovarioles qui différencient progressivement les chambres à œufs, aboutissant à la production des œufs. Le rectangle en pointillé représente le germarium, qui est à l'origine des chambres à œufs et est situé à l'extrémité antérieure de chaque ovariole. Le germarium représente la niche des cellules souches germinales et contient également les cellules germinales prolifératives, actives tout au long de la vie adulte, permettant la production des œufs. Au début du développement, les cellules germinales précoces s'associent intimement aux cellules somatiques voisines, notamment les cellules somatiques de morphologie squameuse et les cellules folliculaires (jaunes). L'ensemble de cellules germinales interconnectées est associé à un autre type de cellules somatiques, les cellules folliculaires somatiques. Ces dernières sont dérivées de deux cellules souches folliculaires somatiques (orange). Lorsque qu'un ensemble de 16 cellules est entouré de cellules folliculaires, il forme une chambre à œufs et bourgeonne à partir du germarium pour continuer sa maturation (Dävring and Sunner, 1973). Une des 16 cellules se développera en ovocyte par le processus de méïose. Les autres cellules se développeront en cellules nourricières, favorisant la croissance de l'ovocyte. D'après (Gleason et al., 2018).

a) Voie primaire

Lors de la voie primaire des piARN (**Figure 24.**), les longs précurseurs d'ARN simple brin des piARN, sont transcrits par l'ARN polymérase II, sous le contrôle du facteur de transcription *Cubitus interruptus* (Goriaux et al., 2014), à partir de locus génomiques spécifiques, appelés cluster de piARN. Ces derniers sont riches en répétitions, incluant des ET et sont dépourvus de gènes codant des protéines. Ces ET ont accumulé de nombreuses mutations, ce qui a induit la perte de leur capacité à transposer dans le génome. Ils représentent une base de données d'ET qui pourront être reconnus et contrôlés par la voie des piARN. Lors de l'invasion par un nouvel ET, dans une population hôte naïve, sa transposition dans le génome sera contrôlée seulement après son intégration dans un cluster, permettant la production de piARN dérivés de cet ET. Le principal cluster somatique de piARN est *flamenco*, qui est le régulateur majeur des rétrotransposons de la famille Gypsy tels que *gypsy*, *ZAM* et *idefix* (Desset et al., 2003, 1999; Guida et al., 2016; Malone et al., 2009; Pélisson et al., 1994). Le locus flamenco d'environ 180 kb est localisé dans la région hétérochromatique du chromosome X et est principalement constitué de fragments de transposons antisens, majoritairement de la classe des rétrotransposons à LTR (Brennecke et al., 2007; Mével-Ninio et al., 2007). Ce cluster de petits ARN interférents apporte une source de piARN dans les ovaires, tandis que dans les tissus somatiques il fournit une source d'endo-siARN (Brennecke et al., 2007; Ghildiyal et al., 2008). De plus, les populations naturelles de *D. melanogaster* peuvent présenter un polymorphisme permissif ou restrictif pour *flamenco*, permettant une régulation maternelle de l'expression de *gypsy* (Pélisson et al., 1997; Prud'homme et al., 1995). Dans le cas des drosophiles permissives pour *flamenco*, la transmision maternelle de *gypsy* a été identifiée, incluant la détection de nouvelles copies provirales de *gypsy* dans la lignée germinale de la descendance (Touret et al., 2014a).

Dans une plus faible proportion, une deuxième source de piARN inclut des gènes codant des protéines (Robine et al., 2009). Le gène codant le facteur de transcription Traffic Jam (tj) est le plus important (Saito et al., 2009). Le gène *traffic jam*, spécifique des cellules somatiques ovariennes, fournit un substrat pour la production de piARN somatiques, par l'intermédiaire de l'extrémité 3'UTR de son transcrit. La protéine codée par *traffic jam* active l'expression de Piwi dans les cellules somatiques ovariennes (Saito et al., 2009).

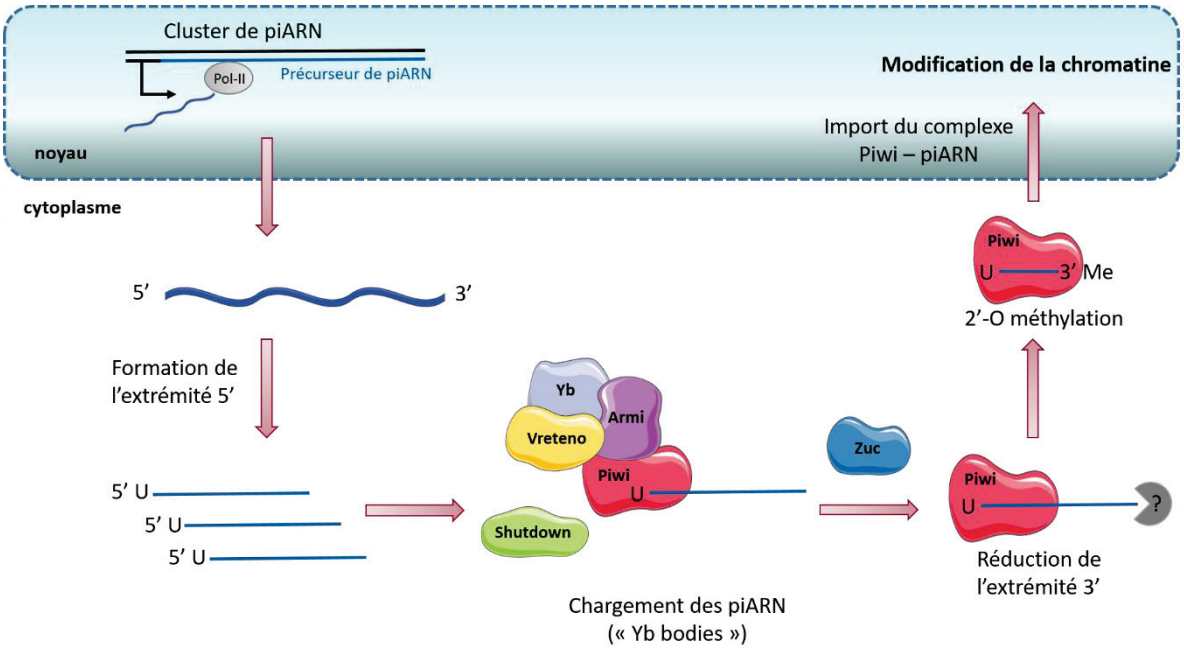


Figure 24. Voie primaire de biogenèse des piARN dans les cellules somatiques ovariennes chez la drosophile. Les piARN sont générés à partir des clusters de piARN, tels que *flamenco*, l'extrémité 5' UTR des ARNm ou le locus spécifique des cellules somatiques ovariennes, appelé *traffic jam*. Ils sont transcrits en long précurseurs de piARN, simple brin, par l'ARN polymérase II puis exportés dans le cytoplasme. La maturation de ces précurseurs en piARN implique un clivage de l'extrémité 5', par une endonucléase non décrite et conduit à la présence d'un uracile à l'extrémité 5', qui permettra sa reconnaissance par Piwi ou Aub (Senti and Brennecke, 2010). Les précurseurs de piARN sont transportés au niveau d'un compartiment cytoplasmique périnucléaire, proche des mitochondries, nommé « Yb bodies ». Le chargement des précurseurs de piARN sur Piwi implique les protéines Yb, Armitage (Armi) et Vreteno, qui interagissent ensemble. La protéine Shutdown est une protéine chaperonne nécessaire au chargement des piARN, son rôle spécifique n'étant pas décrit. L'endonucléase putative Zuc est nécessaire à la maturation des piARN et à la localisation nucléaire de Piwi. L'extrémité 3' des piARN est raccourcie par une exonucléase inconnue, et est modifiée par une 2'-O méthylation, permettant d'avoir des piARN stables entre 23 et 30 nucléotides de longueur. Le complexe Piwi – piARN est ensuite importé dans le noyau où il module l'état chromatinien. D'après (Ross et al., 2014).

Le mécanisme d'export des précurseurs de piARN dans le cytoplasme n'est pas décrit. Il est fortement suggéré que Zuc, une protéine de la membrane externe mitochondriale, permette la formation des extrémités 5' des précurseurs de piARN, par l'intermédiaire de son activité nucléasique (Ipsaro et al., 2012; Nishimasu et al., 2012; Olivieri et al., 2010).

Les précurseurs de piARN présentent une forte préférence pour le résidu uridine au niveau de l'extrémité 5' (Yan et al., 2011), pouvant améliorer la liaison avec Piwi (Kawaoka et al., 2011). Chez la drosophile, les précurseurs de piARN matures entrent dans une structure cytoplasmique spécifique des cellules somatiques ovariennes, le corps périnucléaire Yb appelé « Yb body » (Olivieri et al., 2010; Qi et al., 2011; Saito et al., 2010). Plusieurs protéines sont nécessaires à la production des piARN matures, incluant l'hélicase Armitage, la protéine Vreteno contenant un domaine TUDOR, et la protéine Yb. Ces deux dernières protéines sont

essentielles pour l'expression et la stabilité de Piwi. La protéine Zuc est également nécessaire à cette maturation puisque sa délétion entraîne une accumulation d'intermédiaires de piARN au niveau des cellules somatiques ovariennes (Nishimasu et al., 2012). Piwi prend en charge les piARN matures, par un mécanisme non élucidé mais impliquant la protéine chaperonne Shutdown, qui est essentielle à la biogenèse de l'ensemble des piARN (Olivieri et al., 2012). Le chargement des piARN engendre un changement de conformation de Piwi (Yashiro et al., 2018). Les piARN sont ensuite clivés au niveau de leur extrémité 3' puis méthylés en 2'-O par la méthyltransférase Pimet (Saito et al., 2007), augmentant probablement leur stabilité (Ross et al., 2014). Le complexe Piwi – piARN est ensuite importé dans le noyau où il effectuera ses fonctions de régulation épigénétique. La localisation nucléaire de Piwi requiert l'extrémité N-terminale et son chargement sur les piARN (Stewart, 2007). De plus, Piwi contient un signal de localisation nucléaire bipartite qu'il expose à la protéine Importinα (Impα) lors du chargement des piARN, permettant sa localisation nucléaire (Yashiro et al., 2018). Un mécanisme de régulation épigénétique, dépendant de Piwi, a été identifié chez la drosophile. Il comprend notamment le dépôt de marques répressives de type H3K9me3, pour histone 3 lysine 9 trimethyl, qui conduit au recrutement de la protéine HP1, *heterochromatin protein 1*, au niveau du cluster de piARN ciblé, entraînant une diminution de l'occupation de l'ARN polymérase II (X. A. Huang et al., 2013; Le Thomas et al., 2013). Piwi identifie des cibles complémentaires des piARN et induit une répression transcriptionnelle par des modifications de la structure de la chromatine (Le Thomas et al., 2013). La protéine Maelstrom est également impliquée dans la répression transcriptionnelle médiée par Piwi, chez la drosophile, mais n'est pas impliquée dans le dépôt de marques H3K9me3 (Sienski et al., 2012). D'autres co-facteurs de Piwi sont impliqués dans le contrôle transcriptionnel des ET, en induisant de l'hétérochromatinisation au niveau des locus ciblés, tels que *Panoramix/Silencio*, *Asterix/GTSF1*, et *linker histone H1* (Czech et al., 2013; Dönertas et al., 2013; Handler et al., 2013; Iwasaki et al., 2016; Ohtani et al., 2013; Yu et al., 2015).

b) Boucle d'amplification en ping-pong

La voie de biogenèse secondaire des piARN correspond à la boucle d'amplification des piARN, et est principalement retrouvée au niveau de la lignée germinale. Au niveau des cellules nourricières des ovaires de drosophile, la machinerie impliquant Zuc est combinée avec un mécanisme impliquant les protéines PIWI, ce qui aboutit à ce qu'on appelle le cycle d'amplification en ping-pong. Les piARN transmis maternellement sont pris en charge par les protéines de la famille PIWI, Aub et/ou Piwi et vont être maturés suite au clivage de l'extrémité méthylée à la position 2' O (Kawaoka et al., 2011; Saito et al., 2007). Certains piARN matures sont pris en charge par Piwi et sont transportés dans le noyau où ils vont réprimer transcriptionnellement les transposons (Le Thomas et al., 2013; Sienski et al., 2012). D'autres piARN matures sont pris en charge par Aub et vont ainsi continuer le cycle d'amplification en ping-pong en s'hybridant par complémentarité au niveau de 10 nucléotides au brin d'ARN sens des transposons et en les clivant, générant ainsi des piARN sens (Brennecke et al., 2007;

Gunawardane et al., 2007; Huang et al., 2017). Ces piARN vont ensuite s'associer à la protéine Ago3. Cette dernière cible l'ARN antisens des transcrits par complémentarité de 10 nucléotides en utilisant ces piARN comme guides et en les clivant, permettant la production de piARN antisens qui vont à nouveau pouvoir se lier à Piwi ou Aub et pourront entrer dans la boucle d'amplification (**Figure 25.**) (Czech and Hannon, 2016; Luteijn and Ketting, 2013; Varjak et al., 2018b). Ces piARN secondaires présentent une signature particulière, résultant d'un chevauchement de 10 nucléotides entre les brins sens et antisens (**Figure 26.**) (Antoniewski, 2014; Brennecke et al., 2007). Un biais de séquence pour une uridine à la position 1 (U1) et une adénine à la position 10 (A10) est également spécifique de ces piARN (Huang et al., 2014; C. Li et al., 2009).

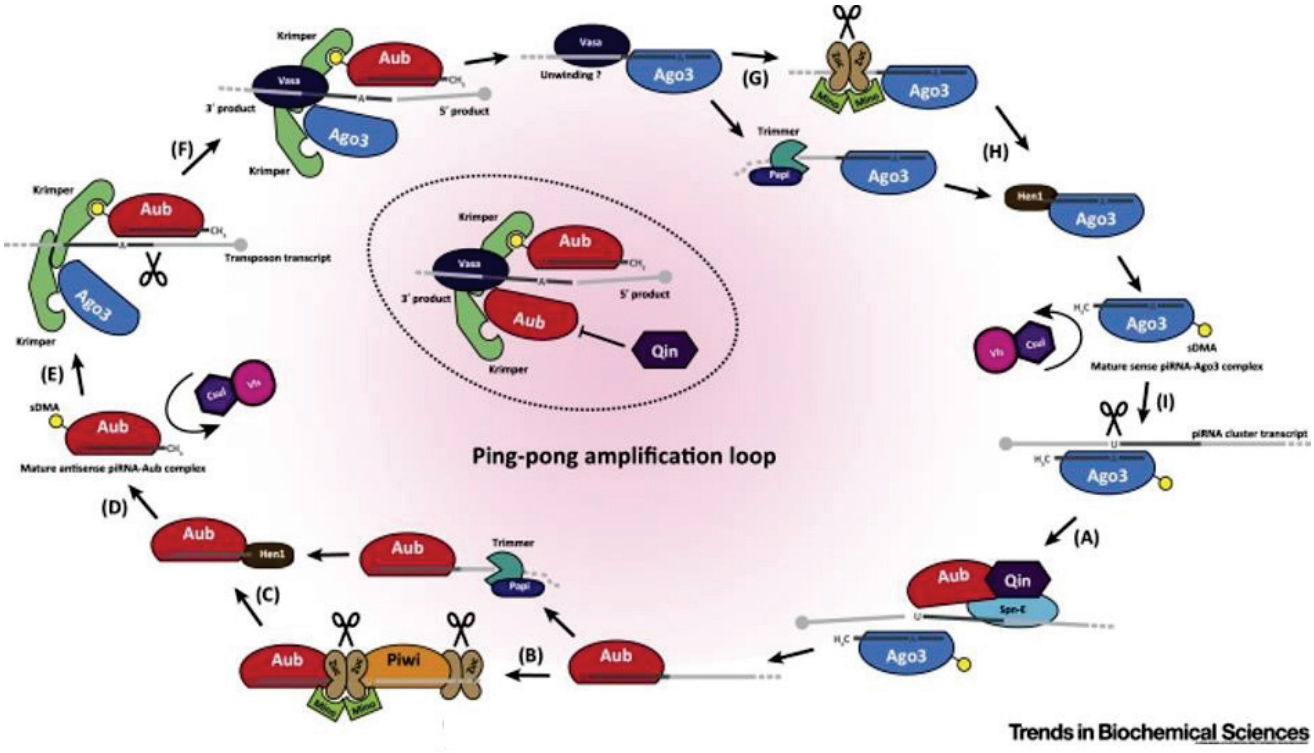


Figure 25. Boucle d'amplification des piARN en ping-pong. (A) Les piARN sens matures sont pris en charge par Ago3 permettant la reconnaissance et le clivage des transcrits de transposons anti-sens provenant des clusters de piARN et induisant la production des intermédiaires de piARN anti-sens avec un uracile en position 5'. Ces derniers sont chargés sur Aub, impliquant la protéine Spindle-E en soutien (Spn-E). **(B)** La maturation des piARN implique notamment le clivage par Zuc, générant une extrémité 3' qui permet une liaison avec la protéine Aub et produisant des piARN qui s'associent à Piwi. La formation de l'extrémité 3' peut également impliquer la protéine Papi. **(C)** La méthyltransférase Hen1 méthyle ensuite l'extrémité 3' de ces piARN anti-sens (2'-O méthylation). **(D)** Les complexes matures de piARN anti-sens liés à Aub sont modifiés à leur extrémité N-terminal par l'ajout d'un motif d'arginine diméthyle symétrique, sDMA. **(E)** Ce complexe est recruté par la protéine Krimper (Krimp) qui à son tour interagit avec une protéine Ago3 libre et rapproche les facteurs impliqués. **(F)** Le complexe piARN-Aub s'hybride par complémentarité partielle de séquence aux transposons ARN sens et les clive, formant ainsi des piARN sens, pris en charge par Ago3 avec l'aide de la protéine Vasa. **(G, H)** De la même façon que pour les complexes piARN-Aub, les piARN chargés sur Ago3 sont maturés. **(I)** Les complexes Ago3-piARN sens clivent à leur tour des transcrits anti-sens, et un autre cyle est engagé. D'après (Czech and Hannon, 2016).

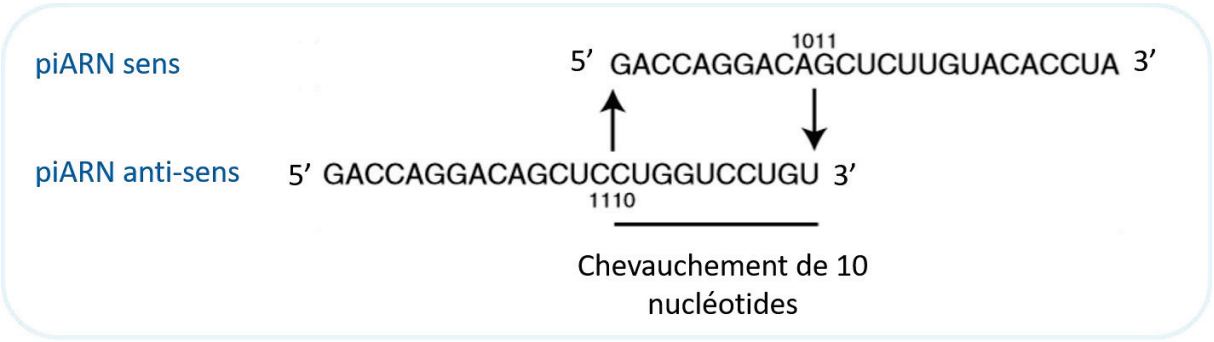


Figure 26. Signature des piARN impliqués dans la boucle d'amplification en ping-pong. Les piARN sens et anti-sens se chevauchent sur 10 nucléotides. Le clivage (représenté par des flèches) de l'extrémité 5' d'un transcrit précurseur long de piARN est effectué par la protéine Piwi, entre les positions 10 et 11 par rapport à l'extrémité 5' du piARN sens complémentaire. D'après (Antoniewski, 2014).

C. piARN somatiques, non gonadiques

Chez la drosophile, en dehors des gonades, Piwi se lie aux chromosomes polytènes de la glande salivaire (Brower-Toland et al., 2007). Récemment, plusieurs travaux ont révélé l'existence d'une voie piARN non gonadique, particulièrement active dans les têtes de *D. melanogaster* déficientes pour la protéine Ago2 (Ghildiyal et al., 2008) et au niveau des têtes sauvages de *D. melanogaster* (Mirkovic-Hösle and Förstemann, 2014; Perrat et al., 2013; Yan et al., 2011). Cependant, il est important de noter que la présence de ces piARN dans les têtes de drosophiles a été depuis remise en question, incluant une étude qui montre que les piARN séquencés par Yan et collaborateurs (2011) et Mirkovic-Hösle and Förstemann (2014) à partir des têtes de drosophiles seraient probablement dus à une contamination des tissus somatiques par les tissus germinaux (Beek et al., 2018). Par ailleurs, il a été montré que les drosophiles portant un transgène qui code une forme mutée de Piwi présentent un défaut d'expression du gène de couleur des yeux *white* (Brower-Toland et al., 2007; Yin and Lin, 2007). De façon plus précise, il a été montré que les protéines PIWI, Ago3 et Aub connues pour contrôler les ET dans la lignée germinale, sont également exprimées au niveau d'une région spécifique du cerveau. La mutation de ces protéines induit une augmentation de l'expression des éléments transposables dans les têtes des drosophiles (Perrat et al., 2013). Par ailleurs, la plupart des insertions d'ET détectées dans les tissus cérébraux des drosophiles ont été remises en question et correspondraient finalement à des chimères techniques produites au moment de la construction des banques de séquençage Illumina (Treiber and Waddell, 2017). Cependant, aucun piARN n'a été détecté dans le thorax de *D. melanogaster* (Lewis et al., 2018). De plus, une voie piARN fonctionnelle somatique, non gonadique, a été décrite au niveau des corps gras chez la drosophile adulte, incluant la présence des effecteurs des piARN, tel que Piwi et les petites molécules d'ARN interférents de 23-30 nucléotides de longueur qui ciblent les ET (Jones et al., 2016). Les drosophiles mutées pour la protéine Piwi ne présentent

plus de piARN dans les corps gras et l'expression des ET est largement augmentée. De plus, l'invalidation de Piwi augmente ainsi les dommages causés à l'ADN et diminue les réserves lipidiques (Jones et al., 2016). Ces résultats indiquent que la voie des piARN somatique, même si elle est marginale chez la drosophile en comparaison avec d'autres arthropodes, est essentielle au bon fonctionnement du métabolisme et à la santé générale d'un organisme.

VI. Interférences entre les piARN et les siARN

Les mécanismes de régulation et la biogenèse de ces deux voies d'ARNi (siARN et piARN) ont été largement étudiés de façon indépendante et ont été présentés dans la partie V (Le mécanisme d'ARN interférence). La régulation des ET par la voie des piARN est bien décrite dans la littérature (Huang et al., 2017; Iwasaki et al., 2015; Ozata et al., 2019; Yamashiro and Siomi, 2018). Cependant, il a été montré que la voie des siARN est également impliquée dans la régulation des ET (Chung et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008). Les clusters génomiques représentant une source pour la synthèse de piARN pourraient être aussi sources de siARN (Chung et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008).

De manière intéressante, plusieurs auteurs suggèrent l'existence d'une éventuelle interaction entre les siARN et les piARN. Chez la drosophile, il a été montré que la délétion de Piwi, au niveau des cellules folliculaires des ovaires, entraînait une augmentation de la production de certains siARN ciblant les ET, permettant de prendre le relais des piARN et d'atténuer l'augmentation du taux de transcrits des ET (Barckmann et al., 2018). D'autres études indiquent une fonction complémentaire entre les siARN et les piARN ciblant les ET au niveau des têtes des drosophiles (Beek et al., 2018). Pour les drosophiles entières mutées pour les séquences codantes des gènes *dcr-2* ou *piwi*, la majorité des familles d'ET ne sont quasiment pas dérégulées alors que chez les doubles mutants *piwi* et *dcr-2*, l'expression de la majorité des ET est augmentée (Beek et al., 2018). Ces résultats suggèrent une fonction en relais des siARN et des piARN dans le contrôle des ET au niveau des tissus somatiques adultes. Ces auteurs suggèrent que la première voie de contrôle des ET est établie par les piARN, alors que les siARN, dépendants de Dcr-2, pourraient être présentés comme une voie de secours, en permettant le contrôle des ET lorsque ces derniers échappent à la première ligne de défense.

1. Le rôle des ET dans l'immunité anti-virale

A. Stimulation de l'immunité antivirale par la voie des siARN

Comme cela a été évoqué partie V.1 (siARN), au temps précoce de l'infection, la transcription inverse de certains fragments de virus à ARN permet la formation d'ADN viral complémentaire à leur génome (ADNv) chez la drosophile (Goic et al., 2013) et chez le moustique (Goic et al., 2016). Il a été suggéré que ces formes d'ADNv soient intégrées dans le génome de l'insecte hôte, au sein de séquences de rétrotransposons (Goic et al., 2013). Les chimères d'ADNv et de rétrotransposons sont transcrits et pris en charge par la machinerie d'ARNi, permettant l'inhibition de la réplication virale en stimulant la production de siARN (Goic et al., 2013). L'inhibition de la transcription inverse bloque la formation d'ADNv et la production de siARN secondaires (Goic et al., 2016, 2013). Ces résultats suggèrent que les ET participent à la persistance et à la tolérance virale chez la drosophile (Goic et al., 2013) et le moustique (Goic et al., 2016). La production d'ADNv a été identifiée lors d'infections de drosophile par les virus SINV, DCV, FHV, et joue un rôle dans l'amplification et la production systémique de siARN antiviraux (**Figure 27.**) (Goic et al., 2013; Poirier et al., 2018). Ces mêmes observations ont été faites lors de l'infection de moustiques *Ae. albopictus* par CHIKV et de moustiques *Ae. aegypti* par DENV (Goic et al., 2016) et ZIKV (Nag and Kramer, 2017), ainsi que *in vitro* lors de l'infection du WNV, DENV, et LACV, dans les lignées d'*Ae. albopictus* C6/36, U4.4 et d'*Ae. aegypti* Aag2 (Nag et al., 2016).

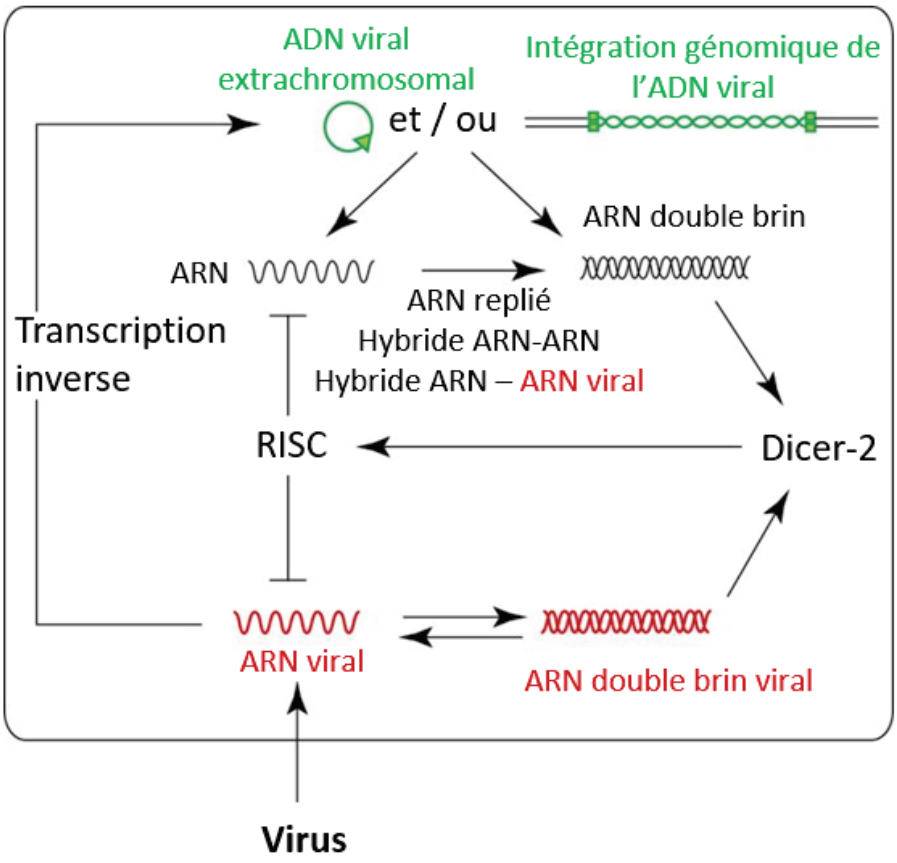


Figure 27. Modèle de la persistance virale chez les insectes. Lors de l'infection virale par un virus ARN, le génome viral se réplique produisant ainsi de l'ARN double brin qui se propage dans l'organisme infecté (rouge). La transcription inverse de ces ARN viraux en ADN (vert) implique une activité de transcriptase inverse cellulaire. L'ADN viral peut s'intégrer dans le génome de l'hôte ou maturer en ADN circulaire extra-chromosomal. Des transcrits ARN et ARN double brin (noir) sont générés à partir des formes ADN virales et peuvent être pris en charge par Dcr-2 et la voie sous-jacente des siARN (impliquant le complexe RISC), qui permettent la production de siARN anti-viraux fonctionnels, favorisant l'immunité anti-virale. D'après (Goic et al., 2013).

B. Induction de l'immunité antivirale par la voie des piARN

Alors que seuls les rétrovirus présentent une étape d'intégration de leur génome dans le génome hôte au cours du cycle viral, des séquences virales de toute nature ont récemment été identifiées comme étant endogénisées. Ces séquences virales endogénisées dans le génome hôte forment des éléments endogènes viraux (EVE) (Aiewsakun and Katzourakis, 2015; Feschotte and Gilbert, 2012) ou NIRVS, pour *Non-Retroviral Integrated RNA Virus Sequences* (Ballinger et al., 2012). Le premier exemple de NIRVS, dans un système animal, tel que la drosophile, inclut la découverte du gène codant la protéine (P), associée à la polymérase du virus sigma, et semble avoir été intégré au rétrotransposon LTR du groupe *blastopia*

(Ballinger et al., 2012). Plus récemment, la présence de séquences de virus ARN non-rétrovirales a été détectée dans le génome du moustique du genre *Aedes* (Palatini et al., 2017) telles que des séquences de Flavivirus et de Rhabdovirus intégrées dans le génome d'*Aedes* et enrichies en clusters de piARN, permettant la production de piARN viraux. Ces séquences d'ADN viral ont été identifiées comme intégrées dans les régions riches en séquences de rétrotransposons à LTR (Chen et al., 2015).

Au cours de l'évolution, des épisodes d'infections répétées de Rhabdovirus chez *Ae. aegypti* semblent être à l'origine de l'accumulation de nombreuses séquences intégrées (Fort et al., 2012). De plus, les séquences d'ADN provenant de Flavivirus (EVE) ont été détectées à la fois à partir de moustiques d'élevage mais aussi de moustiques sauvages, révélant des implications majeures au niveau de l'évolution des génomes de l'hôte (Cook et al., 2006; Crochu et al., 2004). Il a été montré que ces EVE évoluent différemment entre les populations de moustiques en produisant des transcrits et des petits ARN mais pas de protéines, suggérant qu'ils ont un rôle au niveau de l'ARN (Suzuki et al., 2017). Ces piARN dérivés d'EVE pourraient donc expliquer les différences de compétence vectorielle observées chez le moustique (Whitfield et al., 2017). Ces résultats suggèrent une implication de l'activité de transcriptase inverse des rétrotransposons dans le processus d'intégration de l'ADN viral et dans l'induction de l'immunité antivirale sous-jacente par les piARN viraux chez le moustique.

2. Production de piARN viraux et activité antivirale

A. Production de piARN viraux

La voie des piARN a été largement décrite chez la drosophile et sert de matrice pour les études récentes de cette voie chez le moustique. Cependant la voie des piARN semble assez différente entre ces deux organismes (Varjak et al., 2018b). La voie des piARN chez la drosophile est jusqu'à présent basée sur la production de piARN à partir de séquences génomiques et au moins 50% de ces piARN ciblent les ET (Arensburger et al., 2011; Varjak et al., 2018b). Au contraire, malgré un taux d'ET plus important, seulement 19% des piARN ciblent les ET chez *Ae. aegypti* (Arensburger et al., 2011).

De plus, chez le moustique, on trouve des piARN qui sont complémentaires de séquences virales. L'orthologue d'Ago3 est retrouvé chez le moustique contrairement à Aub et Piwi. En effet, il y a plus de gènes codant des protéines PIWI chez les moustiques *Culex*, *Aedes* et *Anopheles* (Lewis et al., 2016). Morazzani et collaborateurs ont montré la production de piARN impliqués dans le cycle d'amplification en ping-pong au niveau des tissus somatiques du moustique (Morazzani et al., 2012). Chez la drosophile, cela semble être plus spécifique aux

tissus germinaux et aux cellules environnantes, même si des piARN réprimant les transposons au niveau du corps gras (Jones et al., 2016) et du cerveau (Ghildiyal et al., 2008; Perrat et al., 2013) ont été décrits chez *D. melanogaster*. Chez *D. melanogaster* et *Ae. aegypti*, les analyses bioinformatiques indiquent que la majorité des piARN dirigés contre les ET révèlent principalement le brin antisens de transcription, incluant le contrôle des ET par des piARN antisens par rapport aux transcrits des ET et mettant en jeu les voies primaire et secondaire des piARN (Arensburger et al., 2011; Brennecke et al., 2007; Kaminker et al., 2002; Vagin et al., 2006). Par ailleurs, la production des piARN par les moustiques semble plus étendue en diversité de cible, incluant un champ d'action également plus important. En plus du contrôle des ET, ces piARN peuvent cibler les ARNm de gènes codants des protéines du moustique ou encore des intermédiaires de réplication virale (cela sera détaillé dans la partie VI.2.B suivante) (Arensburger et al., 2011; Girardi et al., 2017).

Chez la drosophile, la production des piARN viraux est controversée puisque ces derniers ont été identifiés dans les cultures de cellules somatiques ovariennes (Wu et al., 2010), tandis qu'ils n'ont pas été retrouvés dans la drosophile entière, suite à l'infection par SINV, DMelSV, DAV, DXV, NV ou encore DCV (Petit et al., 2016). La production de piARN viraux a été découverte au niveau de cellules somatiques ovariennes (lignée OSS) qui sont infectées de façon persistante par plusieurs virus ARN. Il s'agit de piARN viraux issus de DXV, du virus américain noda, du birnavirus et du tétravirus de drosophile, de DCV et du NV (Wu et al., 2010). Ces piARN viraux ne présentent pas la signature caractéristique des piARN secondaires impliqués dans la boucle d'amplification en ping-pong. De plus, ils ont un fort biais pour l'uridine en position 1 (U1) mais pas le biais pour l'adénine en position 10 (A10). Ils ont une seule polarité sens détectée et les cellules expriment seulement Piwi, ce qui ne permet pas un cycle d'amplification de type ping-pong (C. Li et al., 2009; Malone et al., 2009). Jusqu'à présent, la production de piARN viraux n'a pas été décrite dans d'autres tissus de drosophile (Lewis et al., 2018).

Aucune activité antivirale n'a été décrite pour les piARN viraux chez *D. melanogaster* infectées par DCV, DXV et SINV (Petit et al., 2016). Néanmoins, d'autres auteurs montrent une augmentation de la sensibilité des drosophiles suite à l'infection par DXV pour les mutants *aub* et *piwi* par rapport aux drosophiles sauvages (Zambon et al., 2006). Il semblerait également que le même phénotype soit retrouvé pour des mutants *piwi* infectés par WNV (Chotkowski et al., 2008). Ces derniers résultats sont discutables du fait de l'absence d'information concernant la production de piARN viraux et l'utilisation de drosophiles ayant des fonds génétiques différents (Petit et al., 2016).

B. piARN viraux chez le moustique

Comme mentionné dans la partie V.3.C (piARN somatiques, non gonadiques), la production de piARN somatiques est répandue chez les moustiques, tel que chez *Ae. aegypti*, où ces piARN ciblent les virus (Miesen et al., 2015; Morazzani et al., 2012). *Ae. aegypti* code Ago3 et sept autres protéines de la famille PIWI. Les transcrits Piwi4-6 et Ago3 sont fréquemment détectés dans les cellules Aag2 et en somatique chez le moustique, tandis que les protéines Piwi1-3 sont spécifiques de la lignée germinale (Akbari et al., 2013; Wu et al., 2010). A l'inverse, l'expression des protéines PIWI et leur localisation ne sont pas connues chez *Ae. albopictus*. Cette espèce code deux protéines Ago3 et neuf autres protéines PIWI (Piwi1-9). Seulement Piwi5-7 et les deux protéines Ago3 ont été détectées dans l'intestin moyen des moustiques femelles. Les transcrits des protéines Piwi1-4 ont été identifiés chez le mâle adulte, chez la femelle après le repas sanguin et au niveau de l'embryon. Les transcrits des protéines Piwi5-9 semblent seulement exprimés dans l'embryon (Wang et al., 2018). De nombreux arbovirus produisent des piARN viraux lors de l'infection virale, notamment par les virus DENV (Wang et al., 2018), RVFV (Dietrich et al., 2017), ZIKV (Saldaña et al., 2017), CHIKV (Morazzani et al., 2012) chez leur hôte arthropode. On retrouve la production de ces piARN viraux dans les lignées cellulaires dérivées de ces arthropodes, pour l'ensemble des virus précédemment cités. De plus, à partir de cellules de moucherons *Culicoides*, vecteurs du virus Bluetongue (BTV) et du virus Schmallenberg (SBV), des piARN viraux spécifiques de ces deux virus ont été détectés lors de l'infection (Schnettler et al., 2013b). Dans le cas de l'infection par SINV, des piARN viraux sont produits à partir de différentes lignées cellulaires d'*Aedes* (Aag2, U4.4 et C6/36) (Brackney et al., 2010; Miesen et al., 2015; Vodovar et al., 2012).

Lors de l'étude d'arthropodes non infectés, le génome partiel d'un virus à ARN simple brin de polarité positive, nommé Aedes Anphevirus, a été détecté lors du séquençage du transcriptome d'*Ae. aegypti* et est ciblé à la fois par les siARN, et les piARN présentant la signature caractéristique de l'amplification en ping-pong (Lewis et al., 2018). Chez d'autres arthropodes, tels que *Centruroides sculpturatus*, présentant une faible proportion de piARN viraux ou *Tribolium castaneum*, ne révélant qu'un seul des deux brins des piARN classiques lors de l'analyse bioinformatique, aucune signature ping-pong n'a ainsi été révélée (Lewis et al., 2018). Toutefois, des siARN viraux ont été identifiés chez ces arthropodes. L'ensemble de ces résultats met en lumière des virus ciblés par les siARN et les piARN somatiques, même si la voie des siARN reste probablement la première ligne de défense antivirale chez les arthropodes.

Les EVE ou NIRVS, qui sont des séquences provenant de virus non rétroviraux, insérées dans le génome du moustique, servent de matrice pour la production de ces piARN viraux (Katzourakis and Gifford, 2010; Miesen et al., 2016b). La délétion de Piwi5 en cellules Aag2 engendre une diminution de la production de piARN dérivés de ces EVE, ce qui montre que ces derniers dépendent principalement de Piwi5 chez le moustique *Aedes* (Palatini et al.,

2017). Les protéines Ago3, Piwi4 et Piwi6 sont également impliquées mais dans une moindre mesure (Palatini et al., 2017). La même observation est faite pour la production de piARN à partir des virus SINV (Miesen et al., 2015), SFV (Schnettler et al., 2013a), DENV (Miesen et al., 2016a), et CHIKV (Varjak et al., 2018a). A l'inverse, la production de piARN à partir des ET implique un plus large panel de protéines PIWI, incluant des protéines PIWI qui ne contribuent pas à la production des piARN viraux. Ces résultats indiquent une spécialisation fonctionnelle des différentes protéines PIWI de moustique dans le traitement de l'ARN endogène ou exogène.

Les piARN dérivés des NIRVS présentent généralement une orientation anti-sens avec une uridine en position 1, et peuvent ainsi cibler directement l'ARNm viral. On trouve un biais correspondant à une adénine en position 10 pour les piARN sens particulièrement chez *Aedes albopictus*, indiquant la production de piARN dérivés de NIRVS lors de la boucle d'amplification en ping-pong. Cette production de piARN secondaires est pratiquement exclusive pour les NIRVS dérivés de Rhabdovirus (Palatini et al., 2017; Whitfield et al., 2017). La production de nouveaux piARN peut être induite après transformation des moustiques avec des séquences d'ET (système rapporteur). Cela est en relation avec l'hypothèse mentionnant une intégration d'ADN viraux dans le génome du moustique au niveau des clusters de piARN qui entraîne une production de piARN viraux (**Figure 28.**) (Palatini et al., 2017).

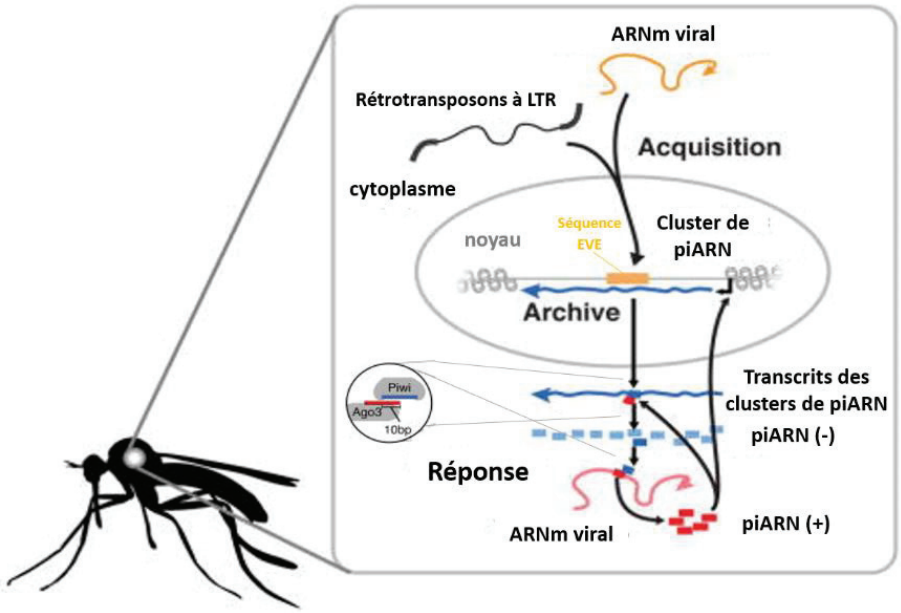


Figure 28. Schéma de l'origine de la production de piARN viraux chez le moustique. La séquence d'un EVE au niveau d'un cluster de piARN est transcrite en ARN anti-génomique puis maturée en piARN. Ces piARN ciblent le génome ou l'ARNm de virus. Dans les deux cas, les piARN sens se lient aux transcrits anti-génomiques par l'intermédiaire d'Ago3 et les piARN anti-sens se lient aux transcrits génomiques avec la protéine Piwi pouvant ainsi générer un biais à la position 10 sur une adénine (A10), et ainsi un chevauchement de 10 nucléotides entre les piARN sens et anti-sens. D'après (Whitfield et al., 2017).

De nombreuses études s'intéressent à l'implication des piARN viraux au niveau des réponses de l'immunité antivirale chez le moustique, classiquement prises en charge par la voie des siARN (Hess et al., 2011; Keene et al., 2004; Schnettler et al., 2013a; Vodovar et al., 2012; Wu et al., 2010). Chez le moustique, principalement du genre *Aedes*, plusieurs membres de la famille des *Togaviridae* (Goic et al., 2016; Miesen et al., 2015; Morazzani et al., 2012; Schnettler et al., 2013a; Vodovar et al., 2012), des *Flaviviridae* (Miesen et al., 2016a; Scott et al., 2010; Varjak et al., 2017), *Phenuiviridae* (Dietrich et al., 2017; Léger et al., 2013) et des *Reoviridae* (Schnettler et al., 2013b) engendrent la production de piARN viraux présentant une activité antivirale (Miesen et al., 2016b; Varjak et al., 2018b). Ce sont des piARN de taille classique entre 23 et 30 nucléotides et présentant une signature ping-pong caractéristique (biais U1 et/ou A10) pour les alphavirus et les phlébovirus, contrairement aux flavivirus qui n'en présentent pas. Pour les virus DENV et SINV, ces piARN viraux sont méthylés à l'extrémité 3' en position O-2' (Miesen et al., 2015, 2016a).

Dans les cellules d'*Ae. aegypti* Aag2 invalidées pour le gène codant la protéine Piwi4 et infectées par le virus de la forêt de Semliki (SFV), la réplication virale est plus importante que dans les cellules contrôles alors que contrairement aux protéines Ago3 et Piwi5, la protéine Piwi4 n'est pas impliquée dans la production de piARN viraux (Schnettler et al., 2013b; Varjak et al., 2017). De plus, il a été montré comme attendu, que dans les cellules Aag2, l'invalidation du gène codant la protéine Dcr2, un des principaux effecteurs de la voie des siARN, révèle une perte de l'activité antivirale, lors de l'infection de ces cellules par le SFV. En l'absence d'une voie siARN fonctionnelle, Piwi4 maintient son activité antivirale, ce qui suggère que l'activité antivirale de Piwi4 est indépendante de la voie des siARN chez le moustique lors de l'infection par SFV (Schnettler et al., 2013a). Par ailleurs, Varjak et collaborateurs ont décrit une association de la protéine Piwi4 avec des protéines membres des voies siARN (Dcr2 et Ago2) et piARN (Ago3, Piwi5 et Piwi 6) (Varjak et al., 2017). Leurs résultats démontrent une interaction complexe entre la voie des siARN et celle des piARN chez *Ae. aegypti.*

On retrouve cette augmentation de réplication virale pour RVFV lors de la délétion d'Ago3 et Piwi4 (Dietrich et al., 2017), et pour ZIKV lors de la délétion de Piwi4 (Varjak et al., 2017), en cellules Aag2, ce qui implique que la voie des piARN a une activité antivirale chez *Ae. aegypti*. Piwi5 et Ago3 semblent également être les seuls membres de la famille PIWI à être essentiels pour la production de piARN viraux lors de l'infection par les virus SINV (Miesen et al., 2015), DENV (Miesen et al., 2016a) et CHIKV (Varjak et al., 2018a) en cellules Aag2. Des piARN viraux sont également produits chez *Anopheles gambiae* lors de l'infection par *O'nyong nyong virus* (ONNV) et la délétion d'Ago3 induit une légère augmentation de la réplication virale (Keene et al., 2004). L'ensemble de ces résultats suggère que l'activité antivirale des piARN est spécifique des piARN viraux produits chez les Arthropodes, et est principalement caractérisée chez le moustique *Aedes*.

MATERIEL ET METHODES

MATERIEL ET METHODES

I. Souches de drosophiles

Nous avons utilisé la lignée Makindu, une souche de *D. simulans*, décrite précédemment par (Fablet et al., 2006). Il s'agit d'une lignée isofemelle fondée à partir d'un échantillonnage d'une population naturelle du Kenya. Les souches de *D. melanogaster* w^{1118} et les mutants *Dcr-2*$^{[L811fsX]}$ (Tomari et al., 2007) proviennent du laboratoire de M-C Saleh (Institut Pasteur Paris). Les mutants de l'import d'ARN double brin, *CG4572*^{c05963}, ont été obtenus à partir du centre de stock de Drosophiles Bloomington (Bloomington *Drosophila* Stock Center). Les drosophiles ont été maintenues à 25°C sur du milieu qui a été autoclavé et contenant pour un litre d'eau osmosée 8,5 g d'agar ; 80 g de farine de maïs ; 80 g de levure morte ; 5,2 g d'antifongique parahydrobenzoate de méthyle (Moldex) et 4 ml d'acide propionique. Conformément aux études précédentes, l'ensemble des expérimentations a été effectué en utilisant des drosophiles femelles âgées de 3 à 6 jours (Goic et al., 2013; Saleh et al., 2009).

II. Validation du génotype chez *D. melanogaster Dcr-2*$^{[L811fsX]}$

Nous avons utilisé les mutants *D. melanogaster Dcr-2*$^{[L811fsX]}$ présentant le génotype suivant : *y[d2] w[1118], P{ry+t7.2 = ey–FLP.N}2 ; Dcr-2[L811fX] (Tomari et al., 2007)*. Ces mutants homozygotes *Dcr-2*$^{[L811fsX]}$ présentent une mutation au niveau de la leucine 811, sur un fond génétique w^{1118}. Suite à un décalage du cadre de lecture, la leucine 811 est remplacée par 19 nouveaux acides aminés, ce qui introduit prématurément un codon stop au niveau du gène codant la protéine Dcr-2. Par conséquent, Dcr-2 n'est plus fonctionnelle dans ces mutants. Ces derniers ne sont plus capables de produire des siARN et de détecter spécifiquement l'ARNm à cliver.

Nous avons extrait l'ADN des mutants *Dcr-2*$^{[L811fsX]}$ et des drosophiles w^{1118}, puis séquencé l'ADN (séquençage par la société GATC), permettant la validation du génotype par rapport au génotype sauvage w^{1118}.

III. Validation du génotype chez *D. melanogaster CG4572^{c05963}*

Nous avons utilisé les mutants *D. melanogaster CG4572^{c05963}* présentant le génotype suivant : *w[1118]; PBac{w[+mC]=PB}CG4572[c05963]* (*Bloomington Drosophila Stock Center*). La fonction du gène *CG4572* (*Coding Gene 4572*) reste encore inconnue. Cependant, il a été précédemment identifié lors du criblage à haut débit du génome de cellules de *Drosophila*, que la protéine CG4572/DORA est impliquée dans la machinerie de l'import d'ARN double brin (Saleh et al., 2009, 2006). Les mutants homozygotes *CG4572^{c05963}* présentent une insertion d'un transposon *PiggyBac*, contenant le marqueur *white*[+] (w[+mC]) au sein du cadre de lecture ouvert du gène *CG4572* pour chaque allèle. Ces mutants homozygotes pour le gène *CG4572* sont viables mais peuvent être peu fertiles. Par conséquent, le balanceur TM6B a été ajouté et est porté par quelques drosophiles, permettant son maintien et une fertilité plus importante. Une sélection du phénotype correspondant aux drosophiles homozygotes pour la mutation (*CG4572^{c05963}*) est nécessaire. Ces insertions restaurent le niveau de pigmentation des yeux. Ainsi, nous avons sélectionné les individus homozygotes pour l'insertion, présentant des yeux fortement colorés (orange foncé) et ne présentant plus de soie au niveau des épaules du thorax. Les individus aux yeux clairs sont hétérozygotes pour l'insertion car ils portent le chromosome balanceur TM6B, Tb[1]. La présence du marqueur dominant Tb est caractérisée par des larves et des pupes très courtes et ramassées.

En parallèle, nous avons extrait l'ADN des mutants *CG4572^{c05963}* et des drosophiles sauvages *w^{1118}*, et effectué une PCR permettant la visualisation du gène *CG4572*. Nous avons utilisé les amorces F 5' GTGGACCCTGGCTTCAAATC 3' et R 5' GATAATTGCGCGTCAGTGGA 3' pour l'amplification de ce gène, incluant l'insertion du transposon *PiggyBac* dans le cas où il est présent.

IV. Détection et élimination de *Wolbachia*

Les *Wolbachia* sont des bactéries endosymbiotiques présentes chez une grande partie des arthropodes terrestres (Hertig and Wolbach, 1924). Depuis 2008, on sait que ces bactéries ont un effet antiviral chez *D. melanogaster*, notamment lors de l'infection par SINV (Bhattacharya et al., 2017), DCV, CrPV, FHV ou NV (Hedges et al., 2008; Teixeira et al., 2008).

Il est nécessaire de s'affranchir de ces bactéries dans le cadre de l'étude des infections virales et des voies d'ARNi. Les bactéries *Wolbachia* ont été éliminées pour l'ensemble des drosophiles, en traitant ces dernières à la tétracycline pendant deux générations. Comme décrit dans le protocole de (Merkling and van Rij, 2015), les drosophiles ont été transférées

dans un milieu standard supplémenté avec 0,25 mg/ml de tétracycline (Sigma Aldrich), en les laissant pondre pendant trois jours. La génération suivante (G2) a été placée dans un nouveau tube supplémenté de tétracycline pendant trois jours. Les drosophiles G3 ont été déposées sur un milieu sans tétracycline et maintenues pendant au moins trois générations avant le début des expérimentations.

L'élimination de *Wolbachia* a été validée par PCR, après extraction individuelle d'ADN et en utilisant les amorces spécifiques de *Wolbachia pipientis* (16S) (Holden et al., 1993) ou du gène *wsp* (codant pour une protéine de surface retrouvée chez un grand nombre d'espèces de *Wolbachia*) (Riegler et al., 2005) (**Tableau 3.**). La PCR a été réalisée en utilisant le kit de PCR Go Taq (Promega). Le gène de ménage de drosophile *rp49* a été utilisé comme témoin d'ADN (Akkouche et al., 2012).

V. Détection et élimination des virus des drosophiles

Les drosophiles de laboratoire sont couramment infectées par des virus persistants, tels que DCV, DAV ou NV (Habayeb et al., 2006; Kapun et al., 2010; Webster et al., 2015). Ces virus induisent les voies d'ARNi de l'hôte, impliquées notamment dans la tolérance virale des drosophiles lors de leur réplication (Bronkhorst et al., 2012; Galiana-Arnoux et al., 2006; van Rij et al., 2006; Wang et al., 2006). Certains d'entre eux ont développé une stratégie d'échappement à la voie d'ARNi en induisant des suppresseurs d'ARNi, et activent de nombreuses voies cellulaires de l'hôte (Marques and Imler, 2016; Merkling and van Rij, 2013). Des mécanismes de tolérance permettent de réduire l'impact de l'infection sur la physiologie de la drosophile en atténuant les réponses immunitaires excessives ou en diminuant les lésions tissulaires (Ayres and Schneider, 2012; Medzhitov et al., 2012). Lors de l'étude des réponses d'un virus particulier, suite à son inoculation expérimentale, ces infections persistantes pourraient avoir un impact et perturber l'analyse des résultats. L'objectif de cette étape est d'éliminer les infections persistantes des drosophiles en traitant les œufs avec de l'eau de javel domestique, comme décrit par Merkling and van Rij, (2015) et en suivant les étapes principales suivantes :

- Préparer des lames contenant du milieu gélosé, vinaigré et sucré puis ajouter de la levure morte pour augmenter le taux de reproduction des drosophiles. Déposer la lame au fond d'une cage adaptée et mettre environ 50 mouches âgées de 3 à 7 jours dans la cage, puis incuber à 25°C pendant 3 à 6 heures afin de laisser les drosophiles pondre des œufs sans formation de larves.
- Fabriquer un récipient pour les œufs, à l'aide d'un Falcon 50 ml coupé et permettant d'accueillir un filtre nylon.
- Collecter les œufs avec un pinceau stérile pour chaque lignée, permettant de s'affranchir de toute contamination éventuelle ou de confusion de génotype. Disposer

les œufs dans le récipient prévu à cet effet et les laver une première fois avec de l'eau déminéralisée.

- Javelliser les œufs avec une solution de javel domestique à 2,6% de chlore actif, diluée au demi dans de l'eau (50/50) pendant 10 minutes, puis rincer trois fois à l'eau douce. Attention à ne pas laisser incuber trop longtemps les œufs dans la solution de javel car cela peut impacter fortement la viabilité des embryons.
- Récolter les œufs déchorionnés du filtre nylon et les disposer sur un petit morceau de papier filtre.
- Placer le morceau de papier filtre contenant les œufs sur le milieu de drosophile standard et incuber à 25°C pendant environ 10 jours, pour l'émergence des adultes.

L'absence de virus a ensuite été validée par RT-PCR, à partir des adultes émergeants. L'ARN a été extrait à partir de 10 mouches, préalablement incubées à -20°C, pendant 15 minutes et en utilisant une méthode d'extraction au Trizol (Thermo Fisher Scientific) couplée au kit RNeasy Mini (Qiagen), selon les recommandations du fournisseur. La transcription inverse d'1 µg d'ARN a été effectuée à l'aide du kit Omniscript (Qiagen) associé à la méthode d'amorçage des ARNm par des oligodT (Invitrogen). La PCR a été réalisée en utilisant le kit de PCR Go Taq (Promega). Les amorces spécifiques utilisées sont référencées dans le (**Tableau 3.**) (Merkling and van Rij, 2015). Les gènes de ménage *rp49* et *Act42A* ont été utilisés comme témoins d'extraction d'ARN et d'amplification de l'ADN complémentaire pour chaque échantillon. L'élimination de ces infections persistantes a été effectuée à partir des lignées de *D. melanogaster* (w^{1118}, $Dcr\text{-}2^{[L811fsX]}$ et $CG4572^{c05963}$) pour l'ensemble des expérimentations et également pour des expériences complémentaires de RT-PCR quantitative, réalisées à partir des drosophiles *D. simulans* Makindu.

Cible à détecter	Séquences des amorces (5'-3')	Taille du produit attendu (pb)
DCV	AAAATTTCGTTTTAGCCCAGAA	250
(Merkling and van Rij, 2015)	TTGGTTGTACGTCAAAATCTGAG	
DAV	AGGAGTTGGTGAGGACAGCCCA	146
(Merkling and van Rij, 2015)	AGACCTCAGTTGGCAGTTCGCC	
NV	ATGGCGCCAGTTAGTGCAGACCT	410
(Merkling and van Rij, 2015)	CCTGTTGTTCCAGTTGGGTTCGA	
Wolbachia wsp	TGGTCCAATAAGTGATGAAGAAAC	610
(Riegler et al., 2005)	AAAAATTAAACGCTACTCCA	
Wolbachia 16S	TTGTAGCCTGCTATGGTATAACT	900
(Holden et al., 1993)	GAATAGGTATGATTTTCATGT	
Act42A	GCGTCGGTCAATTCAATCTT	522
(Merkling and van Rij, 2015)	CTTCTCCATGTCGTCCCAGT	
rp49	GCGCTTGTTCGATCCGTA	188
(Akkouche et al., 2012)	CGGATCGATATGCTAAGCTGT	

Tableau 3. Amorces utilisées pour la détection des infections par les virus persistants et par la bactérie endosymbiotique *Wolbachia* dans les lignées de drosophiles. *Act42A* et *rp49* sont des gènes de ménage, utilisés comme témoins de RT-PCR.

VI. Production du virus Sindbis par génétique inverse

SINV a été produit par génétique inverse en utilisant des cellules de mammifère, BHK-21 (lignée cellulaire de fibroblastes de rein de hamster) selon les étapes suivantes :

1. Préparation du plasmide pour la transcription *in vitro*

Le plasmide correspondant au génome de la souche virale SINV TOTO 1101, traitement 6, population 144 nous a été gracieusement donné par l'équipe « Transmission des alphavirus et pathologie comparée » de l'IVPC. Précédemment, une monocouche de cellules BHK (*Baby hamster kidney*) avait été infectée avec le virus ancestral (Rice et al., 1987), puis environ 10^4 particules virales infectieuses ont été transférées sur des cellules fraîches incluant 25 passages au total et permettant la production du traitement 6 de la souche TOTO 1101 (Morley et al., 2015). Nous avons ensuite amplifié le plasmide de SINV en transformant des bactéries TOP10 compétentes (One Shot™ TOP10 Chemically Competent E. coli, Thermo Fisher). A partir d'un bouillon de culture, nous avons effectué une « maxi préparation », en utilisant le kit NucleoBond® Xtra Midi / Maxi (Macherey-Nagel), selon les recommandations du fournisseur. Le plasmide a été linéarisé avec l'enzyme de restriction XhoI puis purifié à l'aide du kit PCR Clean up (Macherey-Nagel).

2. Transcription *in vitro*

Le plasmide SINV TOTO 1101 a été transcrit *in vitro* en utilisant le kit mMESSAGE mMACHINE, High Yield Capped RNA transcription Kit, SP6 (Life technologies) selon les instructions du fournisseur. L'ensemble du matériel doit être dépourvu de RNases et de DNases pour éviter toute contamination ou perte de matériel biologique. L'ARN a été extrait par la méthode au phénol-chloroforme (Sigma-Aldrich) et précipité à l'isopropanol (Thermo Fisher). Enfin, l'ARN a été visualisé sur un gel agarose 1%.

3. Electroporation des cellules BHK-21

Les cellules BHK-21 ont été ensemencées au préalable dans du DMEM (Dulbecco's Modified Eagle Medium), supplémenté de 10% de sérum de veau fœtal (SVF) et 1% de pénicilline-streptomycine (PS). Le jour de l'expérience (jour 0), les cellules ont été comptées et 5.10^6

cellules ont été utilisées pour l'électroporation. Après avoir centrifugé les cellules 5 minutes à 500 g à 25°C, le surnageant a été éliminé et les cellules ont été lavées en ajoutant 5 ml de PBS (*Phosphate Buffered Saline*). L'opération est réitérée en resuspendant les cellules, cette fois-ci dans 400 µl d'Opti-MEM. Une flasque de 75 cm^2 contenant du DMEM supplémenté de 10% de SVF est nécessaire pour la suite de l'expérimentation.

L'électroporateur, Gene Pulser Xcell™ Electroporation System (Biorad), a été réglé au préalable à 270 V, avec une capacité de 975 µF et un temps de décharge entre 18 et 25 ms. Une fois l'ARN décongelé, 10 µg ont été déposés dans la cuvette d'électroporation puis rapidement mais délicatement supplémentés des 400 µl des cellules en suspension dans l'Opti-MEM pour effectuer l'électroporation (**Figure 29.**). Cette étape permet l'entrée de l'ARN de SINV dans les cellules BHK-21, qui pourra ensuite se répliquer et former des particules virales infectieuses. L'ARN a été resuspendu dans la cuve, puis déposé dans la flasque de 75 cm^2 prévue à cet effet et incubée à 37°C dans une étuve à 5% de CO_2. Après 16h d'incubation (jour 1), le milieu de culture a été remplacé par un milieu de culture frais identique. Environ 24h après (jour 2), la visualisation des effets cytopathiques au microscope indique la présence de particules virales infectieuses de SINV. Le milieu de culture infectieux a été récolté dans un Falcon 50 ml et centrifugé 5 minutes à 500 g. Le surnageant a été déposé dans un nouveau Falcon, aliquoté dans des tubes à vis et stocké à -80°C. La production des stocks contrôle, sans production de particules virales (mock) a été réalisée dans les mêmes conditions. Pour cela, les cellules ont été électroporées sans ARN viral, permettant de récolter l'ensemble des constituants libérés par les cellules lors de l'électroporation mais sans aucune production de particules virales. Chaque aliquot a été décongelé une seule fois puisque les épisodes de congélation/décongélation altèrent le titre viral.

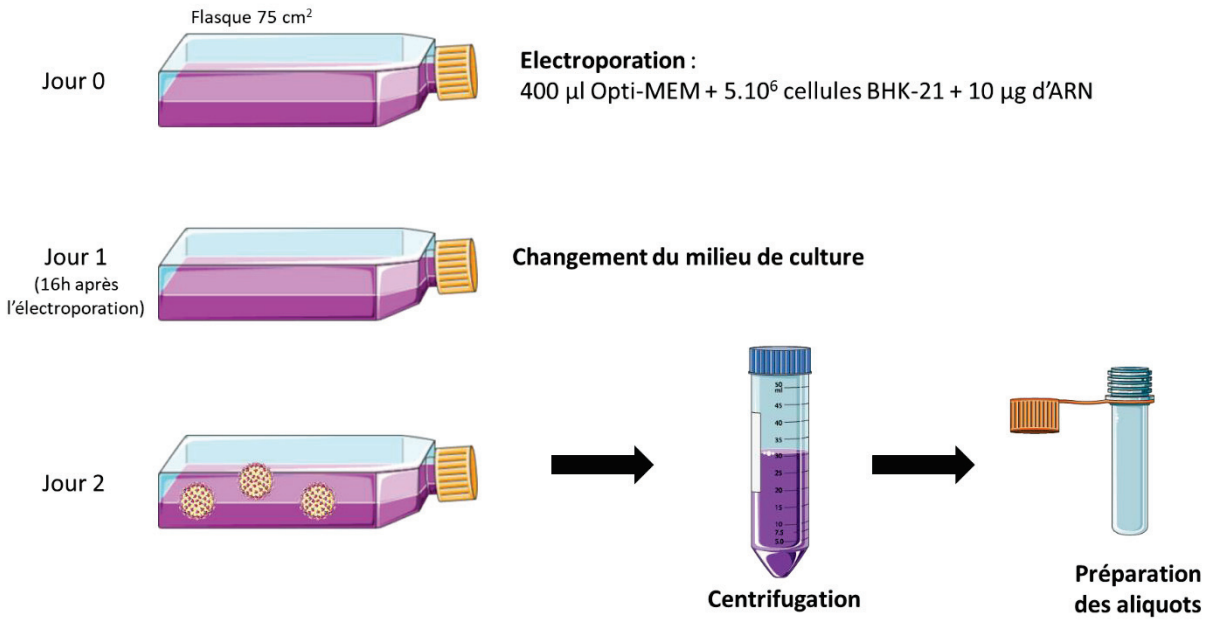


Figure 29. Production de SINV par électroporation des cellules BHK-21. Au jour 0, les cellules BHK-21 ont été électroporées, permettant l'entrée de l'ARN de SINV. Au jour 1, le milieu de culture a été remplacé par un milieu

neuf (DMEM + 10% SVF). Au jour 2, les effets cytopathiques sont visibles au microscope et le milieu infectieux est centrifugé pour ensuite être aliquoté.

VII. Titration des stocks viraux de SINV

L'ensemble des expérimentations a été réalisé à partir de trois stocks viraux de SINV. Les titres viraux de chaque stock ont été déterminés par la méthode de référence de titration des particules virales infectieuses en plage de lyse, nommé PFU pour *plaque-forming unit*, en utilisant des cellules de mammifère permissives Vero E6, et sont calculés en PFU pour 1 ml. Cette lignée cellulaire a été isolée à partir de cellules épithéliales de rein extraites d'un singe vert africain (*Chlorocebus*) et est déficiente en interféron. Ces stocks ont été dilués dans du DMEM permettant d'équilibrer les titres viraux de SINV à $7{,}2.10^7$ PFU/ml. L'ensemble des titres viraux ont été confirmés par une méthode de titration virale basée sur la détermination du seuil, où 50% des cellules de mammifère infectées meurent, connus sous le nom de dose infectieuse pour culture sur tissu à 50% (TCID50, *50% tissue culture infectious dose*). Cette méthode nécessite des dilutions en série par un facteur 10 du stock viral à quantifier et la mise en culture de cellules permissives, telles que les cellules Vero E6 et est basée sur la quantification des puits présentant des effets cytopathiques. Le titre viral est déterminé en TCID50/ml selon la méthode de Reed et Muench's (Reed and Muench, 1938). Le stock contrôle « mock » a été validé dans les mêmes conditions. n = 3 titrations indépendantes (pour chaque méthode de titration virale et chaque stock viral).

VIII. Production du virus sigma (DMelSV)

Le laboratoire de Francis Jiggins (Université de Cambridge) nous a gracieusement donné la lignée *D. melanogaster* naturellement infectée de façon persistante par le virus sigma Hap23 (Carpenter et al., 2007). Cette lignée de drosophile a été obtenue par ces auteurs en croisant, pendant plus de 6 générations, 5 populations de *D. melanogaster* naturellement infectées par le virus sigma (Carpenter et al., 2007; Gay, 1978). Nous avons ensuite extrait le virus à partir de ces drosophiles chroniquement infectées. Pour cela, 100 drosophiles chroniquement infectées ont été broyées dans du PBS supplémenté de 10% de SVF. Après centrifugation pendant 2 minutes à 10000 xg, le surnageant a été filtré à 0,45 µm puis aliquoté pour un stockage à -80°C. Les stocks contrôles (mock), dépourvus de particules virales, sont constitués de PBS supplémenté de 10% de SFV et aliquotés de la même façon que précédemment.

IX. Titration du stock viral de DMelSV

Le virus DMelSV n'induit pas d'effets cytopathiques, ce qui implique qu'il ne peut pas être quantifié par des méthodes classiques de titration virale, impliquant classiquement une quantification des particules virales après culture en cellules permissives et l'analyse des effets cytopathiques. Par ailleurs, aucun anticorps (Ac) n'a été décrit pour l'analyse de DMelSV en FFU (*fluorescent focus units*), mais l'infection par DMelSV peut être confirmée par immunofluorescence en utilisant de l'antisérum anti-DMelSV (Tsai et al., 2008). Par ailleurs, il a été décrit que l'infection par DMelSV cause une paralysie de la drosophile infectée, puis la mort de la drosophile lorsque celle-ci est exposée au dioxyde de carbone (CO_2) (L'Heritier, 1948). A l'inverse, les drosophiles non infectées et exposées au CO_2 se réveillent normalement. Ainsi, le test d'exposition au CO_2 est la méthode de choix pour quantifier DMelSV (Longdon et al., 2012). Nous avons déterminé le nombre d'unités infectieuses du stock de DMelSV en réalisant une série de dilutions de DMelSV (pure, 10^{-1}, 10^{-2}, 10^{-3} et 10^{-4} *versus* condition contrôle « mock ») et infecté expérimentalement par injection intrathoracique, dans un volume de 207 nl, les mouches *D. melanogaster* naïves (L'Heritier, 1957; Longdon et al., 2012). Les drosophiles ont été maintenues sur milieu classique, contenant de l'agar, de la farine de maïs et de la levure morte et incubées à 25°C. Après 14 jours d'infection, les drosophiles ont été soumises à 100% de CO_2 puis déposées sur la glace pendant 15 minutes. Suite au retour à 25°C, les drosophiles infectées par DMelSV sont paralysées de façon irréversible et succombent à l'infection. Nous avons quantifié les drosophiles survivantes 24h après le test CO2 et en se basant sur la méthode de Reed et Muench (Reed and Muench, 1938), le titre viral de DMelSV a été estimé en dose infectieuse par drosophile (DID50, *50% Drosophila infectious dose*).

X. Infection des drosophiles

Les drosophiles ont été maintenues sur un milieu autoclavé contenant principalement de l'agar, de la farine de maïs et de la levure morte (plateforme Arthro-tools) et incubées à 25°C. Dans chaque tube de milieu, 15 femelles et 10 mâles (âgés de 3 à 6 jours) ont été incubés à 25°C pendant 3 jours, permettant aux femelles de pondre. Environ 10 jours après, la nouvelle génération a émergé et les drosophiles femelles âgées de 3 à 6 jours ont été infectées expérimentalement par injection intrathoracique à l'aide de l'appareil Nanoject II (Drummond Scientific). Les mouches *D. melanogaster* w^{1118} ont été utilisées comme contrôle sauvage par rapport aux drosophiles mutantes $Dcr2^{[L811fsX]}$ et $CG4572^{c05963}$ présentant le même fond génétique. Ces drosophiles ont été anesthésiées au CO_2 puis infectées de façon intrathoracique avec 2300 particules infectieuses de SINV (PFU) dans un volume de 32,2 nl, pendant 6 jours avant de passer à la suite des expérimentations. En parallèle, nous avons

injecté de la même façon, 32,2 nl du stock « mock » pour les drosophiles contrôles non infectées par SINV. Les drosophiles Makindu (*D. simulans*) ont été anesthésiées au CO_2, puis infectées après injection intrathoracique avec $1{,}44.10^4$ particules infectieuses de SINV (PFU), pendant 6 jours ; ou avec une dose infectieuse de DMelSV contenue dans un volume de 207 nl, pendant 8 jours, et comparées à la condition « mock » réalisée dans les mêmes conditions. La mortalité des drosophiles au jour de l'injection a été attribuée aux dommages induits par l'injection et n'est pas prise en compte pour la suite des expérimentations. Les drosophiles ont été transférées dans des milieux neufs tous les 2-3 jours. Chaque expérience d'injection virale de SINV a été effectuée 3 fois, de façon indépendante, en utilisant les 3 stocks de SINV.

XI. Courbes de survie et titration virale de SINV

Les courbes de survie ont été réalisées à partir de 15 drosophiles (w^{1118}, $Dcr\text{-}2^{[L811Fx]}$, $CG4572^{c05963}$), à partir du jour de l'infection et jusqu'à 10 jours après infection par SINV, comme décrit précédemment. Le pourcentage de drosophiles vivantes a été déterminé chaque jour et ces drosophiles ont été transférées dans des milieux neufs tous les 2 jours. Les courbes de survie ont été construites en utilisant le logiciel Prism. n = 3 expériences indépendantes.

En parallèle, tout au long de la cinétique d'infection, les titres viraux des drosophiles infectées par SINV ont été déterminés par la méthode de titration virale en dilution limite, basée sur la méthode de Reed et Muench, et exprimés en TCID50/ml (Reed and Muench, 1938). Pour cela, les drosophiles infectées ont été anesthésiées au CO_2 puis transférées dans un eppendorf contenant 1 ml de PBS 1X et 30 µl de détergent (liquide vaisselle). Sous le poste de sécurité microbiologique, ces drosophiles ont été rincées 3 fois avec 1 ml de PBS 1X. Chaque drosophile a été broyée individuellement dans 1 ml de milieu de broyage, composé d'antibiotiques et d'antifongiques (**Tableau 4.**), auquel ont été ajoutées 3 billes magnétiques, en utilisant le broyeur Tissue Lyser II (Qiagen). Le milieu de broyage est préparé extemporanément avant chaque expérience. Les titrations en TCID50 ont été effectuées à partir de dilutions en série ($1/10^{ème}$), de 100 µl de surnageant de broyat de drosophiles puis en ajoutant les cellules Vero E6. n = 3 expériences indépendantes incluant 5 drosophiles par jour post-infection.

Constituants du milieu de broyage	Concentration finale
Milieu DMEM supplémenté de 4% de SVF (Gibco)	1X
Amphotéricine B (PAN Biotech GmbH)	2,5 µg/ml
Nystatine (Gibco)	100 U/ml
Gentamicine (Gibco)	50 g/ml
Pénicilline-streptomycine (Gibco)	50 g/ml

Tableau 4. Préparation de milieu de broyage des drosophiles.

XII. Titration virale de DMelSV

En parallèle des expériences de séquençage à haut débit, nous avons injecté les drosophiles Makindu (*D. simulans*) avec DMelSV de façon intrathoracique, dans un volume de 207 nl. La validation de l'infection a été réalisée à l'aide du test CO_2, du jour 6 au jour 14 après infection par DMelSV. Les drosophiles ont été maintenues sur milieu classique, contenant de l'agar, de la farine de maïs et de la levure morte et incubées à 25°C. A chaque jour de la cinétique d'infection, les drosophiles ont été soumises à 100% de CO_2 puis déposées sur la glace pendant 15 minutes. Suite au retour à 25°C, les drosophiles infectées par DMelSV sont paralysées de façon irréversible et succombent à l'infection. Le taux de drosophiles mortes a été quantifié, permettant de déterminer le pourcentage d'infection de ces drosophiles à chaque jour de la cinétique d'infection.

XIII. Dissection des drosophiles

A 6 jours après infection par SINV ou 8 jours après infection par DMelSV, les drosophiles ont été anesthésiées au CO_2 puis les ovaires ont été disséqués à l'aide d'une minutie de dissection en étain. Les paires d'ovaires et les carcasses de drosophiles (corps de la drosophile sans les ovaires) ont été déposées dans 1 ml de PBS 1X. Les têtes de drosophiles ont été disséquées de la même façon. Les expérimentations suivantes ont été réalisées immédiatement après l'étape de dissection.

XIV. Séquençage du transcriptome (RNA-seq)

L'ARN total a été extrait à partir de 30 paires d'ovaires ou 30 carcasses, à 6 jours ou 8 jours après l'infection par SINV ou DMelSV respectivement, en utilisant une méthode

d'extraction au Trizol (Thermo Fisher Scientific) couplée au kit RNeasy Mini (Qiagen), et en suivant les recommandations du fournisseur. L'ARN total purifié a été ensuite traité à la Turbo DNase (Ambion), permettant l'élimination de toute trace d'ADN. La qualité de l'ARN a été validée au Bioanalyzer (Agilent). Les banques de séquençage ont été préparées par les plateformes spécialisées, incluant la sélection des ARN polyadénylés. Ces banques ont été effectuées en utilisant le kit TruSeq RNA Sample Prep Kit v2. De plus, le séquençage a été réalisé en paire (*paired-end*), selon la technologie Illumina, permettant le séquençage à partir de chaque extrémité des fragments. Les banques enrichies ont été séquencées sur les plateformes Hiseq 4000 NGS (Illumina NGS) par la société Macrogen (Corée) pour les échantillons de Makindu et GenomEast IGBMC (France) pour tous les autres échantillons. La longueur des lectures est de 100 (GenomEast) ou 151 (Macrogen) nucléotides. Les extrémités des données brutes ont été nettoyées pour la qualité en utilisant le pipeline UrQt (Modolo and Lerat, 2015) et les adaptateurs de séquençage ont été éliminés à l'aide du pipeline trim Galore (Lindgreen, 2012) (**Figure 30.**). Les données de séquençage nettoyées ont été alignées sur les gènes de référence répertoriés sur FlyBase (ftp://ftp.flybase.net/releases/FB2019_01/dsim_r2.02/fasta/dsim-all-gene-r2.02.fasta.gz et ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/dmel_r6.26_FB2019_01/fasta/dmel-all-gene-r6.26.fasta.gz) en utilisant TopHat2 (Kim et al., 2013). Les alignements ont ensuite été triés à l'aide de SAMtools (H. Li et al., 2009) et les lectures alignées ont été comptées avec eXpress (Roberts et al., 2011). En parallèle, nous avons compté les lectures correspondants aux ET et aux virus en utilisant le module TEcount de TEtools (Lerat et al., 2017). La liste des séquences des ET incluait des séquences d'ET obtenues à partir de *D. melanogaster* et *D. simulans*, correspondant à 237 familles et sous-familles d'ET annotés. Pour cela, nous avons utilisé un fichier fasta contenant les séquences de toutes les insertions d'ET du génome séquencé chez *D. melanogaster* et chez *D. simulans*, incluant des séquences d'insertion d'ET collectées manuellement. Pour chacune des insertions, la famille d'ET correspondante a été renseignée dans un fichier texte appelé « rosette ». Ces deux fichiers sont disponibles en libre accès sur ftp://pbil.univ-lyon1.fr/pub/datasets/Roy2019. Les comptages de lecture obtenus pour les gènes d'une part et les ET d'autre part ont ensuite été concaténés pour constituer le tableau de comptage de lecture complet. Ce dernier a été directement analysé à l'aide du package R, DESeq2 (version 1.18) (Love et al., 2014), qui normalise selon la taille des banques. Pour évaluer si la moyenne des log2 *fold change* (ratio) des quantités de transcrits d'ET entre deux conditions est différente de zéro, le test de Wilcoxon sur données appariées a été utilisé (*Wilcoxon paired tests*). Les données ont été déposées sous le numéro d'accession PRJNA540249 pour les données de SINV. (Les données concernant le RNA-seq de DMelSV ne sont pas encore déposées).

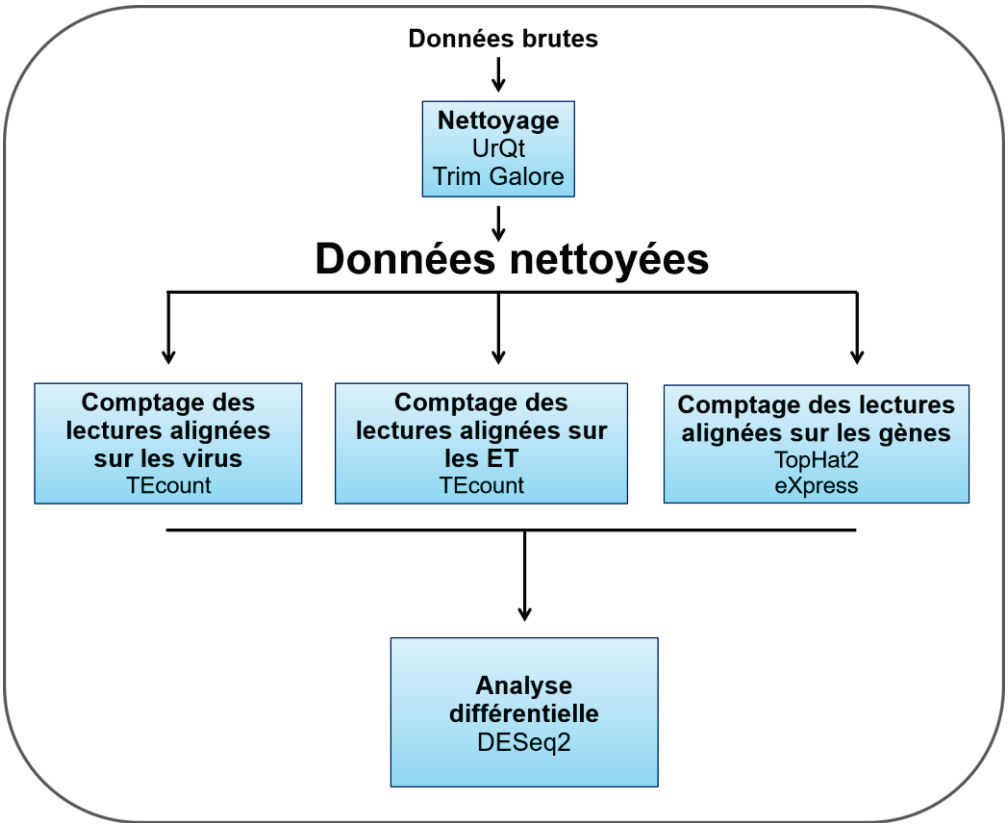


Figure 30. Analyse des données de séquençage du transcriptome (RNA-seq). Les données brutes ont été nettoyées puis alignées sur les gènes, les ET et les virus, permettant de constituer un tableau de comptage global, analysé en utilisant DESeq2.

XV. Analyse de données de séquençage publiques

Les données de séquençage du transcriptome (RNA-seq) de Merkling et collaborateurs, correspondant au numéro d'accession GSE56013 (A, B, E, F, I et J) (Merkling et al., 2015), ont été analysées de la même façon que précédemment, permettant notamment l'étude des transcrits d'ET et des virus. Ils se réfèrent à l'article de (Merkling et al., 2015) et ont été séquencés sur les plateformes Illumina HiSeq 2000 (Baseclear BV, Leiden, the Netherlands). Le RNA-seq a été réalisé en un seul réplicat, permettant une analyse d'exploration. Il s'agit du transcriptome de mouches *D. melanogaster* entières, infectées par DCV ou non infectées (condition mock). Les corps gras ont été isolés par dissection de la carcasse des drosophiles et élimination de l'intestin et des ovaires (Merkling et al., 2015). Le transcriptome des corps gras a également été analysé pour les drosophiles infectées par DCV *versus* les drosophiles mock, non infectées. Les mêmes drosophiles ont été mutées pour l'une des trois H3K9 méthyltransférases, appelée *G9a*, et les transcriptomes de ces mutants entiers et des corps gras correspondants ont été séquencés (Merkling et al., 2015). L'ensemble de ces drosophiles,

présentant un fond génétique identique, avait été affranchi de toute infection virale chronique et de la bactérie endosymbiotique *Wolbachia* (Merkling et al., 2015).

XVI. Séquençage des petits ARN interférents (Small RNA-seq)

Les petits ARN interférents ont été extraits à partir de 50 paires d'ovaires ou 50 carcasses, à 6 jours ou 8 jours après l'infection par SINV ou DMelSV respectivement, en utilisant des colonnes de chromatographie à échange d'anions HiTrap Q HP (GE Healthcare Life Sciences), en se référant à la méthode décrite par (Grentzinger and Chambeyron, 2014). Cette méthode permet une sélection des petits ARN interférents qui sont impliqués dans des complexes nucléoprotéiques, étant probablement des petits ARN interférents fonctionnels et exclut par conséquent les produits de dégradation d'ARNm.

Voici les étapes permettant l'isolement des petits ARN interférents :

- Les échantillons de paires d'ovaires et de carcasses ont été broyés dans 300 µl d'une solution HQ 0,1M d'acétate de potassium (Hépes 20 mM, KOH 20 mM, EDTA 0,2 mM, MgCl2 1,5 mM, glycérol autoclavé 10%, AcK 0,1M, Pefabloc® SC 0,1 mg/ml, DTT 1 mM). Une fois le broyage terminé, 250 µl de solution HQ 0,1 ont été ajoutés puis les échantillons ont été centrifugés pendant 30 minutes à 10000 g, 4°C.

- Un support de fixation avec tige a été utilisé pour maintenir la colonne. Chaque colonne a été activée avec 5 ml d'une solution HQ 1M d'acétate de potassium puis équilibrée avec 7 ml d'une solution HQ 0,1M d'acétate de potassium. Lors de ces étapes, il est primordial de ne pas faire entrer de bulles d'air dans la colonne car cela pourrait compromettre l'attachement des protéines sur la colonne.

- Le surnageant de l'échantillon broyé a été déposé dans la colonne activée et équilibrée (**Figure 31.**). Puis, 900 µl de solution HQ 0,1M ont été déposés dans la colonne et le filtrat a été récupéré. Après avoir ajouté 100 µl de solution HQ 4M, cette dernière solution a été déposée dans la colonne permettant la récolte de la première élution. Le mix HQ (composé de 900 µl de solution HQ 0,1M + 100 µl de de solution HQ 4M) a été ensuite déposé dans la colonne et la deuxième élution a été récoltée.

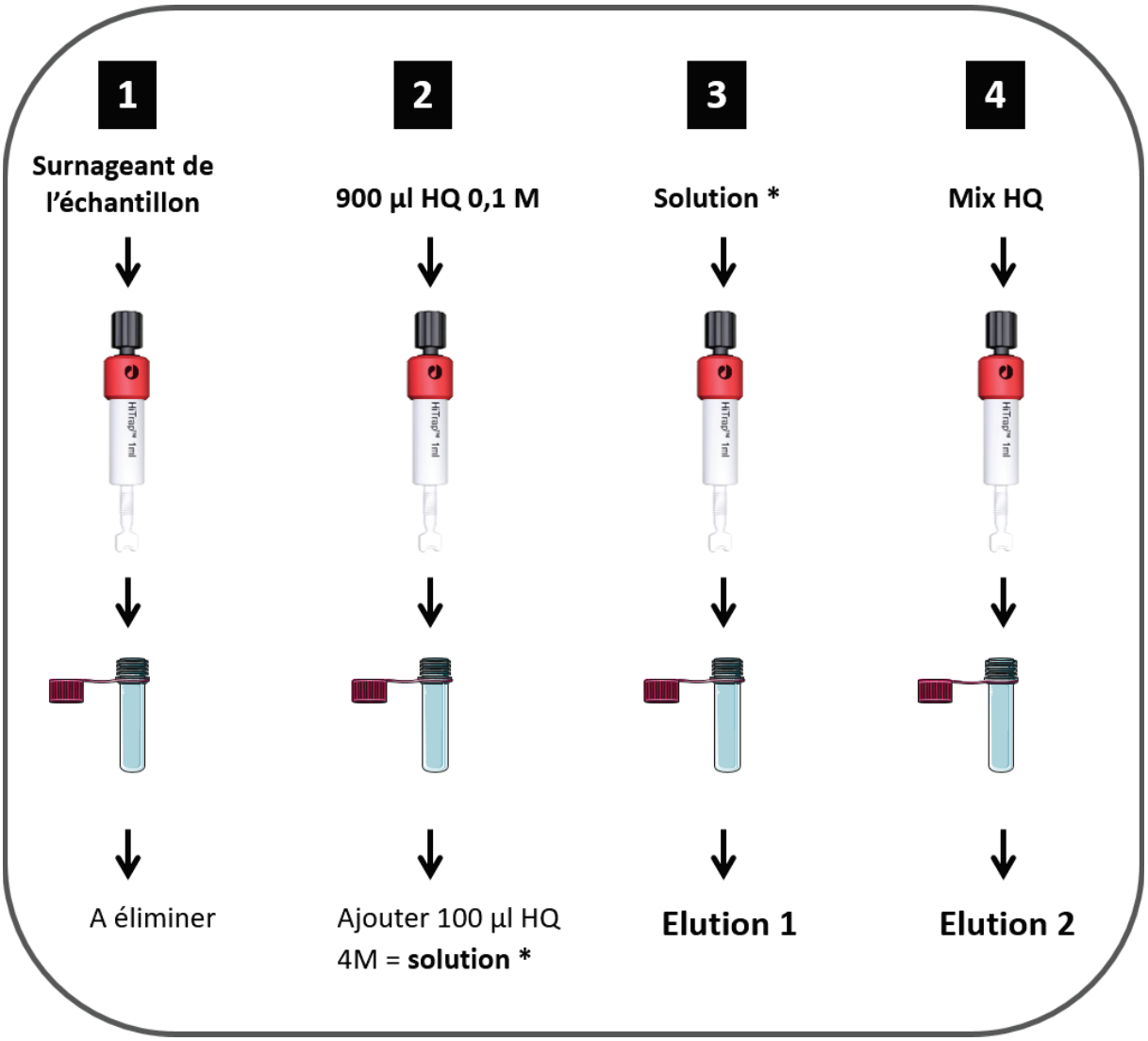


Figure 31. Isolement des petits ARN interférents. Les quatre étapes clés permettent d'isoler les petits ARN interférents *via* 2 élutions (étapes 3 et 4).

- Chaque colonne peut être utilisée plusieurs fois pour l'extraction des petits ARN interférents d'un même tissu. Cependant, il est nécessaire de la laver avec 2 ml de solution HQ 1M puis 1ml de solution HQ 4M, puis de l'équilibrer avec 2 ml de solution HQ 0,1M.

- Les petits ARN interférents ont été purifiés par la méthode d'extraction en phénol-chloroforme (Sigma-Aldrich) et quantifiés au Qubit en utilisant le kit Qubit microRNA Assay (Thermo Fisher).

La sélection sur gel de la taille des petits ARN (18-50 pb) a été effectuée par la plateforme de séquençage, GenomEast IGBMC (France), favorisant la sélection des petits ARN interférents d'intérêt (siARN et piARN). Les petits ARN purifiés ont été utilisés pour la préparation de la banque (SQ00/SIL-04-SR), puis le séquençage HiSeq 4000 (Illumina NGS) a été effectué à partir d'une seule extrémité (*single-end*) de chaque fragment, par la plateforme GenomEast IGBMC. A partir des données brutes, d'une longueur de 50 nucléotides par lecture, les adaptateurs ont été éliminés (TGGAATTCTCGGGTGCCAAGGAACTCCAGTCACTTA) à l'aide de cutadapt (**Figure 32.**) (Martin, 2011). Le nombre de lectures comptées a été normalisé par ceux des miARN (séquences des miARN extraites de Flybase : dsim-all-miRNA-r2.02.fasta.gz pour les échantillons Makindu (*D. simulans*) et dmel-all-miRNA-r6.25.fasta.gz pour les échantillons de *D. melanogaster*). L'abondance des miARN n'est pas impactée par l'infection, et permet de prendre en compte la variation des petits ARN les plus représentés entre les différents tissus de drosophile. En utilisant le pipeline PRINSEQ (Schmieder and Edwards, 2011), les lectures ont été filtrées par tailles comprises entre 23 et 30 nucléotides, considérés comme des piARN et par tailles de 21 nucléotides, considérés comme des siARN. En utilisant une version modifiée, publiée dans (Fablet et al., 2019) du module TEcount de TEtools (Lerat et al., 2017), les lectures ont été alignées sur la liste des séquences d'ET (identique à celle utilisée lors de l'analyse du RNA-seq) dans les directions sens et antisens. Pour évaluer si la moyenne des log2 *fold change* (ratio) des quantités de petits ARN interférents (piARN et siARN) entre deux conditions est différente de zéro, le test de Wilcoxon sur données appariées a été utilisé (*Wilcoxon paired tests*). Les données ont été déposées sous le numéro d'accession PRJNA540249 pour les données de SINV. (Les données concernant le Small RNA-seq de DMelSV ne sont pas encore déposées).

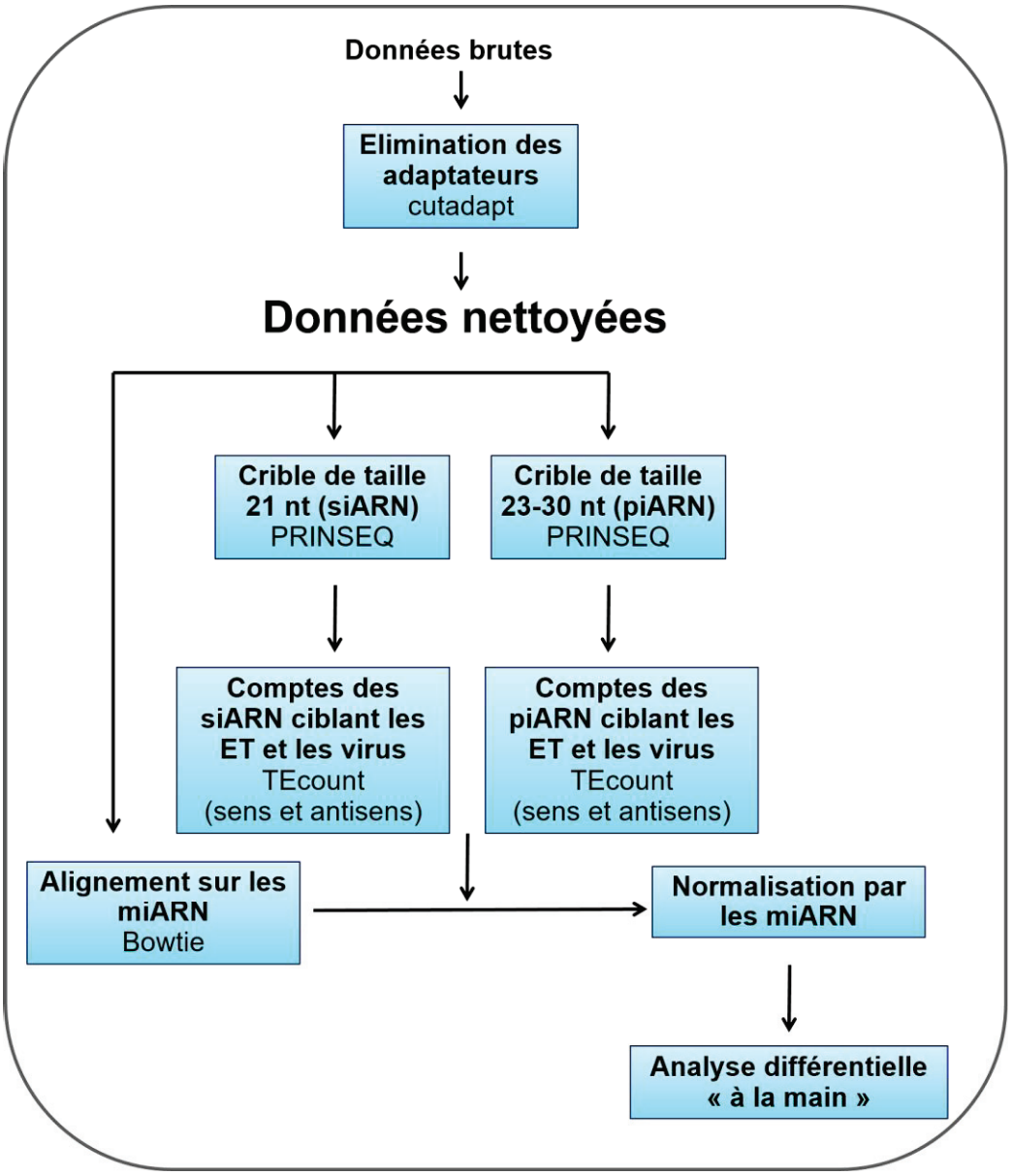


Figure 32. Analyse des données de séquençage des petits ARN interférents (Small RNA-seq). Les données brutes ont été nettoyées puis criblées selon leur taille afin de distinguer les piARN et les siARN. L'ensemble des comptes des lectures des petits ARN interférents a été normalisé par les comptes des lectures des miARN.

XVII. Titrations virales de SINV à partir des différents tissus de drosophiles

En parallèle des expériences de RNA-seq et Small RNA-seq, les titrations virales ont été effectuées à partir des paires d'ovaires, des carcasses, des têtes et des drosophiles entières, de façon individuelle (excepté pour les têtes, titrées par groupe de 5). De la même façon que précédemment, chaque échantillon est broyé dans 1 ml de milieu de broyage (**Tableau 4.**) et

les titres viraux sont exprimés en TCID50/ml selon la méthode de Reed et Muench (Reed and Muench, 1938). Les titrations en TCID50 ont été effectuées à partir de dilutions en série (1/10ème), de 100 µl de surnageant de broyat de l'échantillon puis en ajoutant les cellules Vero E6. n = 3 expériences indépendantes incluant chacune 10 échantillons individuels de carcasses ou de paires d'ovaires. Trois pools de cinq têtes ont été titrés pour chacune des 3 expériences.

XVIII. Détection de DAV par RT-PCR à partir des drosophiles choniquement infectées par DMelSV Hap23

L'ARN total a été extrait à partir de 5 drosophiles chroniquement infectées par DMelSV Hap23, en utilisant une méthode d'extraction au Trizol (Thermo Fisher Scientific) couplée au kit RNeasy Mini (Qiagen), selon les recommandations du fournisseur. L'ARN total purifié est ensuite traité à la Turbo DNase (Ambion), permettant l'élimination de toute trace d'ADN. La transcription inverse de 500 ng d'ARN a été effectuée à l'aide du kit Omniscript (Qiagen) associé à la méthode d'amorçage des ARNm par des oligodT (Invitrogen). La PCR a été réalisée en utilisant le kit de PCR Go Taq (Promega). Le gène de ménage *actine* a été utilisé comme témoin d'extraction d'ARN et d'amplification de l'ADN complémentaire pour chaque échantillon. Les amorces spécifiques de DAV sont référencées dans le **Tableau 3**. n = 2 expériences indépendantes (5 drosophiles par échantillon d'ARN).

XIX. Transcription inverse et PCR quantitative (RT-qPCR)

L'ARN total a été extrait à partir de 30 paires d'ovaires ou 30 carcasses (corps de la drosophile sans les ovaires), ou à partir de 30 têtes, 30 carcasses (corps de la drosophile sans les ovaires et sans la tête) à 6 jours ou 8 jours après l'infection par SINV ou DMelSV, respectivement. Pour cela, une méthode d'extraction au Trizol (Thermo Fisher Scientific) couplée au kit RNeasy Mini (Qiagen) a été utilisée, en suivant les recommandations du fournisseur. L'ARN total purifié est ensuite traité à la Turbo DNase (Ambion), permettant l'élimination de toute trace d'ADN. La transcription inverse a été réalisée à partir de 200 ng (carcasses/têtes) ou 500 ng (ovaires) d'ARN total en utilisant le kit SuperScript IV VILO Master Mix (Thermo Fisher Scientific), selon les recommandations du fournisseur. Les ADN complémentaires ont été traités à la RNaseH (Thermo Fisher Scientific), permettant le clivage de la chaîne d'ARN à partir de l'extrémité 3' des duplex ADN/ARN libérant l'ADN simple brin. Les qPCR (Quantitative Polymerase Chain Reaction) ont été effectuées à l'aide du kit LightCycler® 480 SYBR Green I Master (Roche Life Science), selon les instructions du fabricant. La liste des amorces spécifiques utilisées pour les quantifications est donnée dans le

Tableau 5. Les quantifications ont été normalisées par l'ARNm des gènes *rp49* et/ou *alcohol deshydrogenase* (*adh*). Les analyses statistiques ont été réalisées avec le logiciel R et en utilisant le test de Wilcoxon-Mann-Whitney.

Cible à détecter	Séquences des amorces (5'-3')
accord	Fw 5' TTCACCCGTCGAAAGACTTC 3' Rv 5' GCCGTGAAAGAGTTCGAAAG 3'
gypsy **(Touret et al., 2014b)**	Fw 5' GGCTCATTGCCGTTAAACAT 3' Rv 5' TCTTCCTTCTTTCGCTGAGG 3'
412	Fw 5' TTGATGGGCAAAAGATCCAT 3' Rv 5' TTGCTGGAATTGTCGTTTCA 3'
tirant **(Akkouche et al., 2012)**	Fw 5' ACA CGT TCC CTG AAC AGA CG 3' Rv 5' GAA CGT TAC CAA TCC GAG CA 3'
Doc6	Fw 5' ACATCATGTGCCCATTGTGC 3' Rv 5' AAATTCGGCATGGGCATGTC 3'
Doc	Fw 5' GCACCTGGTTTCGAAGTGAT 3' Rv 5' GTAGCCAGGCGATTTTCTTG 3'
F element	Fw 5' ACGTCGCTTACGCAGAGAAT 3' Rv 5' GTTTTGTGCCTGTTGGTGTG 3'
SINV	Fw 5' TGCGTGATTTCGTCCGTGTA 3' Rv 5' ACTGGTCGGCCAACGTTTAT 3'
DMelSV	Fw 5' ATGTAACTCGGGTGTGACAG 3.' Rv 5' CCTTCGTTCATCCTCCTGAG 3'
rp49 **(Akkouche et al., 2012)**	Fw 5' CGGATCGATATGCTAAGCTGT 3' Rv 5' GCG CTT GTT CGA TCC GTA 3'
adh **(Fablet, 2014)**	Fw 5' CCGTGGTCAACTTCACCAGCTC 3' Rv 5' TCCAACCAGGAGTTGAACTTGTGC 3'

Tableau 5. Amorces utilisées pour la détection des transcrits des ET et des virus SINV et DMelSV. Les gènes de ménage *rp49* et *adh* ont été utilisés comme témoins de RT-qPCR, permettant la normalisation des quantifications.

XX. Immunomarquage d'ovaires et de coupes de drosophiles infectées par le virus Sindbis et microscopie confocale

1. Préparation des blocs d'OCT contenant les tissus à couper

Les drosophiles entières ont été anesthésiées à 6 jours après infection par SINV et déposées dans un eppendorf contenant 1 ml de paraformaldéhyde (PFA) à 2%, pendant 4 heures à température ambiante et sous agitation. En parallèle, les têtes de drosophiles ont été disséquées et traitées de la même façon. La fixation au paraformaldéhyde permet l'inactivation du virus. Après 3 lavages au PBS 1X, les drosophiles entières ou les têtes sont incubées 16h à 4°C, sous agitation, dans une solution de sucrose à 12% (PBS 1X). Elles sont ensuite déposées dans des petits dispositifs à blocs d'OCT et stockés à -80°C. n = 2 expériences indépendantes, incluant 6 drosophiles ou 10 têtes de drosophiles chacune.

2. Préparation des coupes de tissus

Les coupes de 14 µm d'épaisseur ont été réalisées à l'aide d'un cryostat (LEICA CM3050S) à -21°C, puis ont été déposées sur des lames SuperFrost Plus (Thermo Scientific).

3. Immunomarquage : détection de la capside de SINV

Les coupes ont été immédiatement fixées au PFA 0,5%, pendant 20 minutes à température ambiante puis lavées 3 fois au PBS 1X. La perméabilisation des coupes a été effectuée avec une solution de PBS Triton 0,3% pendant 30 minutes. Le marquage de la capside de SINV nécessite une incubation pendant la nuit, à 4°C, dans la solution d'Ac primaire contenant l'Ac anti-capside de SINV (clone C42) (Greiser-Wilke et al., 1989), dilué au 1/300ème dans du PBS + NGS5% +Triton 0,1%. Une chambre humide a été créée pour éviter l'évaporation de la solution sur les lames. Après 3 lavages au PBS Triton 0,3%, les coupes ont été incubées pendant 2 heures à température ambiante et à l'abri de la lumière, dans la solution d'Ac secondaire, contenant l'Ac anti-GFP, A488 de chèvre anti-souris (A-11029, Invitrogen), diluée au 1/500ème dans du PBS + NGS5% +Triton 0,1%. Les coupes ont été rincées 3 fois avec du PBS 1X avant de procéder au montage des coupes entre lame et lamelle dans le milieu de montage Vectashield incluant le DAPI, permettant la coloration de l'ADN. Les coupes ont été conservées à 4°C à l'abri de la lumière. Au moins 2 expériences indépendantes ont été réalisées pour chaque immunomarquage.

L'immunomarquage des ovaires inclut une fixation au PFA 4%, pendant 4 heures à température ambiante puis 3 lavages au PBS 1X. La perméabilisation des ovaires a été effectuée avec une solution de PBS Triton 0,5% pendant 1 heure. La suite de l'immunomarquage a été réalisée de la même façon que pour les coupes de drosophiles.

4. Observation des coupes au microscope confocal

Les acquisitions ont été effectuées en utilisant un microscope confocal LSM 510 META (Carl Zeiss) et analysées à l'aide du logiciel ZEN. Les acquisitions en lumière blanche ont été traitées par la méthode de correction d'ombrage (*shading correction*) et de correction d'alignement (*Stitching correction*).

RESULTATS EXPERIMENTAUX (I)

RESULTATS EXPERIMENTAUX (I)

I. Les voies d'ARN interférence siARN et piARN interagissent et modulent la transcription des éléments transposables lors de l'infection virale chez *Drosophila*

(Article en préparation pour une soumission sous un format de lettre au journal *Nature*)

1. Introduction

Les insectes ne sont pas seulement exposés aux prédateurs ou aux agents pathogènes, tels que les virus. En effet, l'ensemble des génomes des insectes contient des éléments mobiles de l'ADN, appelés éléments transposables (ET), qui sont des parasites génomiques, dont certains présentent des similarités structurales avec les rétrovirus, incluant notamment une séquence de transcriptase inverse. Ajoutés à cette similarité de structure, les mécanismes principaux de contrôle des ET et des virus sont engendrés par la machinerie d'ARNi. Chez les insectes, les ET sont contrôlés par la voie des piARN (*Piwi-interacting RNA*) dans les gonades, tandis que la voie des siARN (*small interfering RNA*) contrôle les ET dans le soma et a un rôle primordial dans le contrôle systémique des infections virales. L'ensemble de ces informations suggère un dialogue entre les virus et les ET, qui pourrait impliquer ces voies d'ARNi. La caractérisation de ces interactions potentielles est d'une grande importance pour l'étude des interactions entre les insectes et les virus et pour l'amélioration des connaissances de la diversité génétique et l'évolution des génomes des insectes.

Dans ce chapitre, nous allons explorer le transcriptome et le répertoire des petits ARN interférents chez deux espèces de drosophiles, *D. melanogaster* et *D. simulans* lors de l'infection expérimentale par un arbovirus (ARthropod-BOrne VIRUS), le virus Sindbis (SINV). La lignée Makindu de *D. simulans* a été utilisée dans cette étude ; elle provient d'une population naturelle africaine et est une des lignées précisément caractérisées pour ses ET (Akkouche et al., 2013, 2012; Fablet et al., 2019). L'ensemble des drosophiles a été disséqué afin de séparer les ovaires -le tissu où la voie des piARN est décrite pour le contrôle des ET chez la drosophile- des carcasses (corps de la drosophile sans les ovaires), où la voie des siARN endogènes semble active dans le contrôle des ET (Jones et al., 2016; Zuo et al., 2016). L'objectif de ce chapitre est de déterminer si l'infection virale impacte la voie de contrôle génomique des ET, entraînant une modulation du niveau de la transcription des ET, dans les carcasses et dans les ovaires. Pour cela, nous avons utilisé des techniques de séquençage à haut débit, de

type RNA-seq pour l'analyse du transcriptome et de type Small RNA-seq pour l'analyse des petits ARN interférents.

Tout d'abord, nous avons caractérisé la réplication de SINV dans chacune des lignées de drosophiles utilisées. Seules les femelles ont été infectées puis analysées, permettant l'étude dans les ovaires, où la voie des piARN est décrite pour contrôler les ET. L'infection par SINV n'induit pas de mortalité chez les drosophiles sauvages (*D. melanogaster* et *D. simulans*). Nous avons ensuite étudié dans les carcasses la réponse attendue de la machinerie d'ARNi contre SINV, médiée principalement par les siARN. Nous avons ensuite déterminé l'impact de l'infection par SINV sur l'abondance des transcrits d'ET et sur le répertoire des petits ARN interférents ciblant ces ET dans ces deux espèces de drosophiles. L'analyse transcriptomique inclut celle des transcrits des virus et celle de 237 familles et sous-familles d'ET. Dans un second temps, nous avons utilisé des mutants de *D. melanogaster*, permettant de décrypter les mécanismes sous-jacents mis en jeux lors de ces potentielles modulations. Pour cela, nous avons utilisé des mutants de la voie de l'import (*uptake*) de l'ARN double brin ($CG4572^{c05963}$), impliquée dans la diffusion systémique des siARN (Saleh et al., 2006) et également des mutants de la voie des siARN ($Dcr2^{[L811fsX]}$) (Tassetto et al., 2017). Après avoir étudié la survie de ces mutants lors de l'infection par SINV, nous avons caractérisé la réplication virale. De la même façon que précédemment, le transcriptome a été étudié chez les mutants $Dcr2^{[L811fsX]}$ et le répertoire des petits ARN interférents a été analysé chez les mutants $Dcr2^{[L811fsX]}$ et $CG4572^{c05963}$, et à partir des carcasses et des paires d'ovaires.

Quatre questions principales sont posées dans cette première partie :

- L'activité des ET est-elle modulée au cours de l'infection par SINV ?
- La modulation des ET est-elle associée à un changement du taux de production des petits ARN interférents lors de l'infection virale ?
- La voie de l'import des ARN double brin est-elle impliquée dans ces modulations ?
- La voie des siARN impacte-elle la production des piARN ciblant les ET et l'activité sous-jacente des ET ?

2. Infection par le virus Sindbis chez *D. melanogaster*

A. Identification et élimination de *Wolbachia* et des infections virales chroniques chez *D. melanogaster*

Depuis quelques années, il a été démontré que l'élimination des infections persistantes dans les stocks des drosophiles était primordiale pour étudier la réponse virale suite à des inoculations expérimentales (Merkling and van Rij, 2015). En effet, les lignées de drosophiles

de laboratoire sont souvent infectées par de nombreux virus, tels que DCV, DAV et NV. Ces virus à ARN peuvent être inducteurs ou suppresseurs des voies d'ARNi (Lopez et al., 2018; Mierlo et al., 2012), et vont également activer des voies au niveau de la physiologie et le métabolisme de l'hôte rendant difficile l'interprétation des résultats concernant le virus d'intérêt (Chtarbanova et al., 2014; Eleftherianos et al., 2011; Merkling and van Rij, 2013). Par conséquent, les œufs de l'ensemble des drosophiles *D. melanogaster* (w^{1118}, $Dcr2^{[L811fsX]}$ et $CG4572^{c05963}$) ont été traités à l'eau de javel domestique pour l'élimination de ces infections persistantes, qui a ensuite été validée par RT-PCR (**Figure S1.**) et lors de l'analyse transcriptomique (RNA-seq) des transcrits viraux des drosophiles w^{1118} et $Dcr2^{[L811fsX]}$ (**Figure 33.**). De nombreux virus d'insectes ont été analysés par RNA-seq, et aucune infection chronique n'a été révélée, mis à part SINV qui a été injecté expérimentalement. Par ailleurs, il est intéressant de noter que le nombre de transcrits de SINV est supérieur dans les carcasses par rapport aux ovaires (le log2 *fold change* n'est pas précisé car la normalisation est spécifique du tissu), incluant également un plus grand nombre de transcrits de SINV chez les mutants $Dcr2^{[L811fsX]}$ par rapport aux drosophiles sauvages w^{1118} en carcasses (moyenne du log2 *fold change* = 4,10) et en ovaires (moyenne du log2 *fold change* = 4,05) ; (**Figure 33.**).

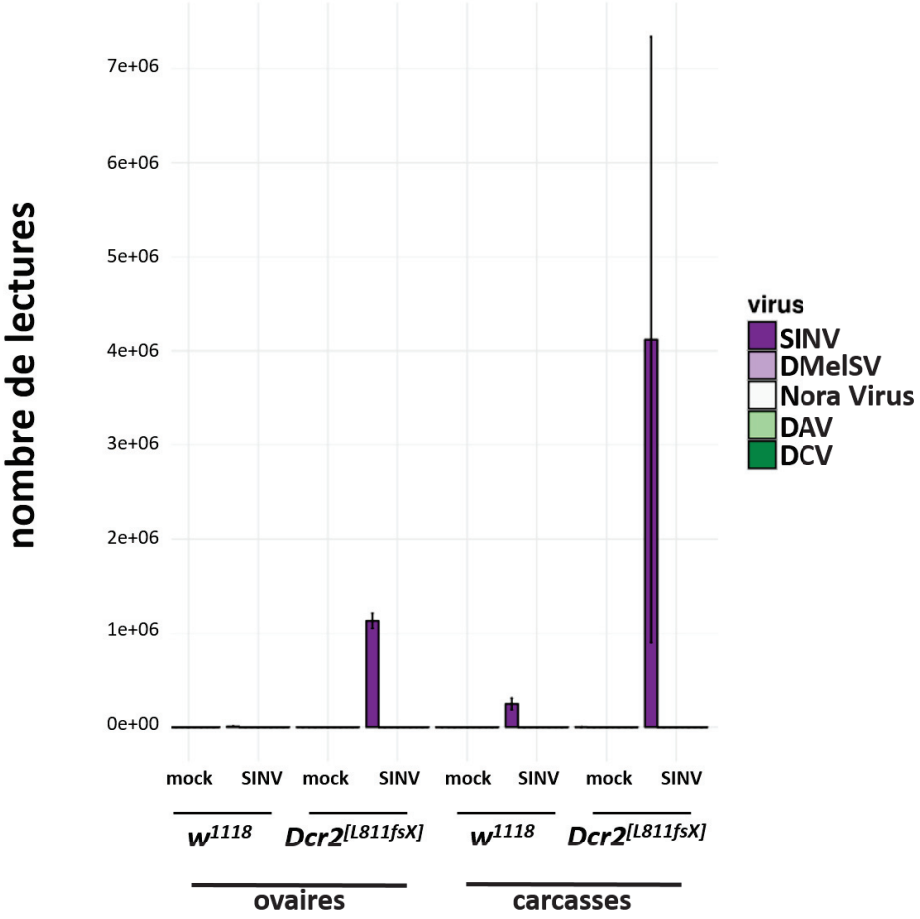


Figure 33. Absence de détection de transcrits de virus spécifiques de *D. melanogaster* (w^{1118} et $Dcr\text{-}2^{[L811fsX]}$) après javellisation des œufs de drosophile. Détermination du nombre de lectures (normalisation par DESeq2

selon la taille des banques) correspondant aux transcrits viraux déterminé par RNA-seq après javellisation des œufs de drosophiles, pour les infections chroniques suivantes : DMelSV, NV, DAV, DCV, dans les tissus (carcasses et ovaires) issus des drosophiles infectées expérimentalement par SINV ou non infectées (mock). n = 2 expériences indépendantes (30 tissus par échantillon). (La condition *mock* correspond à la condition non infectée par SINV).

De plus, l'infection par la bactérie endosymbiotique *Wolbachia* affecte fortement la réplication des virus à ARN (Monsanto-Hearne and Johnson, 2018; Teixeira et al., 2008) et de façon intéressante il a été rapporté recemment que *Wolbachia* induit la résistance à SINV après infection expérimentale (injection) chez la drosophile, incluant une augmentation de l'expression du gène de méthyltransférase *Mt2* (Bhattacharya et al., 2017). Ces bactéries ont donc été éliminées pour toutes les drosophiles *D. melanogaster*, après un traitement à la tétracycline (**Figure S2.**).

Ainsi, dans cette première partie, nous avons validé l'élimination de *Wolbachia* et des virus spécifiques de drosophiles chez *D. melanogaster*, incluant les lignées w^{1118}, $Dcr2^{[L811fsX]}$ et $CG4572^{c05963}$ (absence de données de RNA-seq). Nous nous sommes ensuite intéressés dans un premier temps aux drosophiles contrôles w^{1118} et nous avons commencé par analyser la réplication de SINV dans ce modèle de drosophile exempt de *Wolbachia* et d'infection virale naturelle.

B. Etude de la réplication du virus Sindbis chez *D. melanogaster* w^{1118}

La cinétique de réplication de SINV à partir des drosophiles entières w^{1118} (**Figure 34 A.**) a été effectuée en utilisant la méthode de titration virale en TCID50. On observe une augmentation accrue de la réplication virale de SINV jusqu'à 4 jours après l'infection (médiane au jour 2 = 1,5 log10(TCID50/ml) ; médiane au jour 4 = 3,2 log10(TCID50/ml)) puis une phase plateau où la réplication de SINV ralentit et devient constante (médiane au jour 6 = 3,1 log10(TCID50/ml) ; médiane = 3,5 log10(TCID50/ml) aux jours 8 et 10 après infection).

Pour la suite des expériences, nous avons sélectionné le jour 6 après infection, qui correspond au début de la phase plateau de réplication virale, suggérant que les voies d'ARNi sont engagées. Les titres viraux de SINV ont été déterminés à partir des drosophiles disséquées afin de séparer les paires d'ovaires des carcasses (**Figure 34 B.**). La dissection ne permet pas la détection du virus dans l'inoculum au jour 0, en particulier pour les paires d'ovaires, en raison du seuil de détection de la technique de TCID50 fixée à 1,33 log10(TCID50/ml). Les titres viraux de SINV au niveau des carcasses (médiane = 2,8 log10(TCID50/ml) sont comparables à ceux déterminés en drosophiles entières au jour 6 après l'infection. Par ailleurs, SINV est très

faiblement détecté dans les ovaires des w^{1118} (5 drosophiles positives sur 20). L'analyse du transcriptome et du répertoire des petits ARN interférents se fera de façon tissu spécifique puisque l'implication des siARN et des piARN est différente dans les tissus somatiques et gonadiques. En effet, la voie des siARN est ubiquitaire et est principalement impliquée dans l'immunité antivirale alors que les piARN sont majoritairement exprimés au niveau des gonades des organismes eucaryotes où ils protègent l'intégrité du génome.

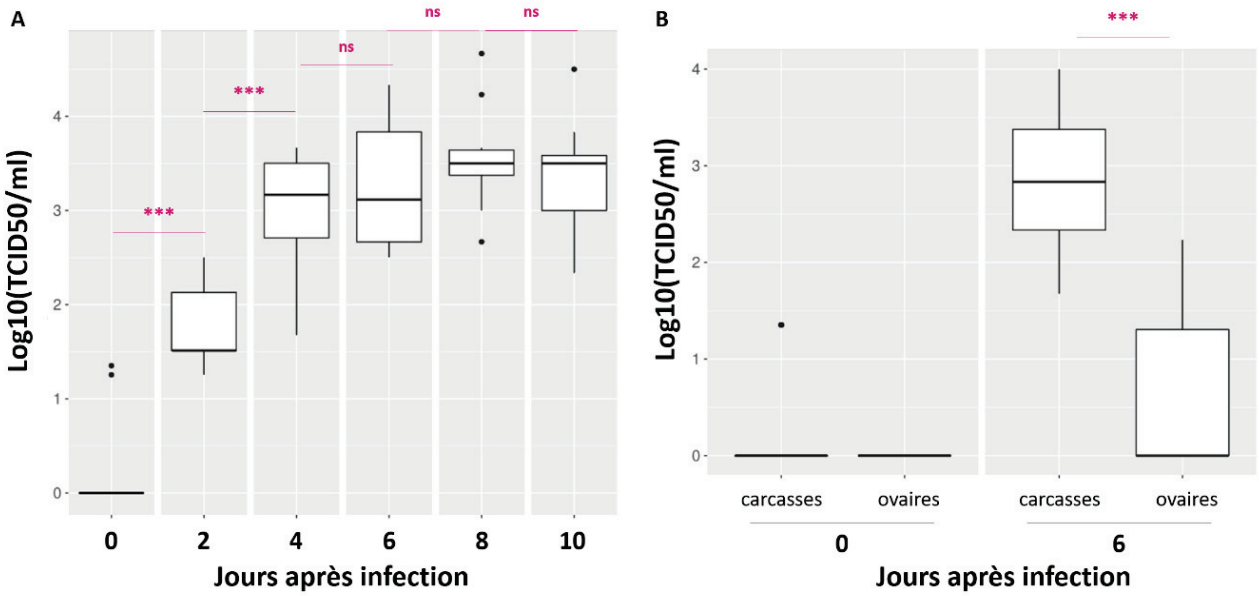


Figure 34. Cinétique de réplication de SINV chez *D. melanogaster*. Les drosophiles w^{1118} ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de drosophiles infectées entières **(A)**, tout au long de la cinétique de réplication (du jour de l'infection (jour 0) à 10 jours après infection) et à partir de broyats individuels de carcasses et de paires d'ovaires **(B)** au jour de l'infection et à 6 jours après l'infection. (A) n = 3 expériences indépendantes (5 drosophiles par condition). (B) n = 2 expériences indépendantes (10 drosophiles par condition). Test statistique de Wilcoxon-Mann-Whitney *; ns : non significatif ; ***P* < 0,001.

Ces résultats ont été confirmés par immunofluorescence en utilisant un anticorps anti-capside de SINV puisqu'aucune détection de la capside de SINV n'a été faite dans les ovaires de *D. melanogaster* w^{1118} (**Figure 35.**).

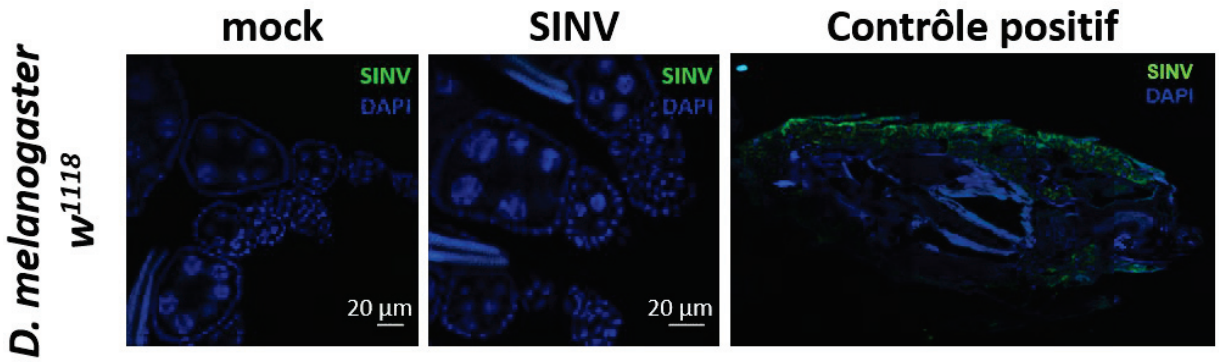


Figure 35. Absence de détection de SINV dans les ovaires de *D. melanogaster* w^{1118}. Les drosophiles w^{1118} ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl puis les paires d'ovaires ont été disséquées au jour 6 après infection. Le contrôle positif correspond à une coupe d'abdomen de drosophile infectée par SINV. La technique d'immunofluorescence a été réalisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488). L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 10 paires d'ovaires. (mock : condition non infectée ; SINV : condition infectée par SINV).

SINV se réplique chez les drosophiles w^{1118}, et plus précisément dans les carcasses, alors que ce virus ne semble pas se répliquer dans les ovaires. La cinétique de réplication virale nous a permis de sélectionner le temps d'infection pour les expériences suivantes, incluant des titres viraux suffisants. Par la suite, la production des petits ARN interférents viraux et des petits ARN interférents dirigés contre les ET sera explorée chez les drosophiles w^{1118}, aboutissant à l'identification des modulations des quantités des petits ARN interférents ciblant les ET et des transcrits des ET lors de l'infection par SINV par rapport à la condition non-infectée.

C. Modulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis chez *D. melanogaster* w^{1118}

a) Caractérisation de la réponse antivirale de type ARNi lors de l'infection par le virus Sindbis (w^{1118})

La méthode utilisée pour isoler les petits ARN interférents a permis la sélection des petits ARN interférents impliqués dans des complexes nucléoprotéiques, étant probablement des petits ARN interférents fonctionnels et exclut par conséquent les produits de dégradation d'ARNm (Grentzinger and Chambeyron, 2014). Nous avons séparé les lectures par taille,

permettant la distinction entre les petits ARN d'une longueur de 21 nucléotides et ceux de 23-30 nucléotides, et nous les avons considérés comme étant des siARN et des piARN respectivement. Comme attendu, lors de l'infection par SINV, à six jours après infection, la réponse de type ARNi implique principalement la voie des siARN (**Figure 36 A.**). De plus, la production des petits ARN de 21 nuléotides, correspondant aux siARN et ciblant SINV est plus importante dans les carcasses où le virus se réplique, par rapport aux ovaires où il ne semble pas y avoir de réplication virale. Une proportion équivalente de brin sens et anti-sens de ces siARN a été révélée dans les carcasses et dans les ovaires, lors de l'infection par SINV. La distribution des siARN ciblant SINV est homogène le long du génome de SINV, d'environ 11500 paires de bases, incluant des points chauds et des points froids de production (**Figure 36 B.**). Par ailleurs, lors de l'infection par SINV, les petits ARN d'une taille entre 23 et 30 nucléotides, correspondant aux piARN, et ciblant SINV, ont été très faiblement produits chez *D. melanogaster* à partir des carcasses et des ovaires (**Figure 36 A.**). Ces derniers ne présentent pas de signature « ping-pong » spécifique des piARN secondaires, incluant un chevauchement de 10 nucléotides entre le brin sens et le brin antisens des piARN (**Figure S3.**).

Ces résultats sont en accord avec la littérature (Petit et al., 2016). Lors de l'infection par SINV, la réponse de type ARNi chez *D. melanogaster* a été analysée et permet de valider le modèle d'infection, incluant la réplication virale de SINV et l'induction de la voie des siARN. Nous allons maintenant nous intéresser à l'impact de l'infection par SINV sur l'activité des ET et la production des petits ARN interférents ciblant les ET chez *D. melanogaster* w^{1118}.

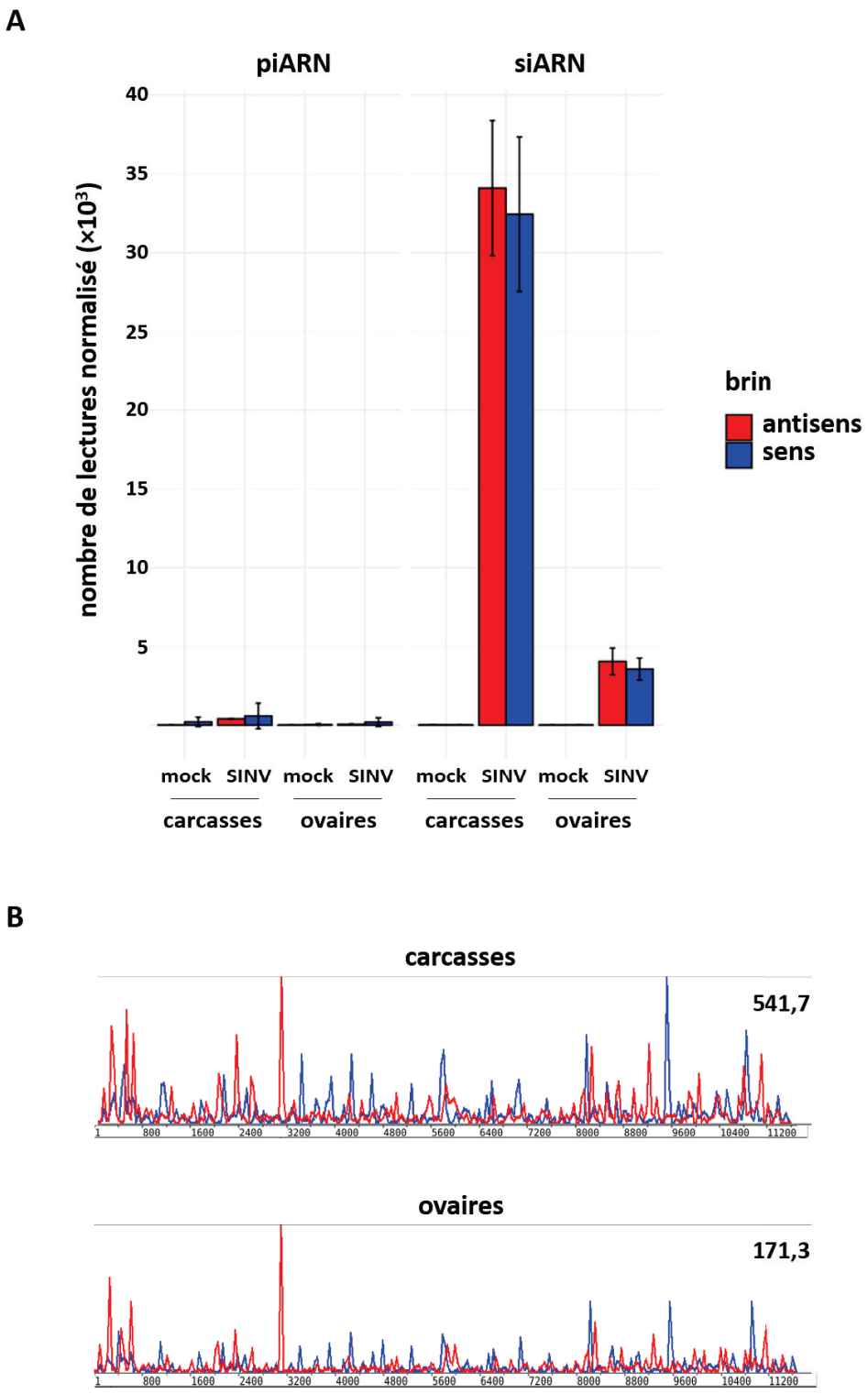


Figure 36. Analyse des siARN et des piARN ciblant SINV chez *D. melanogaster* w^{1118}. (A) Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des w^{1118}. L'ensemble des comptes des lectures a été normalisé par le nombre de lectures des miARN. **(B)** Distribution des petits ARN interférents de 21 nulcéotides ciblant SINV, le long du génome viral, incluant les brins sens (bleu) et antisens (rouge). La valeur numérique en haut à droite correspond au nombre de lectures pour le pic le plus haut. n = 2 expériences indépendantes (50 tissus par échantillon).

b) Modulation de l'abondance des transcrits d'ET et de la production des petits ARN interférents ciblant les ET chez *D. melanogaster*

Dans cette partie nous allons déterminer les effets engendrés lors de l'infection par SINV sur les quantités de transcrits des ET et sur la production des petits ARN interférents dirigés contre les ET chez *D. melanogaster* w^{1118}. L'analyse des comptes des lectures pour 237 familles et sous-familles d'ET a révélé une réduction globale significative de l'abondance des transcrits d'ET, dans les carcasses lors de l'infection par SINV en comparaison avec les échantillons non infectés (**Figure 37 A.** ; moyenne log2 *fold change* (ratio) = -0,32 ; *p-value* = 1,8e-24). Ces résultats sont associés à une augmentation significative des piARN (**Figure 37 B.** ; moyenne log2 *fold change* = 0,61 ; *p-value* = 1,1e-35) et des siARN (**Figure 37 C.** ; moyenne log2 *fold change* = 0,61 ; *p-value* = 2,1e-30) ciblant les ET. La diminution des quantités de transcrits d'ET est significativement dépendante de l'augmentation des productions de piARN (*p-value* = 6,9e-4) et de siARN (*p-value* = 3,1e-3), incluant une dépendance significative entre ces deux classes de petits ARN interférents (*p-value* = 4,7e-6) ; (**Figure S4**). De plus, de façon intéressante, nous avons identifié dans les carcasses non infectées (**Figure 38 A.**) et infectées par SINV (**Figure 38 B.**) des piARN dirigés contre différentes familles d'ET, telles que *doc* et *accord* qui présentent une signature « ping-pong » spécifique des piARN impliqués dans la voie secondaire.

Dans les ovaires, aucune particule infectieuse de SINV n'a été détectée par immunofluorescence (**Figure 35.)**. Dans cet organe, lors de l'infection des drosophiles w^{1118}, on retrouve une diminution significative de l'abondance des transcrits d'ET (**Figure 37 D.** ; moyenne log2 *fold change* = -0,10 ; *p-value* = 5,3e-14) mais beaucoup moins importante que celle observée dans les carcasses. Cependant, l'impact de l'infection par SINV sur la production des piARN et siARN ciblant les ET est opposé et atténué par rapport aux carcasses, incluant une diminution significative mais modérée des quantités de piARN (**Figure 37 E.** ; moyenne log2 *fold change* = - 0,23 ; *p-value* = 4,4e-37) et des siARN (**Figure 37 F.** ; moyenne log2 *fold change* = -0,12 ; *p-value* = 2,2e-26). Ces résultats ne sont pas totalement élucidés, incluant notamment le fait que l'activité des ET soit modulée de la même façon que les petits ARN interférents lors de l'infection par SINV, incluant une diminution des quantités de transcrits d'ET et des petits ARN interférents.

Par ailleurs, afin de déterminer si la modulation de la production des petits ARN interférents ciblant les ET impliquait également une modulation au niveau des quantités de transcrits des gènes codant les effecteurs de ces voies d'ARNi, nous avons comparé le niveau de transcrits d'un ensemble de ces gènes incluant *piwi*, *aub*, *ago3*, *zuc* (voie des piARN) et de *dcr2*, *ago2* (voie des siARN) entre la condition infectée par SINV et la condition non infectée (mock). Aucune modulation significative n'a été révélée dans les carcasses (**Figure S5 A.**) ni dans les ovaires (**Figure S5 B.**) indiquant que la modulation de la production des petits ARN interférents

n'engage pas au niveau ARN les gènes impliquée dans ces voies d'ARNi (évaluée par DESeq2 ; *p-value* > 0,05).

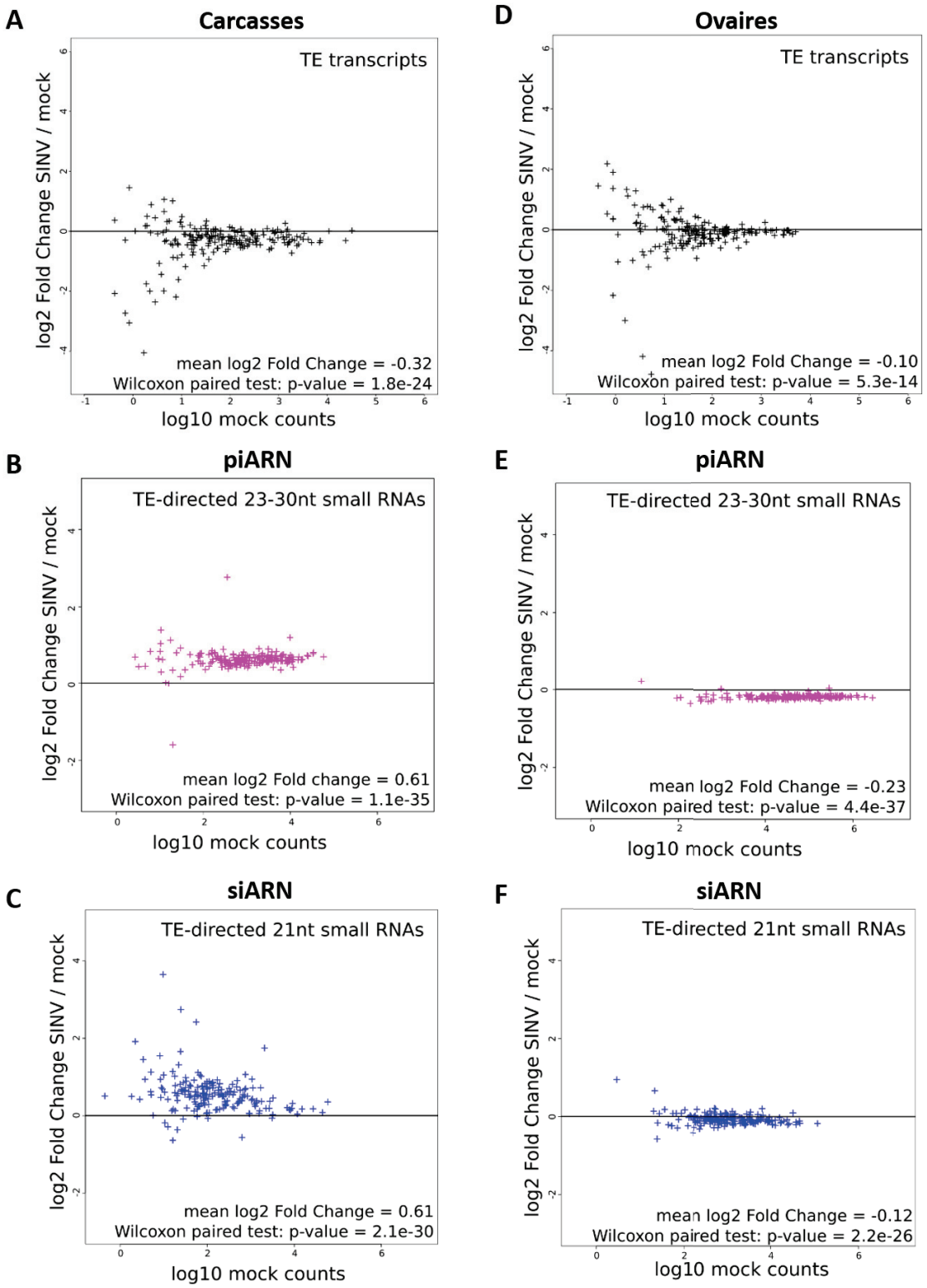


Figure 37. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez *D. melanogaster* (w^{1118}) lors de l'infection par SINV. Chaque point représente le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par SINV et la condition non infectée (mock), pour

les transcrits d'ET **(A, D)**, pour les piARN **(B, E)** et pour les siARN **(C, F)**. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). n = 2 expériences indépendantes.

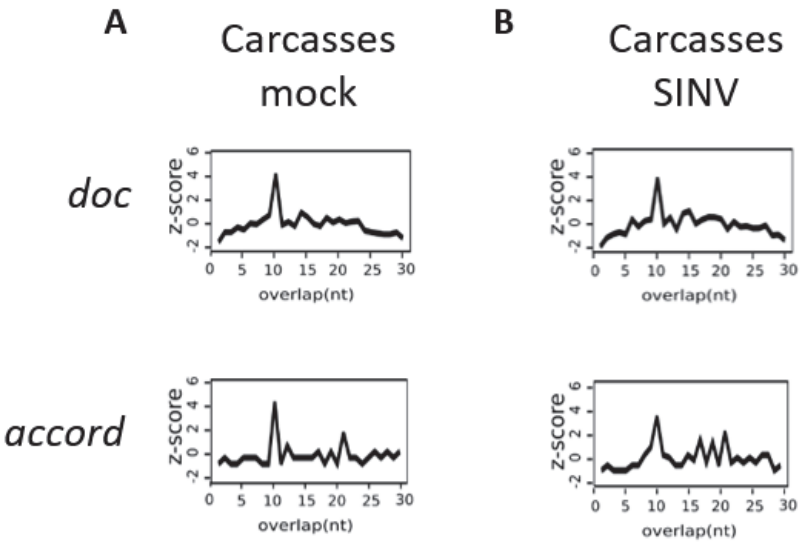


Figure 38. Signature ping-pong des piARN ciblant les ET dans les carcasses de *D. melanogaster* w^{1118}. Détection de la signature ping-pong pour les piARN ciblant *doc* et *accord* dans les carcasses non infectées **(A)** et dans les carcasses infectées par SINV **(B)**. Le z-score permet de standardiser les mesures entre les différentes conditions ; il correspond à la transformation de la variable « nombre de paires de lectures alignées ». Le chevauchement (*overlap*) de 10 nucléotides (nt) entre le brin sens et antisens des piARN est visible au niveau du pic à 10 nt. Ce pic est significatif si le z-score est supérieur à 2. n = 2 expériences indépendantes.

L'ensemble de ces résultats permet de fournir la première démonstration de la modulation des transcrits d'ET et des petits ARN interférents ciblant ces ET, lors d'une infection virale. Un résultat intéressant inclut également la présence de piARN dirigés contre les ET en somatique (c'est-à-dire dans les carcasses), impliquant une signature caractéristique des piARN secondaires ciblant les ET. Par la suite, dans le but de consolider et d'élargir les résultats, nous avons analysé l'infection de SINV en utilisant une autre espèce de drosophile (*D. simulans*) puis déterminé les modulations au niveau de l'activité des ET et de la production des petits ARN interférents lors de l'infection.

3. Infection par le virus Sindbis chez *D. simulans*

L'ensemble des résultats obtenus chez *D. melanogaster* a été comparé à ceux issus d'une lignée de *D. simulans* d'origine naturelle, appelée Makindu. L'activité et le nombre de copies de certains ET, tels que les rétrovirus endogènes (ERV) présentent une grande variabilité entre

les espèces. L'équipe de Cristina Vieira a notamment montré cela en utilisant l'ERV *tirant* chez *D. simulans* (Akkouche et al., 2012). De plus, la lignée Makindu est une des lignées qui a été caractérisée pour ses ET (Akkouche et al., 2013, 2012; Fablet et al., 2019). Par conséquent, nous allons déterminer l'impact de l'infection sur les transcrits d'ET et les petits ARN interférents ciblant les ET en contexte naturel, en utilisant une lignée très bien caractérisée d'une espèce proche de *D. melanogaster*.

A. Identification des infection chroniques et élimination de *Wolbachia* chez *D. simulans*

Nous avons révélé l'infection chronique du NV chez les drosophiles Makindu infectées par SINV et non infectées par SINV (mock), par RT-PCR (**Figure S6.**) et à partir de nos analyses de RNA-seq au niveau des carcasses (**Figure 39.**). Ces deux méthodes ne révèlent aucune infection par les virus DMelSV, DAV et DCV. Notons également que le nombre de transcrits de NV est très inférieur à celui de SINV dans les carcasses des Makindu (moyenne du log2 *fold change* = -2,88). De plus, de la même façon que chez *D. melanogaster* w^{1118}, le nombre de transcrits de SINV est supérieur dans les carcasses par rapport aux ovaires (**Figure 39.**).

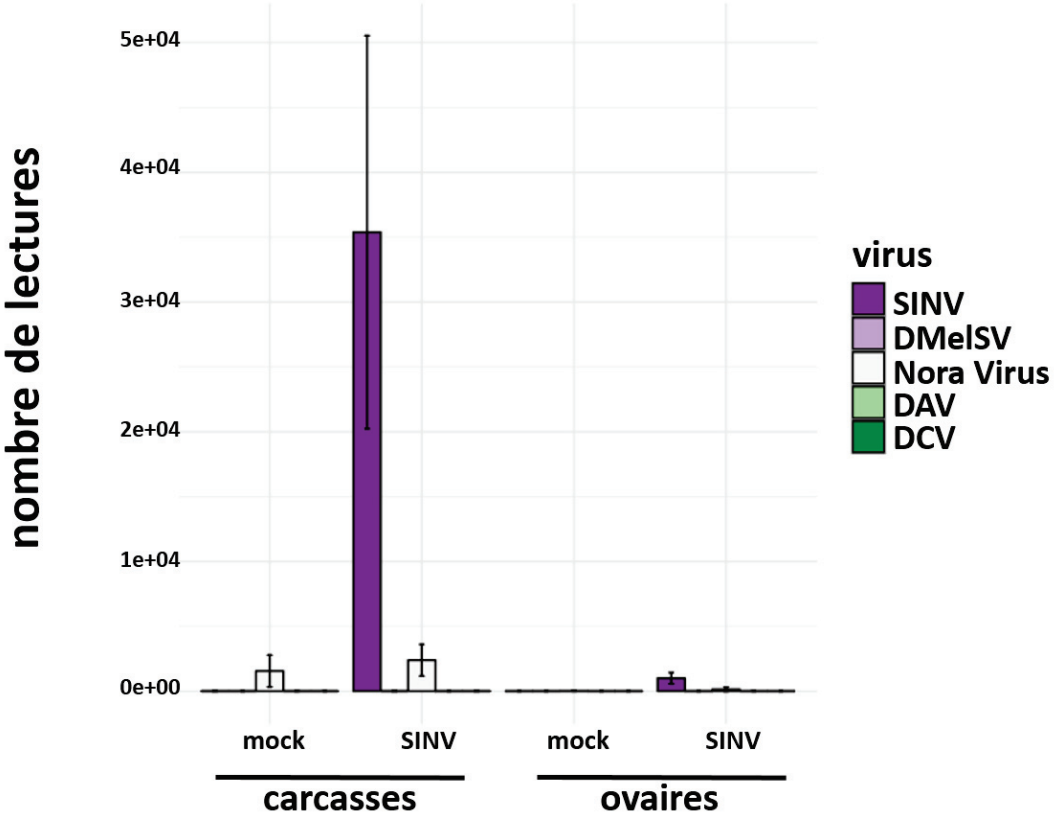


Figure 39. Détection des transcrits viraux chez (*D. simulans*) Makindu. Détermination du nombre de lectures (normalisation par DESeq2 selon la taille des banques) correspondant aux transcrits viraux par RNA-seq pour les

infections chroniques suivantes : DMelSV, NV, DAV, DCV, dans les tissus (carcasses et ovaires) issus de drosophiles infectées expérimentalement par SINV ou non infectées (mock). n = 2 expériences indépendantes (30 tissus par échantillon).

Par ailleurs, de la même façon que pour *D. melanogaster*, nous avons traité les drosophiles Makindu à la tétracycline pour éliminer la bactérie endosymbiotique *Wolbachia*, qui est connue pour interférer avec la réplication de nombreux virus (Monsanto-Hearne and Johnson, 2018; Teixeira et al., 2008). L'élimination de *Wolbachia* a été validée par RT-PCR (**Figure S7.**).

Dans cette partie nous avons détecté l'infection naturelle des Makindu par NV. Par la suite, la réplication de SINV a été explorée chez *D. simulans* Makindu afin de caractériser ce modèle d'un point de vue virologique, incluant la validation de la réplication virale dans ce modèle. En effet, l'analyse de l'infection virale est nécessaire avant d'étudier les modulations d'activité des ET et de production des petits ARN interférents sous-jacentes.

B. Etude de la réplication du virus Sindbis chez *D. simulans*, Makindu

Nous avons observé une très faible réplication de SINV chez les drosophiles Makindu, en utilisant la même dose infectieuse (2300 PFU par drosophile) que pour les expériences issues de *D. melanogaster*, à 6 jours après infection virale (**Figure 40.** ; médiane inférieure à 1 log10(TCID50/ml). Par conséquent, nous avons utilisé une dose infectieuse plus importante (14900 PFU par drosophile), pour l'ensemble des expériences avec les drosophiles Makindu, permettant d'obtenir des titres infectieux comparables à ceux obtenus chez *D. melanogaster* (mouches entières) au jour 6 après infection (**Figure 40.** ; médiane = 3,7 log10(TCID50/ml)). De plus, comme chez *D. melanogaster*, SINV se réplique principalement dans les carcasses des drosophiles et est très faiblement détecté dans les ovaires des Makindu au cours de la cinétique de réplication virale, c'est-à-dire du jour de l'infection à 10 jours après l'infection par SINV (**Figures 41.**). En effet, dans les ovaires, nous avons détecté les particules virales de SINV seulement pour quelques drosophiles au cours de la cinétique de réplication virale, représentant de 3% à 13% des drosophiles infectées. A l'inverse, dans les carcasses, on observe une augmentation accrue de la réplication virale de SINV jusqu'à 4 jours après l'infection (médiane au jour 4 = 2,6 log10(TCID50/ml) puis une phase plateau où la réplication de SINV ralentit et devient constante (médiane au jour 6 = 3 log10(TCID50/ml) ; médiane = 3,7 log10(TCID50/ml) aux jours 8 et 10 après infection). Ces résultats montrent que SINV se réplique chez les drosophiles Makindu. L'ensemble des expériences suivantes a encore une

fois été réalisé au jour 6 après infection, qui correspond à la phase exponentielle de réplication virale chez Makindu.

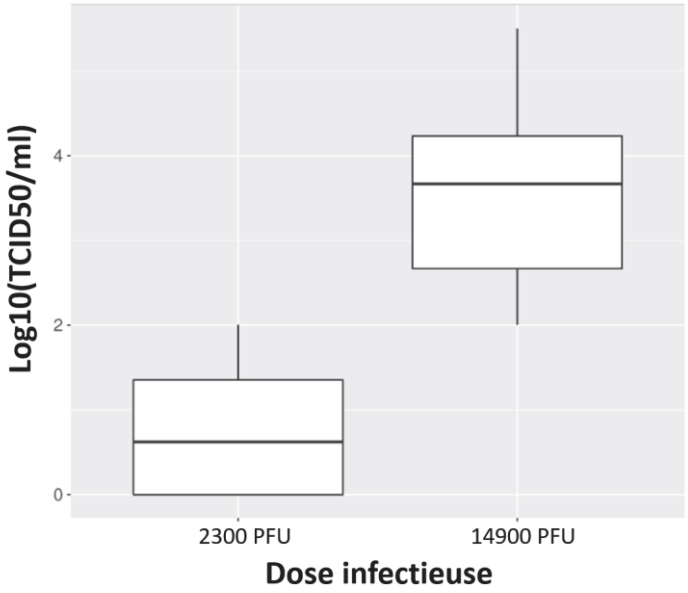


Figure 40. Réplication de SINV chez (*D. simulans*) Makindu à 6 jours après infection. Les drosophiles ont été infectées avec une dose de 2300 ou 14900 PFU de SINV par drosophile, dans un volume d'injection de 32,2 nl ou 207 nl, respectivement. Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de drosophiles infectées au jour 6 après infection. n = 1 expérience (10 drosophiles par condition).

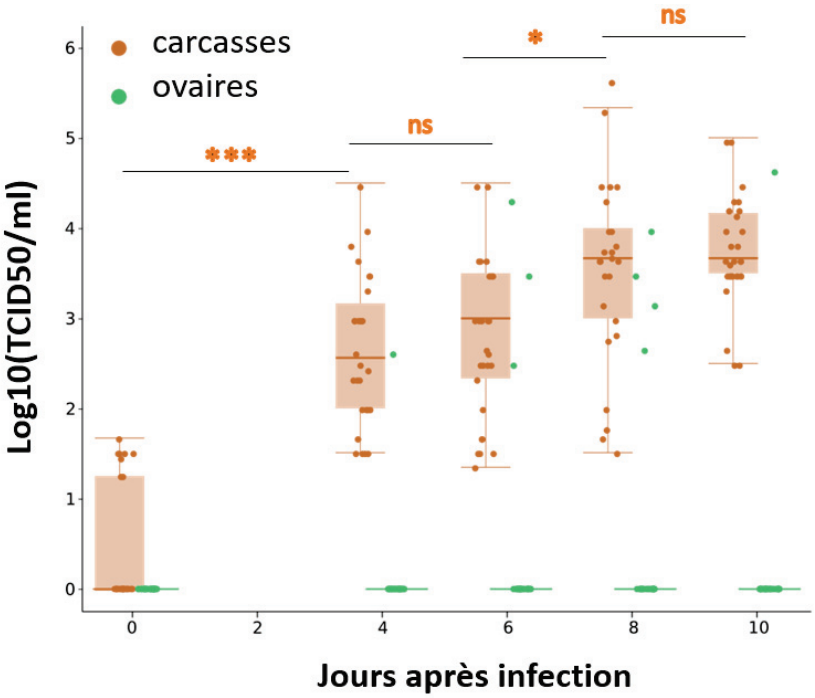


Figure 41. Cinétique de réplication de SINV chez (*D. simulans*) Makindu. Les drosophiles Makindu ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl. Les titres viraux de SINV

ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de carcasses et de paires d'ovaires, tout au long de la cinétique de réplication (du jour de l'infection à 10 jours après infection). n = 3 expériences indépendantes (10 drosophiles par condition). Test statistique de Wilcoxon-Mann-Whitney *; ns : non significatif ; *P* < 0,05 *; ***P* < 0,001.

De plus, le marquage par immunofluorescence de la capside de SINV confirme les résultats de titration virale, puisqu'aucun marquage spécifique n'a été détecté dans les ovaires au jour 6 après l'infection virale (**Figures 42.**). Le tropisme de SINV chez (*D. simulans*) Makindu a ensuite été étudié plus en détail au jour 6 après infection. Nous avons réalisé des coupes de drosophiles, et localisé les particules virales de SINV par immunofluorescence, en utilisant un anticorps anti-capside de SINV (**Figure 43.**). Ces analyses révèlent une détection de SINV principalement au niveau du thorax (**Figure 43 A et B.**) et des têtes (**Figure 43 A et C.**), soit proche de la zone d'injection du virus.

Dans cette partie, après avoir validé la réplication de SINV chez Makindu, nous avons localisé plus précisément les particules virales au niveau du thorax et de la tête des Makindu. Ultérieurement, les petits ARN interférents ciblant SINV ont été explorés dans les carcasses et les ovaires.

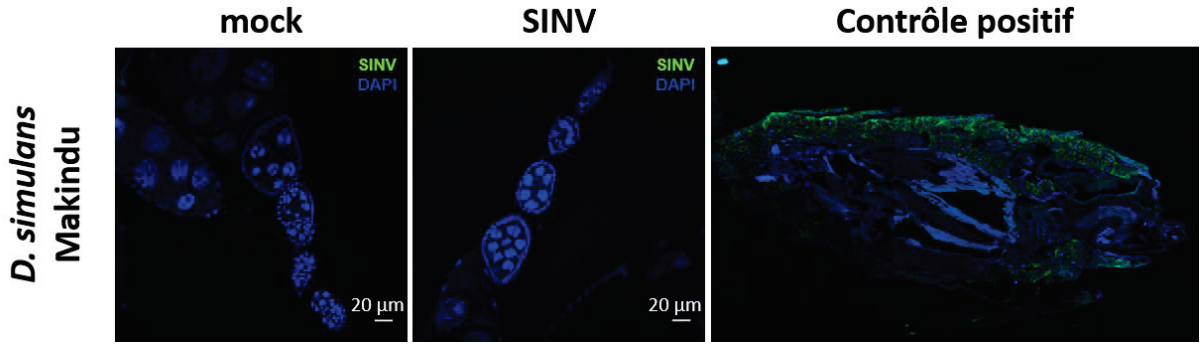
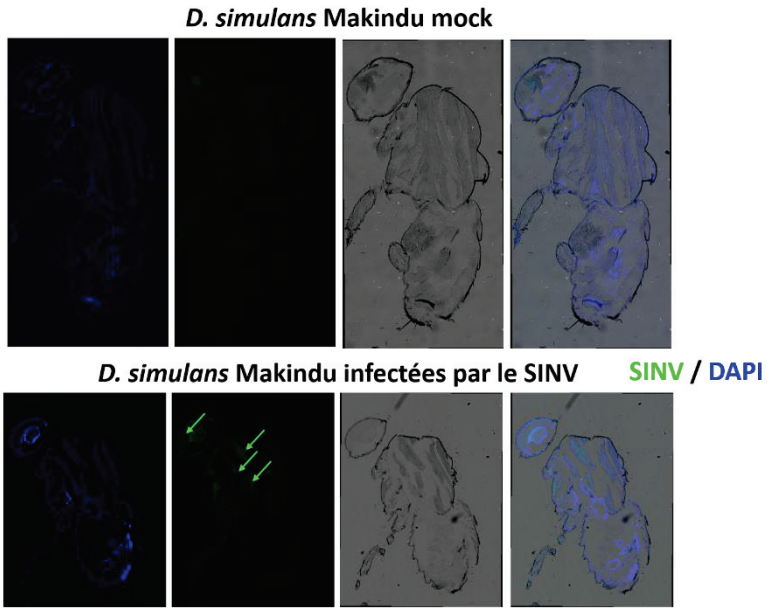
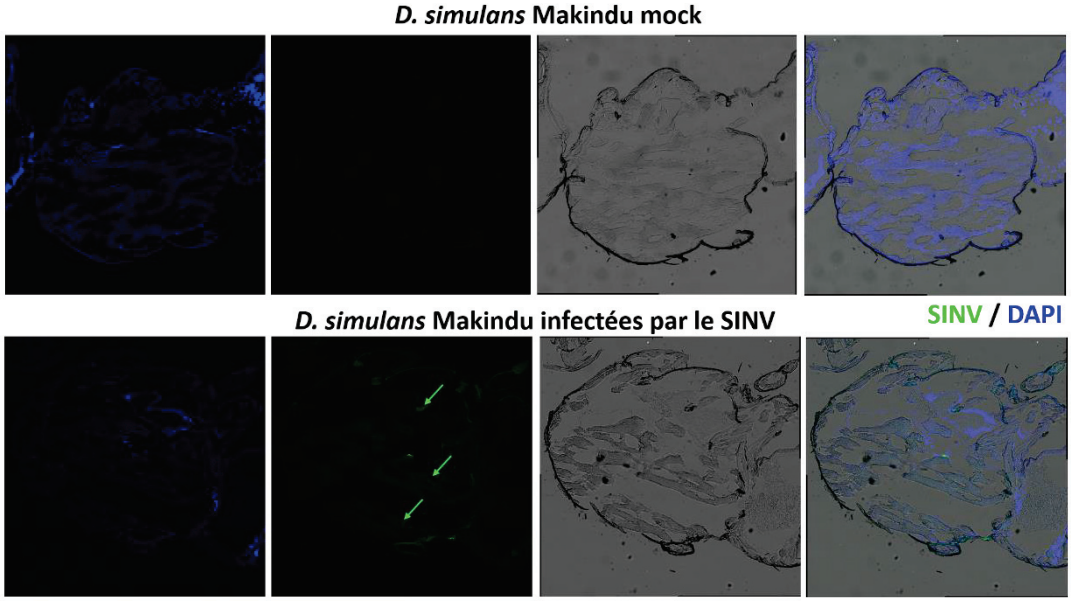


Figure 42. Absence de détection de SINV dans les ovaires de *D. simulans* Makindu. Les drosophiles ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl ; les paires d'ovaires ont été disséquées au jour 6 après infection. Le contrôle positif correspond à une coupe d'abdomen de drosophile infectée par SINV. La technique d'immunofluorescence a été utilisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488). L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 10 paires d'ovaires.

A


D. simulans Makindu mock
D. simulans Makindu infectées par le SINV
SINV / DAPI

B


D. simulans Makindu mock
D. simulans Makindu infectées par le SINV
SINV / DAPI

C

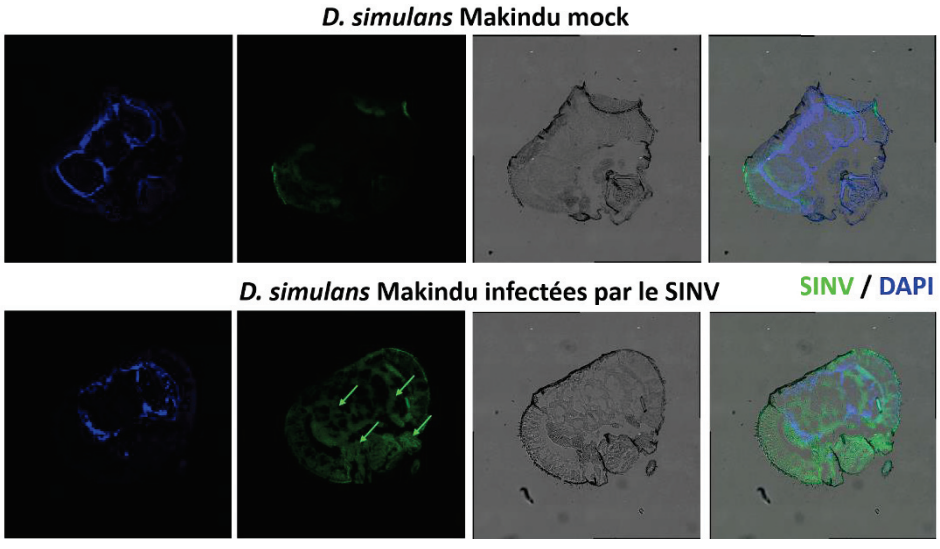

D. simulans Makindu mock
D. simulans Makindu infectées par le SINV
SINV / DAPI

Figure 43. SINV est principalement détecté au niveau du thorax et des têtes chez Makindu. Les drosophiles ont été infectées avec une dose de 14900 PFU de SINV, dans un volume d'injection de 207 nl. La technique d'immunofluorescence a été utilisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488) au jour 6 après l'infection. L'ADN a été marqué en bleu (DAPI). **(A)** Drosophiles Makindu entières ; **(B)** Thorax ; **(C)** Têtes. Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 3 drosophiles.

C. Modulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis chez *D. simulans*, Makindu

a) Caractérisation de la réponse de type ARNi lors de l'infection par le virus Sindbis (Makindu)

La réponse de type ARNi chez les drosophiles Makindu est comparable à celle observée chez *D. melanogaster* au jour 6 après infection par SINV. On retrouve une réponse de type ARNi, principalement médiée par les siARN (**Figure 44.**). Lors de l'infection par SINV, la production de siARN ciblant SINV présente une proportion équivalente de brin sens et antisens mais est plus importante dans les carcasses que dans les ovaires (**Figure 44 A.**). De plus, ces siARN sont distribués de façon homogène le long du génome viral (**Figure 44 B.**). L'analyse de ces résultats révèle également que la production des siARN dans les carcasses chez Makindu (**Figure 44 A.**) semble inférieure à celle des carcasses de *D. melanogaster* w^{1118} (**Figure 36 A.**). Par ailleurs, les piARN ciblant SINV ne sont pratiquement pas produits dans les carcasses et dans les ovaires, incluant principalement la détection des brins sens (**Figure 44 A.**).

L'analyse de l'infection des Makindu par SINV nous a permis de valider la réplication de SINV et la production de siARN viraux dans ce deuxième modèle, qui va maintenant être intéressant notamment pour sa comparaison avec les drosophiles w^{1118} à propos de la modulation de l'activité des ET et de la production des petits ARN interférents ciblant les ET lors de l'infection virale.

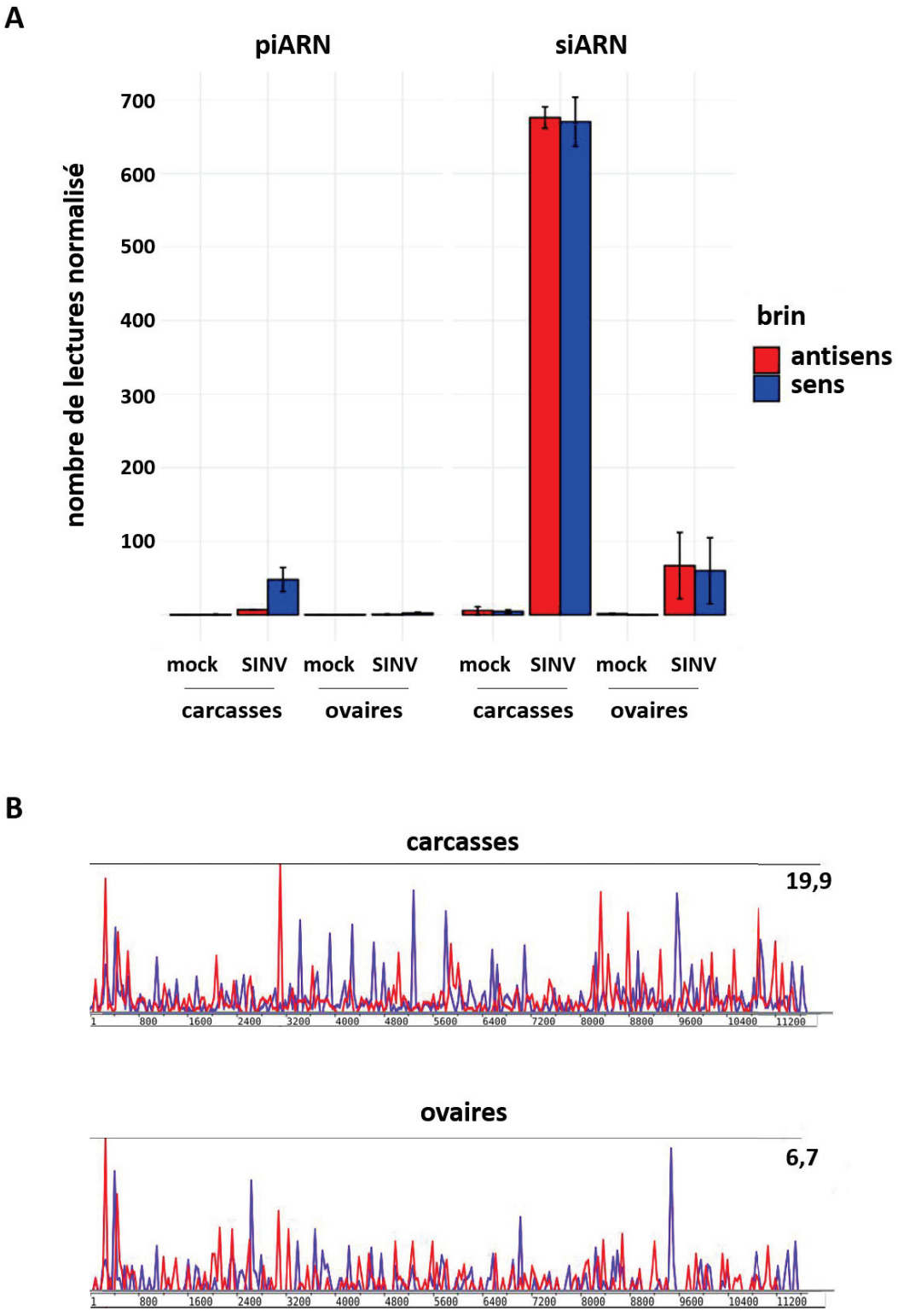


Figure 44. Analyse des siARN et des piARN ciblant SINV chez *D. simulans* Makindu. (A) Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles Makindu. L'ensemble des comptes des lectures a été normalisé par le nombre de lectures de miARN. **(B)** Distribution des petits ARN interférents de 21 nulcéotides ciblant SINV, le long du génome viral, incluant les brins sens (bleu) et antisens (rouge). La valeur

numérique en haut à droite correspond au nombre de lectures pour le pic le plus haut. n = 2 expériences indépendantes (50 tissus par échantillon).

b) Modulation de l'abondance des transcrits d'ET et de la production des piARN ciblant les ET chez *D. simulans* Makindu

Dans les carcasses, l'analyse des transcrits d'ET et des petits ARN interférents ciblant ces ET, lors de l'infection par SINV, à jour 6 après infection, révèle des résultats comparables à ceux observés chez *D. melanogaster*. Il s'agit d'une diminution significative des quantités de transcrits d'ET dans les carcasses (**Figure 45 A.** ; moyenne log2 *fold change* = -0,24 ; *p-value* = 7,6e-21), associée à une augmentation significative de la production des petits ARN interférents d'une longueur entre 23 et 30 nucléotides (piARN) (**Figure 45 B.** ; moyenne log2 *fold change* = 0,40 ; *p-value* = 1,9e-35) et des petits ARN interférents de 21 nucléotides (siARN) (**Figure 45 C.** ; moyenne log2 *fold change* = 0,42 ; *p-value* = 2,1e-33). Les résulats de RNA-seq ont été confirmés par RT-PCR quantitative (échantillons issus des mêmes expériences) pour un ensemble d'ET dérégulés et suffisamment exprimés pour être détectés, incluant *accord*, *gypsy*, *412*, *Doc6*, *Doc* et *F element* (**Figure 46.**). Le ratio de l'expression relative de chacun de ces gènes entre la condition infectée par SINV et la condition non infectée (mock) se situe entre -0,23 et - 0,76. L'expression de l'ERV *tirant* n'a pas été impactée lors de l'infection par SINV à partir des données de RNA-seq (évaluée par DESeq2 ; moyenne du log2 *fold change* = 0,14 ; *p-value* = 0,9) et a été validée en RT-qPCR, utilisé comme témoin négatif. De plus, comme dans les carcasses de *D. melanogaster* w^{1118}, nous avons détecté dans les carcasses des Makindu non infectées (**Figure 47 A.**) et infectées par SINV (**Figure 47 B.**) des piARN dirigés contre différentes familles d'ET, telles que *412* et *doc* présentant une signature « ping-pong » caractéristique des piARN secondaires impliqués dans la boucle d'amplification des piARN.

Cependant, dans les ovaires, les résultats sont contrastés par rapport à ceux décrits chez *D. melanogaster* où l'on a observé une faible diminution des niveaux de transcrits d'ET et des petits ARN interférents ciblant les ET lors de l'infection par SINV. Dans les ovaires des Makindu, aucune modulation significative de l'abondance des transcrits d'ET n'a été détectée (**Figure 45 D.** ; moyenne log2 *fold change* = -0,12 *; p-value* = 0,11), alors que la production des petits ARN interférents est complètement dérégulée, incluant une diminution significative de la production des petits ARN interférents de 23 à 30 nucléotides (**Figure 45 E.** ; moyenne log2 *fold change* = -1,92 ; *p-value* = 4,9e-37) et à l'inverse, une augmentation significative des petits ARN interférents de 21 nucléotides (**Figure 45 F.** ; moyenne log2 *fold change* = 0,60 ; *p-value* = 1,1e-30). Cette réponse en miroir entre les deux types de petits ARN interférents pourrait expliquer l'absence de modulation des quantités de transcrits d'ET. Les ovaires semblent protégés de la modulation de transcription d'ET chez *D. melanogaster*, où on avait un faible effet de l'infection par SINV et chez *D. simulans* où on n'a aucun effet.

Par ailleurs, les modulations au niveau des quantités de petits ARN interférents ciblant les ET n'impliquent pas de modulation du niveau de transcrits des gènes codant les principaux effecteurs de ces voies d'ARN interférence (évaluée par DESeq2 ; *p-value* > 0,05), tels que *piwi*, *aub*, *ago3*, *zuc* (**Figures S8 A. ;** voie des piARN) et *dcr2*, *ago2* (**Figures S8 B. ;** voie des siARN) lors de l'infection dans les carcasses et les ovaires.

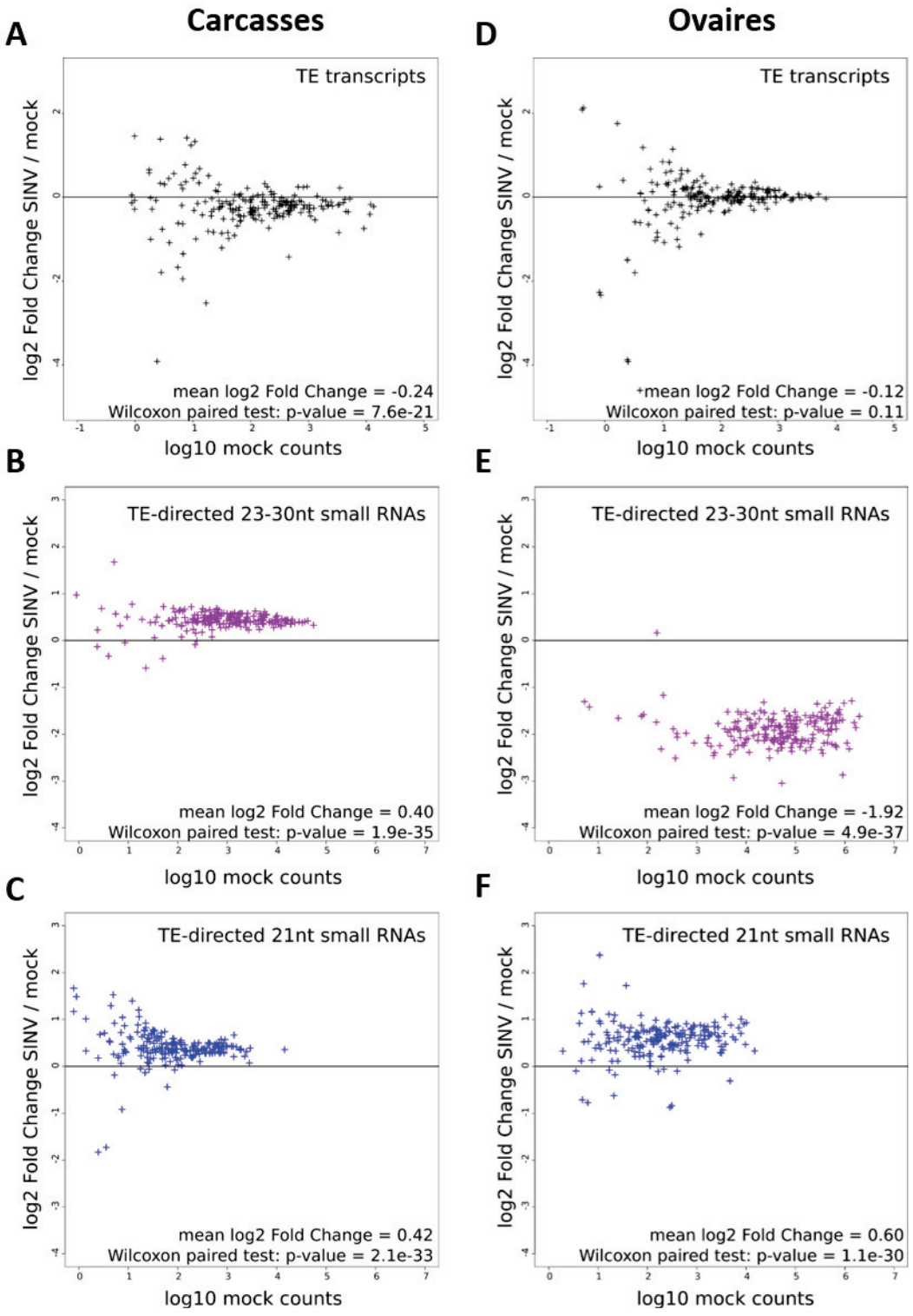


Figure 45. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez Makindu lors de l'infection par SINV dans les carcasses (A-C) et dans les ovaires (D-F). Chaque point représente le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par SINV et la condition non infectée (mock), pour les transcrits d'ET **(A, D)**, pour les piARN **(B, E)** et pour les siARN **(C, F)**. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux

comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). n = 2 expériences indépendantes.

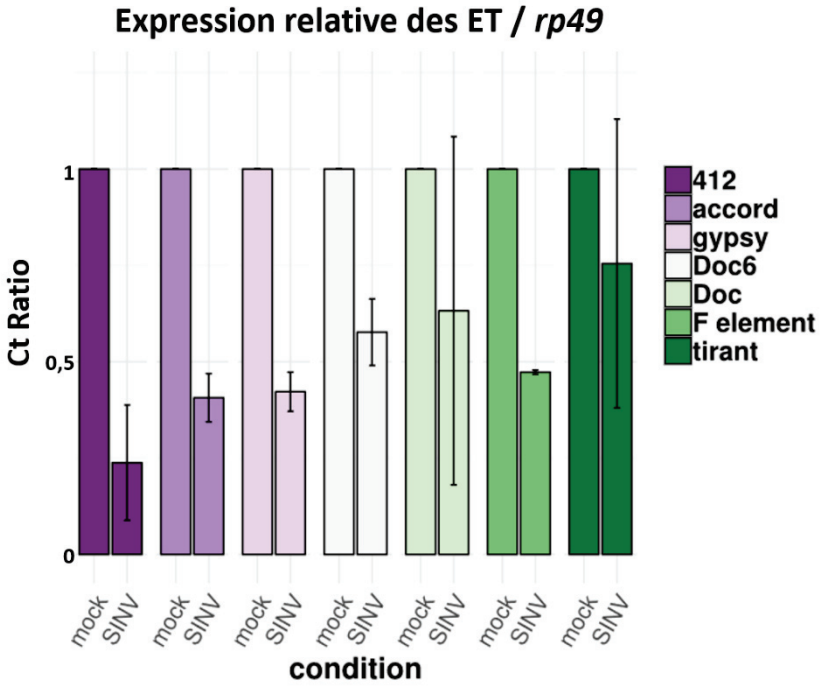


Figure 46. Diminution de l'expression des ET chez Makindu lors de l'infection par SINV. Les drosophiles ont été infectées avec une dose de 14900 PFU/ml de SINV dans un volume d'injection de 207 nl. Les niveaux d'expression des familles des ET *412, accord, gypsy, Doc6, Doc*, *F element* et *tirant* ont été déterminés par RT-PCR quantitative à partir des mêmes échantillons d'ARN que pour le RNA-seq. Les données sont représentées en expression relative par rapport au gène de ménage *rp49*. n = 2 expériences indépendantes (30 carcasses par échantillon d'ARN).

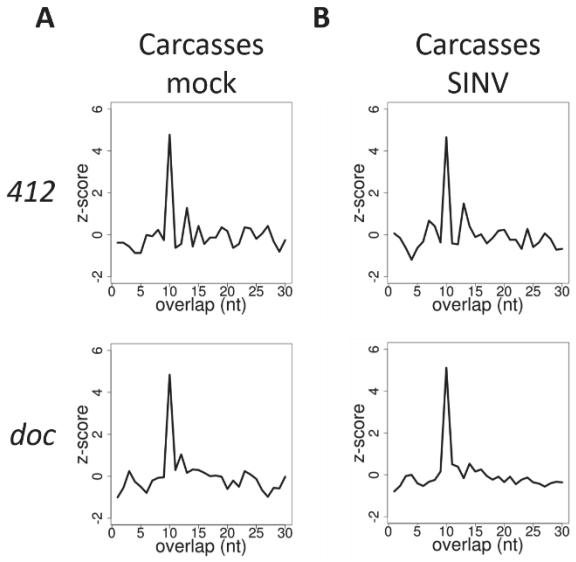


Figure 47. Signature ping-pong des piARN ciblant les ET dans les carcasses de *D.simulans* Makindu. Détection de la signature ping-pong pour les piARN ciblant *412* et *doc* dans les carcasses non infectées **(A)** et dans les

carcasses infectées par SINV **(B)**. Le z-score permet de standardiser les mesures entre les différentes conditions ; il correspond à la transformation de la variable « nombre de paires de lectures alignées ». Le chevauchement (*overlap*) de 10 nucléotides (nt) entre le brin sens et antisens des piARN est visible au niveau du pic à 10 nt. Ce pic est significatif si le z-score est supérieur à 2. n = 2 expériences indépendantes.

Enfin, l'analyse globale du transcriptome et notamment de l'ensemble des gènes différentiellement exprimés, au niveau des carcasses des Makindu, nous a permis de visualiser les processus biologiques dérégulés lors de l'infection par SINV. Nous avons réalisé des analyses de « Gene Ontology (GO)», qui permettent de déterminer s'il existe un enrichissement en certaines fonctions parmi les gènes différentiellement exprimés par rapport à des gènes pris aléatoirement selon différents niveaux, dans de nombreuses voies de la biologie moléculaire et cellulaire, pour une espèce donnée (Harris et al., 2004). Ces analyses révèlent que parmi les gènes différentiellement exprimés, on trouve un enrichissement en gènes associés au métabolisme de l'ARN lors de l'infection par SINV par rapport à la condition non infectée (**Figure S9.**).

c) Caractérisation de la réponse de type ARNi vis-à-vis de l'infection naturelle des Makindu par NV

Les résultats obtenus chez *D. melanogaster* w^{1118} et *D. simulans* Makindu sont comparables dans les carcasses, ce qui renforce nos observations et ainsi l'impact de l'infection par SINV sur les ET et leurs voies de contrôle génomique, qui ne semble pas dépendant de l'espèce de drosophile. Les infections chroniques détectées chez les drosophiles Makindu et plus particulièrement l'infection par NV ne semblent pas perturber l'effet de SINV. NV est un virus persistant, fréquemment retrouvé chez les drosophiles de laboratoire et les drosophiles sauvages (Habayeb et al., 2006; Webster et al., 2015). Lors d'une infection persistante, chaque virus se réplique au niveau de cellules spécifiques et est contrôlé par la voie des siARN qui pourrait être stimulée à cette occasion (Goic et al., 2013). Par conséquent, nous avons analysé les petits ARN interférents ciblant l'ensemble des virus et nous avons détecté une production très importante de petits ARN interférents d'une longueur de 21 nucléotides (siARN) ciblant NV dans les carcasses et dans les ovaires des Makindu (**Figure 48 A.**). La quantité de siARN ciblant NV est nettemment supérieure à celle ciblant SINV, injecté expérimentalement, dans les carcasses (moyenne du log2 *fold change* = 4,64) et dans les ovaires (moyenne du log2 *fold change* = 3,24), lors de l'infection par SINV. De plus, en condition infectée par SINV ou en condition non infectée, ces siARN présentent une proportion comparable des brins sens et antisens dans les carcasses et les ovaires (**Figure 48 B.**).

Nous avons également identifié des petits ARN interférents de 23 à 30 nucléotides ciblant NV dans les carcasses et les ovaires, incluant la condition infectée par SINV ou la condition non infectée (**Figure 48A.**) mais ne présentant pas de signature caractéristique des piARN impliqués dans la boucle d'amplification (données non montrées). De plus, l'analyse bioinformatique de ces petits ARN interférents de 23 à 30 nucléotides, ciblant NV révèle une proportion beaucoup plus importante de brins sens par rapport aux brins antisens, dans les carcasses et les ovaires infectés ou non par SINV (**Figure 48 B.**). Ces résultats sont en accord avec la littérature, en drosophile entière (Petit et al., 2016) et au niveau des cellules somatiques ovariennes (H. Li et al., 2009; Malone et al., 2009; Wu et al., 2010).

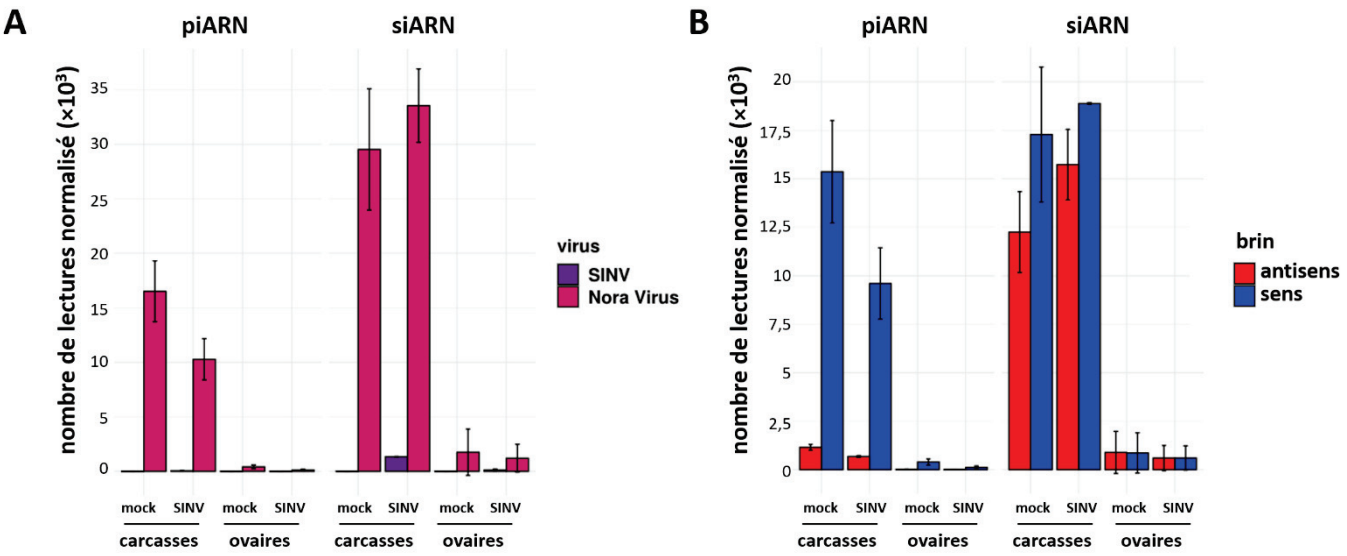


Figure 48. Analyse des siARN et des piARN ciblant NV chez Makindu lors de l'infection expérimentale par SINV. **(A)** Nombre de lectures des petits ARN interférents viraux (SINV, NV) de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des Makindu. **(B)** Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents ciblant NV. L'ensemble des comptes des lectures a été normalisé par le nombre de lectures des miARN. n = 2 expériences indépendantes (50 tissus par échantillon).

Dans ce chapitre, nous avons révélé dans les carcasses infectées par SINV une diminution des quantités de transcrits des ET associée à une augmentation de production des petits ARN interférents (siARN et piARN) chez *D. simulans* Makindu. Ces résultats sont en accord avec ceux obtenus chez *D. melanogaster* w^{1118} lors de l'infection virale et permettent de consolider les résultats. De plus, en somatique, on retrouve des piARN dirigés contre les ET qui présentent la signature spécifique des piARN secondaires, impliqués dans la boucle d'amplification en ping-pong. On note également que la modulation des petits ARN interférents n'implique pas de modulation au niveau des transcrits des gènes codant les effecteurs principaux de ces voies d'ARNi chez les drosophiles w^{1118} et Makindu. Par ailleurs, l'infection chonique par NV détectée dans les Makindu ne permet pas d'affirmer une non-spécificité d'espèce de ces

modulations lors de l'infection virale. Néanmoins, ces résultats révèlent des modulations comparables entre deux espèces de drosophile au niveau des carcasses, où SINV se réplique, ce qui renforce nos analyses. L'ensemble des petits ARN interférents viraux (SINV et NV) détectés est très peu produit et présente une quantité supérieure de brin sens par rapport au brin antisens. Par la suite, l'étude des mécanismes sous-jacents impliqués dans les modulations de l'activité des ET et de la production des petits ARN interférents a été effectuée chez *D. melanogaster*. Ainsi, nous avons d'abord déterminé si la voie de l'absorption de l'ARN double brin est impliquée dans les modulations observées, concernant les quantités de petits ARN interférents ciblant les ET et leur impact sur la transcription des ET.

4. Mécanismes impliqués dans la dérégulation des voies des petits ARN interférents et de la transcription des éléments transposables lors de l'infection par le virus Sindbis

Dans le but de décrypter les mécanismes impliqués dans la modulation de la transcription des ET et la production des petits ARN interférents spécifiques de ces ET, nous avons poursuivi les expériences en utilisant le modèle *D. melanogaster*, incluant de nombreux mutants disponibles dont deux que nous avons utilisés. Il s'agit des mutants $CG4572^{c05963}$, déficients pour la machinerie de l'import d'ARN double brin (Saleh et al., 2009, 2006), et des mutants $Dcr2^{[L811fsX]}$ défaillants pour la production de siARN (Lee et al., 2004). L'ensemble de ces drosophiles a été traité à la tétracycline pour s'affranchir de la bactérie endosymbiotique *Wolbachia* (**Figure S2.**) et les œufs ont été javellisés pour éliminer les particules virales provenant d'infections chroniques (**Figure S1.**) (Merkling and van Rij, 2015).

A. Impact de la voie d'absorption d'ARN double brin sur la dérégulation des ET et des petits ARN interférents ciblant les ET

Il a été précédemment démontré que la réponse antivirale des siARN dépendait d'une diffusion systémique d'intermédiaires viraux à l'état d'ARN double brin, qui ensuite engendrait une production de siARN chez *D. melanogaster* (Saleh et al., 2009). Cette diffusion systémique implique la voie de l'import d'ARN double brin, incluant l'entrée de l'ARN double brin dans les cellules, et diffusant jusqu'aux cellules non infectées. La protéine codée par *CG4572*, qui est une sérine-carboxypeptidase, est un effecteur de cette voie (Saleh et al., 2009). En effet, cette protéine est un composant important dans la mise en place d'une réponse antivirale optimale chez *Drosophila* puisqu'elle est impliquée dans le transport d'ARN double brin permettant une diffusion systémique de siARN, jusqu'aux tissus non infectés (Saleh et al., 2009). En effet, les mutants $CG4572^{c05963}$ ne sont pas capables d'engendrer une

réponse antivirale systémique lors de l'infection par SINV (Karlikow et al., 2014). Il a été décrit que cette protéine, appelée aussi DORA est localisée dans des compartiments endosomaux tardifs (Karlikow et al., 2016). Ces mêmes auteurs montrent que DORA est également retrouvée à l'intérieur de vésicules contenant des fragments d'ARN viraux, qui sont capables d'immuniser des drosophiles sauvages contre l'infection, en présence de DORA.

Dans un premier temps, nous avons validé le génotype des mutants *CG4572*^{c05963} en sélectionnant les drosophiles présentant le phénotype attendu, incluant des yeux fortement colorés et une absence de soie au niveau des « épaules du thorax » et correspondant à l'insertion du transposon *PiggyBac*, contenant le marqueur *white*[+] au sein du cadre de lecture ouvert du gène *CG4572* pour chaque allèle (**Figure S10.**). Nous avons ensuite réalisé les mêmes expériences que pour la lignée *w*^{1118}, qui correspond au fond génétique des mutants *CG4572*^{c05963}.

a) Etude de la réplication du virus Sindbis chez *D. melanogaster CG4572*^{c05963}

De la même façon que pour les drosophiles *D. melanogaster w*^{1118}, nous avons infecté les femelles *CG4572*^{c05963}, par injection intrathoracique, avec 2300 PFU de SINV par drosophile. De plus, nous n'avons pas observé de mortalité suite à l'infection par SINV chez les mutants *CG4572*^{c05963} (**Figure 49.** ; *w*^{1118} et *CG4572*^{c05963} = 6% de mortalité, correspondant aux conditions d'élevage classiques). Ces résultats sont en accord avec ceux des travaux de thèse de (M. Karlikow, 2016), pour ce qui est de l'infection par SINV. A l'inverse, lors des infections par les virus ARN de polarité positive et non-enveloppés comme DCV, FHV, et CrPV, les mutants *CG4572*^{c05963} sont hypersensibles à ces virus et présentent une durée de vie plus courte que chez les *w*^{1118} ou même que chez les mutants *Dcr2*$^{[L811fsX]}$ (M. Karlikow, 2016). Par ailleurs, il a été montré que les mutants *CG4572*^{c05963} sont hypersensibles à l'infection par SINV lorsque pour ces derniers les infections virales chroniques naturelles n'ont pas été éliminées (Saleh et al., 2009).

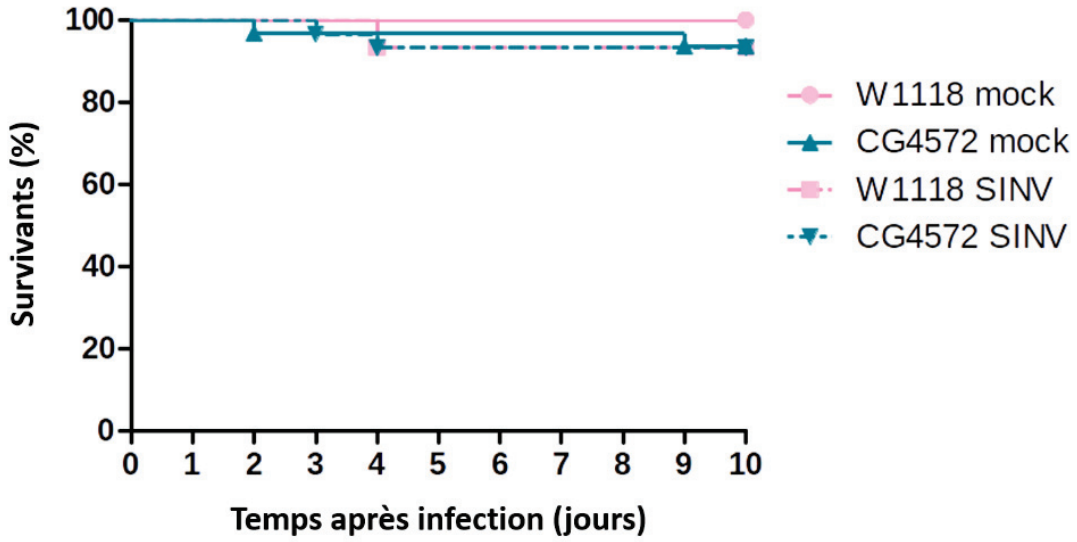


Figure 49. Absence de différence de survie entre les mutants *CG4572^{c05963}* et les drosophiles contrôles *w^{1118}*. Les drosophiles ont été infectées avec une dose de 2300 PFU de SINV, dans un volume d'injection de 32,2 nl. Les drosophiles vivantes ont été quantifiées tout au long de la cinétique de réplication (du jour de l'infection à 10 jours après infection). Les courbes en bleu représentent les mutants *CG4572^{c05963}*, les courbes en rose représentent les drosophiles contrôles *w^{1118}*. Les drosophiles non infectées, mock (traits pleins) et infectées par SINV (traits en pointillés). n = 3 expériences indépendantes (15 drosophiles par condition).

De plus, les résultats de réplication virale sont en accord avec les tests de survie puisqu'aucune différence significative n'a été observée entre les mutants *CG4572^{c05963}* et les drosophiles contrôles *w^{1118}* (**Figure 50.** ; test de Wilcoxon-Mann-Whitney ; *p-value* > 0,05). De la même façon que pour les drosophiles *w^{1118}*, l'infection de SINV chez les mutants *CG4572^{c05963}* inclut une phase exponentielle de réplication virale jusqu'à 4 jours après l'infection (médiane au jour 2 = 2,2 log10(TCID50/ml) ; médiane au jour 4 = 3,1 log10(TCID50/ml) puis une phase plateau où la réplication de SINV ralentit et devient constante (médiane au jour 6 = 3,5 log10(TCID50/ml) ; médiane = 3,2 log10(TCID50/ml) aux jours 8 et 10 après infection). Comme précédemment, les drosophiles ont été disséquées pour analyser l'influence de l'infection par SINV sur la production des petits ARN interférents et les quantités de transcrits des ET dans les carcasses et dans les ovaires. Aucune différence significative de réplication virale de SINV n'a été observée entre les drosophiles contrôles *w^{1118}* et les mutants *CG4572^{c05963}* (test de Wilcoxon-Mann-Whitney ; *p-value* > 0,05) au niveau des carcasses (**Figure 51 A.** ; légende marron) et des ovaires (**Figure 51 B.** ; légende kaki) au jour 6 après l'infection. De plus, encore une fois, la capside de SINV n'a pas été détectée au niveau des ovaires des mutants *CG4572^{c05963}* (**Figure S11.**).

L'ensemble de ces résultats révèle une réplication comparable de SINV entre les drosophiles contrôles *w^{1118}* et les mutants *CG4572^{c05963}*, n'induisant pas de mortalité suite à l'infection

virale. Par la suite, nous avons exploré la production des petits ARN interférents ciblant SINV et identifié les points de divergence entre ces deux lignées.

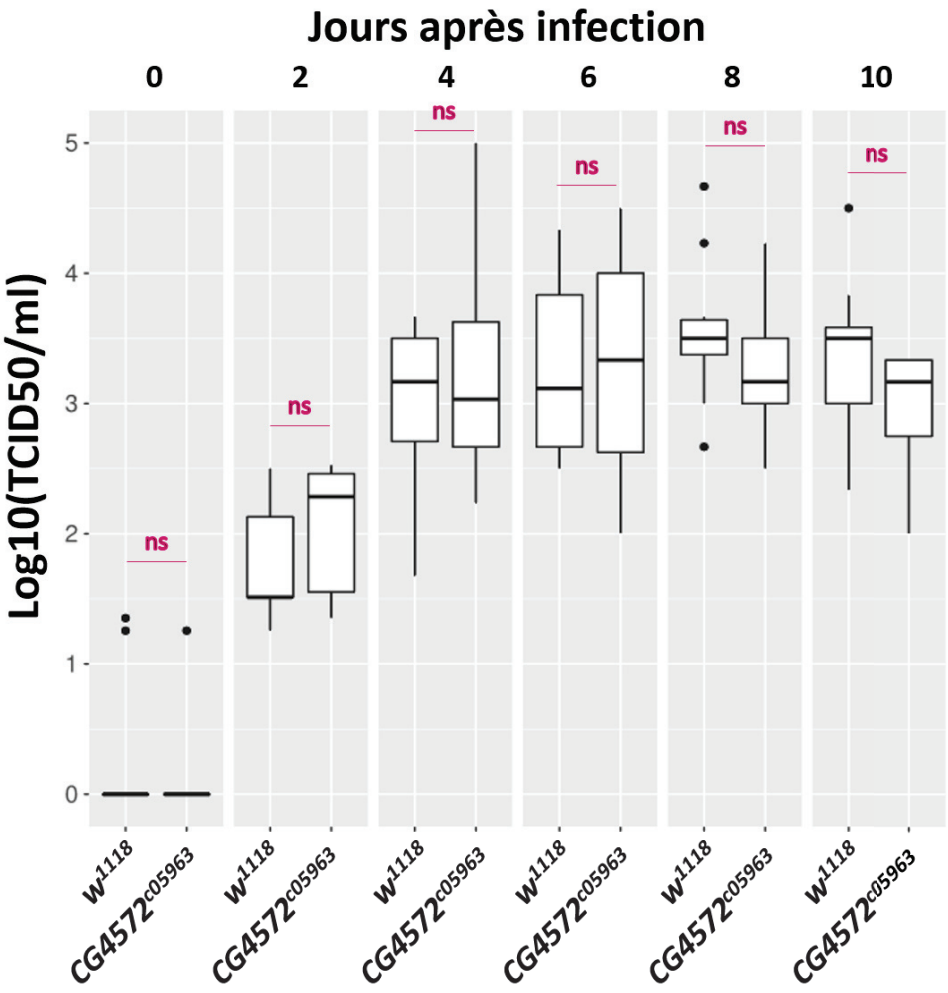


Figure 50. Absence de différence de réplication entre les mutants *CG4572*^{c05963} par rapport aux contrôles *w*^{1118}. Les drosophiles *w*^{1118} et *CG4572*^{c05963} ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de drosophiles infectées, tout au long de la cinétique de réplication (du jour de l'infection à 10 jours après infection). n = 2 expériences indépendantes (5 drosophiles par condition). Test statistique de Wilcoxon-Mann-Whitney ; ns : non significatif ; (*P < 0,05).

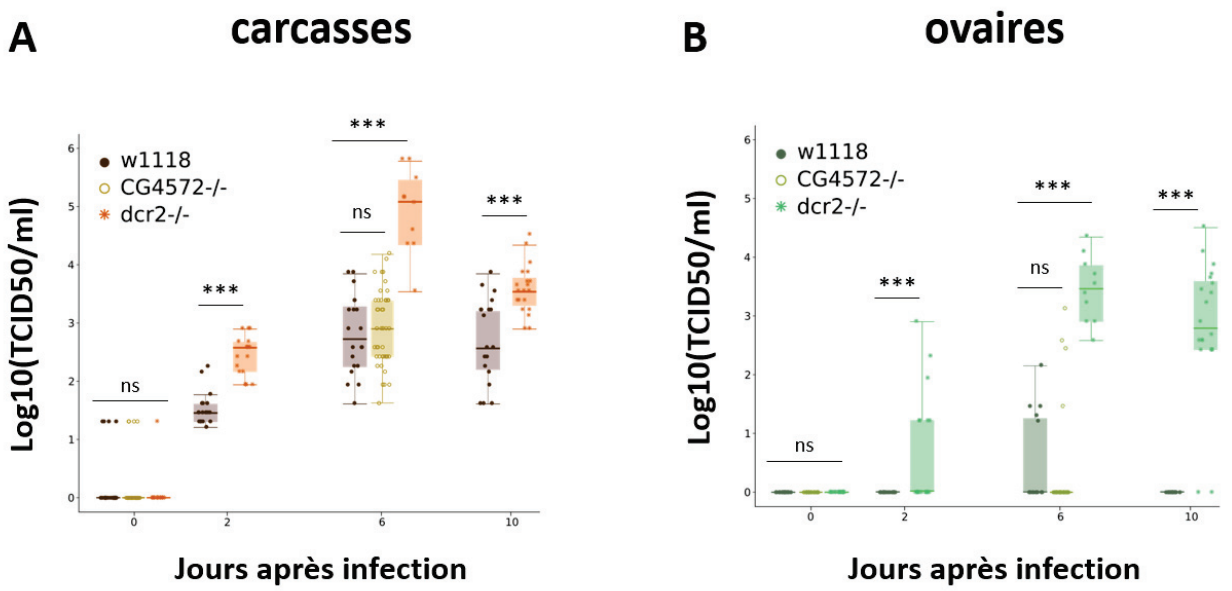


Figure 51. Réplication de SINV chez *D. melanogaster* w^{1118}, $CG4572^{c05963}$ et $Dcr2^{[L811fsX]}$. Chaque drosophile a été infectée avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de carcasses **(A)** et de paires d'ovaires **(B)** issues de drosophiles infectées. n = 2 expériences indépendantes (10 drosophiles par condition). Test statistique de Wilcoxon-Mann-Whitney ; ns : non significatif ; ***P < 0,001.

b) Modulation de la production des petits ARN interférents ciblant le virus Sindbis chez *D. melanogaster* $CG4572^{c05963}$

Comme attendu, la réponse de type ARN interférence chez les mutants $CG4572^{c05963}$ révèle une plus faible production des siARN ciblant SINV par rapport aux drosophiles w^{1118}, dans les carcasses (moyenne du log2 *fold change* = -1,51) et les ovaires (moyenne du log2 *fold change* = -2,02), à 6 jours après l'infection (**Figure 52 A.**). Chez les mutants $CG4572^{c05963}$, ces siARN sont moins nombreux mais présentent toujours des proportions équivalentes de brins sens et antisens dans les carcasses et les ovaires (**Figure 52 A**.). En outre, nous avons également observé une distribution homogène des siARN ciblant SINV le long du génome viral (**Figure 52 B.**). De plus, une infime production de piARN viraux a été détectée, incluant très majoritairement la détection de brin sens, ce qui est comparable à ce qui avait été observé chez les w^{1118} (**Figure 52 A.**). Par ailleurs, nous avons détecté une quantité non négligeable de siARN ciblant DAV dans un des deux réplicats des mutants $CG4572^{c05963}$ (**Figure S12.**) alors que nos expériences de RT-PCR n'ont révélé aucun transcrit de DAV (**Figure S1 B.**), suggérant un contrôle efficace par les siARN. Pour autant, aucune différence n'a été observée pour l'ensemble des analyses entre ces deux réplicats.

Nous avons validé le modèle *CG4572*^{c05963}, présentant notamment une faible production de siARN viraux lors de l'infection par SINV. Nous avons ensuite analysé l'impact de l'invalidation de la voie de l'import de l'ARN double brin (mutants *CG4572*^{c05963}) par rapport à la modulation de la production des petits ARN interférents observée lors de l'infection virale dans la lignée *w*^{1118}.

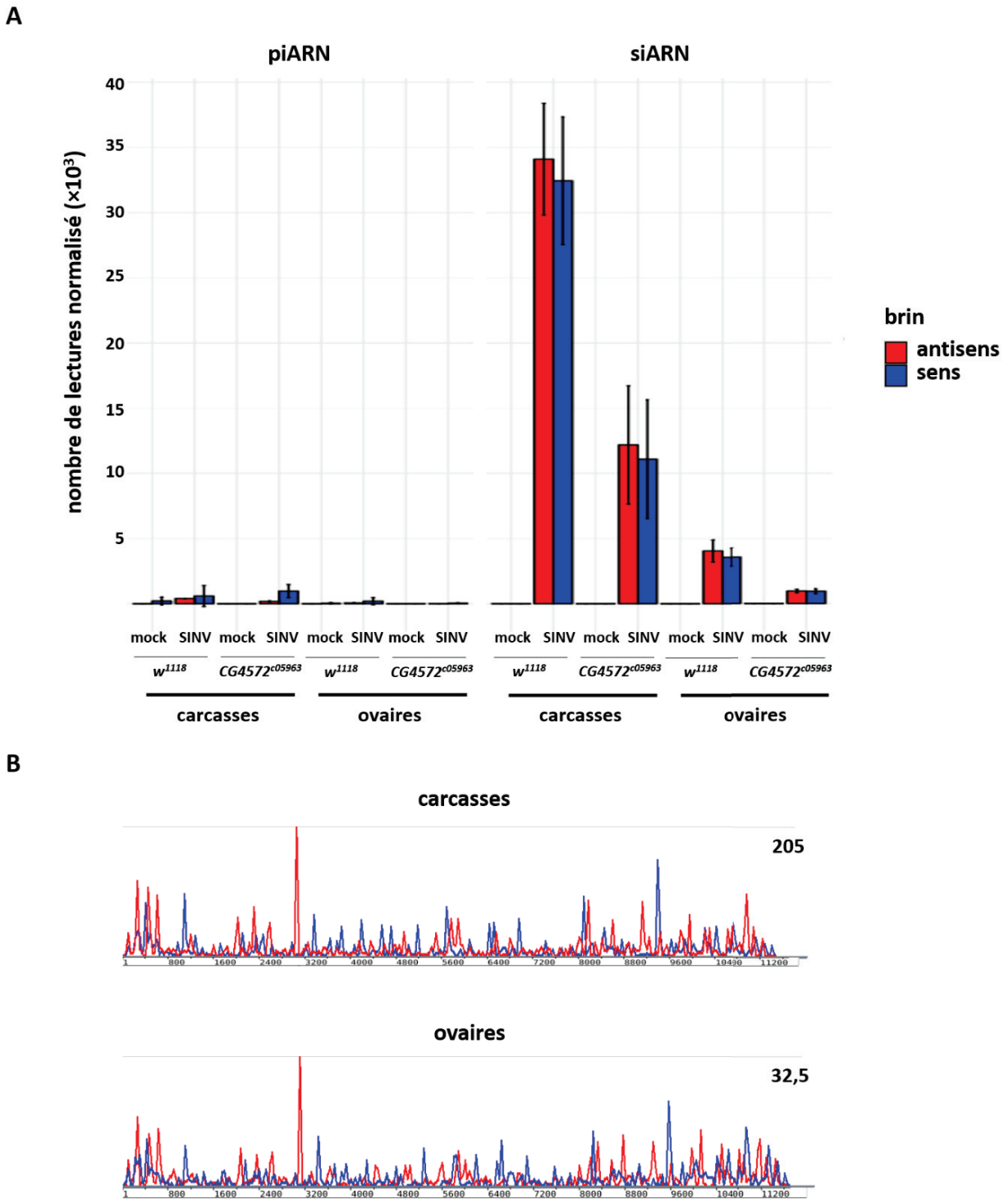


Figure 52. Analyse des siARN et des piARN ciblant SINV chez *D. melanogaster w*^{1118} *et CG4572*^{c05963}. (A) Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux de 21 nucléotides (siARN)

et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles sauvages w^{1118} et des mutants $CG4572^{c05963}$. L'ensemble des lectures a été normalisé par les comptes des lectures des miARN. **(B)** Distribution des petits ARN interférents de 21 nulcéotides ciblant SINV, le long du génome viral, incluant les brins sens (bleu) et antisens (rouge) chez les mutants $CG4572^{c05963}$. La valeur numérique en haut à droite correspond au nombre de lectures pour le pic le plus haut. n = 2 expériences indépendantes (50 tissus par échantillon).

c) Modulation de la production des petits ARN interférents ciblant les ET chez *D. melanogaster* $CG4572^{c05963}$

L'analyse des petits ARN interférents (Small RNA-seq) -ciblant les ET- chez les mutants $CG4572^{c05963}$ a été réalisée selon le même protocole que pour les drosophiles w^{1118}. La lignée $CG4572^{c05963}$ contient les mêmes ET que la lignée w^{1118}.

Carcasses :

Cette analyse a révélé la quasi-disparition de la modulation des quantités des petits ARN interférents de 23-30 nucléotides (**Figure 53 A.**) et des petits ARN interférents de 21 nucléotides (**Figure 53 B.**), suite à l'infection par SINV dans les carcasses des mutants $CG4572^{c05963}$ par rapport aux drosophiles sauvages (w^{1118}). En effet, ces observations incluent une diminution de 84 % et 66% de la moyenne du log2 du ratio de la quantité respective des piARN et des siARN, entre la condition infectée et la condition non infectée. Ces modulations restent significatives mais révèlent un très faible effet sur la production des piARN (**Figure 53 A.** ; points rouges ; moyenne log2 *fold change* = 0,10 ; *p-value* = 3,0e-7) et des siARN (**Figure 53 B.** ; points rouges ; moyenne log2 *fold change* = 0,21 ; *p-value* = 1,1e-6). De plus, lorsqu'on compare les mutants $CG4572^{c05963}$ aux drosophiles w^{1118}, on a observé une augmentation significative mais modérée de la production des piARN dans un contexte non infectieux (**Figure S13 A.** ; moyenne log2 *fold change* = 0,12 ; *p-value* = 4,3e-3) et une diminution significative des piARN lors de l'infection par SINV (**Figure S13 B.** ; log2 *fold change* = -0,40 ; *p-value* = 6,7e-15). Tandis que pour les siARN, de la même façon que pour les siARN ciblant SINV, une diminution significative de la production des siARN a été identifiée en contexte non infectieux (**Figure S13 C.** ; moyenne log2 *fold change* = -0,80 ; *p-value* = 5,1e-19) et lors de l'infection par SINV (**Figure S13 D.** ; moyenne log2 *fold change* = - 1,19 ; *p-value* = 6,1e-24). Ces résultats suggèrent que l'augmentation des petits ARN interférents ciblant les ET, associée à une diminution des quantités de transcrits d'ET, observée dans les carcasses w^{1118} suite à l'infection par SINV, implique la voie de l'import des ARN double brin.

Ovaires :

A l'inverse, dans les ovaires, il n'y a quasiment pas de différence entre les mutants *CG4572*^{c05963} et les drosophiles sauvages *w*^{1118}. En effet, de la même façon que chez *w*^{1118}, nous observons une légère diminution des petits ARN interférents de 23-30 nucléotides (**Figure 53 C.** ; points rouges ; moyenne log2 fold change = -0,28 ; p- value = 3,2e-37) et de 21 nucléotides (**Figure 53 D.** ; points rouges ; moyenne log2 fold change = -0,15 ; p- value = 8,0e-30). Ces résultats suggèrent que la voie de l'import des ARN double brin ne détermine pas les quantités de petits ARN interférents dirigés contre les ET dans les ovaires lors de l'infection par SINV. Par ailleurs, lorsqu'on compare les mutants *CG4572*^{c05963} aux drosophiles *w*^{1118}, une augmentation significative de la production des piARN dirigés contre les ET a été révélée en condition non infectée (**Figure S14 A.** ; moyenne log2 *fold change* = 0,65 ; *p-value* = 2,4e-25) et lors de l'infection par SINV (**Figure S14 B.** ; moyenne log2 *fold change* = 0,58 ; *p-value* = 3,1e-22). De plus, de la même façon que pour les siARN ciblant SINV, nous avons détecté une diminution de la production des siARN dirigés contre les ET dans un contexte non infectieux (**Figure S14 C.** ; moyenne log2 *fold change* = -0,52 ; *p-value* = 7,1e-17) et lors de l'infection par SINV (**Figure S14 D.** ; moyenne log2 *fold change* = - 0,52 ; *p-value* = 1,2e-16). Ainsi, nous avons observé un effet opposé de l'invalidation de la voie de l'import sur la production des petits ARN interférents dirigés contre les ET, incluant une augmentation des piARN et une diminution des siARN. L'ensemble de ces résultats suggère que la légère diminution des petits ARN interférents ciblant les ET observée dans les ovaires des *w*^{1118} n'implique pas la voie d'import de l'ARN double brin.

Dans cette partie, la voie de l'import de l'ARN double brin ne semble pas avoir la même implication au niveau des modulations de production des petits ARN interférents observées dans les carcasses et les ovaires. En effet, seules les carcasses ne présentent pas le même profil de modulation des quantités de petits ARN interférents, suite à l'infection dans les mutants *CG4572*^{c05963} par rapport aux drosophiles *w*^{1118}. Dans les carcasses, l'effet de l'infection virale sur la production des petits ARN interférents semble atténué lorsque la voie de l'import de l'ARN double brin est invalidée. Par ailleurs, la voie de l'import de l'ARN double brin étant impliquée dans la réponse antivirale des siARN, nous avons ensuite déterminé si les siARN sont impliqués dans la modulation de la production des piARN et de l'activité des ET.

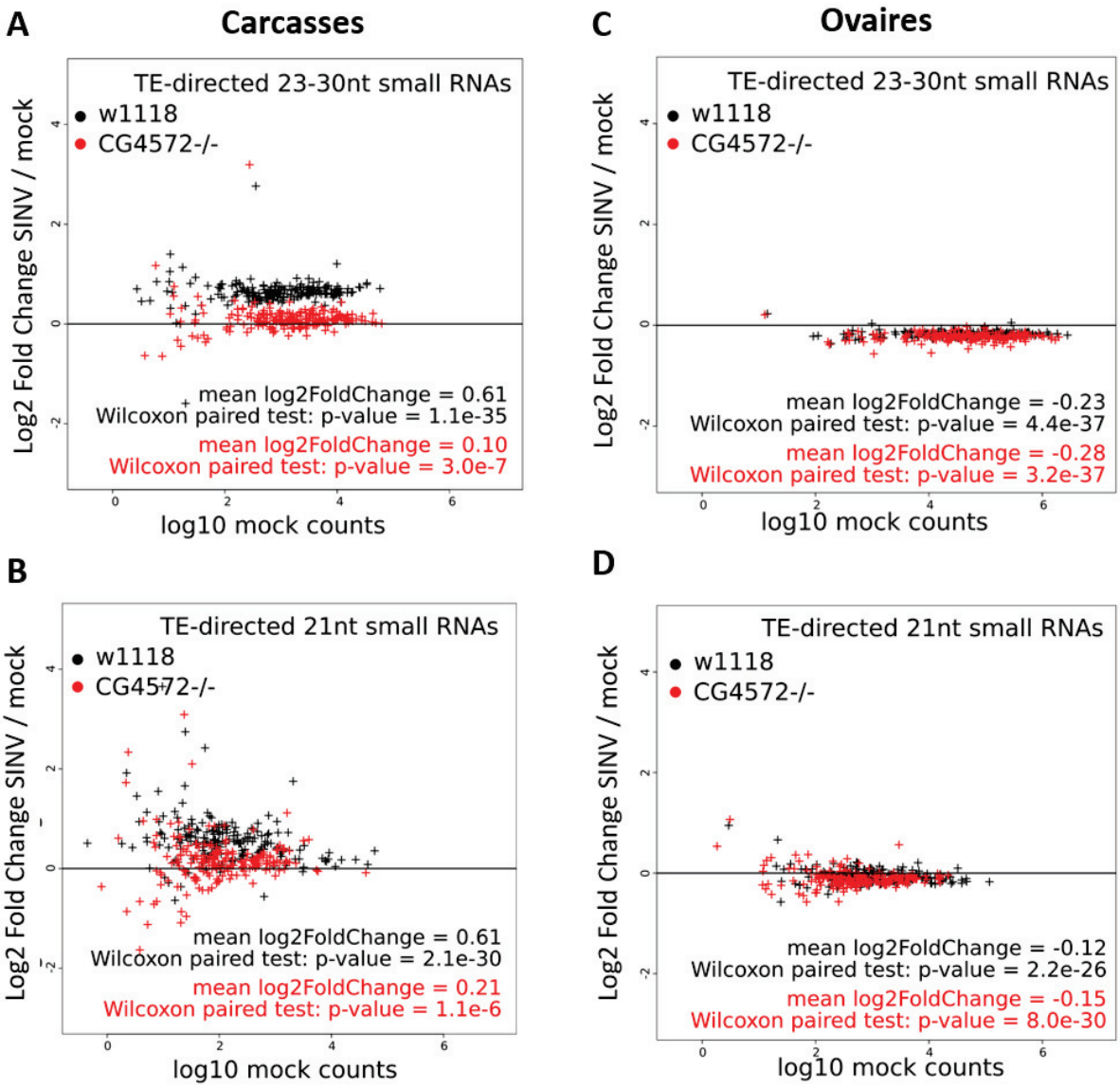


Figure 53. Analyse de la modulation des productions de petits ARN interférents ciblant les ET chez *D. melanogaster CG4572*^{c05963} (CG4572-/-), lors de l'infection par SINV, dans les carcasses (A-B) et dans les ovaires (C-D). Chaque point représente le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par SINV et la condition non infectée (mock), pour les « piARN » **(A, C)** et pour les « siARN » **(B, D)**. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). n = 2 expériences indépendantes.

B. Impact de la voie des siARN sur la dérégulation des ET et des petits ARN interférents ciblant les ET

Les résultats précédents suggèrent que la modulation des transcrits d'ET que nous observons serait due à une augmentation des quantités de petits ARN interférents et notamment des siARN, dans les carcasses. Par conséquent, nous avons étudié les conséquences de l'absence de production de siARN fonctionnels sur les modulations observées précédemment. Pour cela, nous avons utilisé les mutants *Dcr-2*$^{[L811fsX]}$, qui sont des drosophiles présentant le même fond génétique que les drosophiles sauvages *D. melanogaster w*^{1118} mais avec une mutation au niveau de la Leucine 811 de Dcr-2, qui a été validée (**Figure S15.**). Cette protéine est essentielle à la production de complexes double brin d'ARN appelés siARN, présentant une taille de 21 nucléotides et représentant l'une des principales défenses au niveau de l'immunité antivirale chez les insectes. Par conséquent, n'ayant plus une protéine Dcr-2 fonctionnelle, ces mutants ne sont plus capables de produire des siARN double brin et de détecter spécifiquement l'ARNm viral à cliver (Lee et al., 2004; Liu et al., 2006).

a) Etude de la réplication du virus Sindbis chez les mutants *D. melanogaster Dcr-2*$^{[L811fsX]}$

En accord avec la littérature (Lamiable et al., 2016b; Petit et al., 2016; Saleh et al., 2009; Tassetto et al., 2017), nous avons infecté les drosophiles *w*^{1118} et *Dcr-2*$^{[L811fsX]}$ avec une dose de 2300 PFU du virus Sindbis (SINV). Comme attendu, SINV induit une forte mortalité des mutants *Dcr-2*$^{[L811fsX]}$ infectés, et cela à partir du neuvième jour après l'infection, jusqu'à atteindre 100% de mortalité au dixième jour après l'infection (**Figure 54.**). Le faible pourcentage de mortalité observé dans les conditions non infectées est représentatif de celui retrouvé classiquement lors de l'élevage de drosophiles en laboratoire.

La cinétique de réplication de SINV a été comparée entre les drosophiles *w*^{1118} et les mutants *Dcr-2*$^{[L811fsX]}$. SINV se réplique abondamment dans les mutants *Dcr-2*$^{[L811fsX]}$ incluant une augmentation constante de la réplication de SINV jusqu'à 8 jours après l'infection (**Figure 55.** ; médiane au jour 2 = 2,9 log10(TCID50/ml) ; médiane au jour 4 = 4,7 log10(TCID50/ml) ; médiane au jour 6 = 5,6 log10(TCID50/ml) ; médiane au jour 8 = 6,7 log10(TCID50/ml) puis une phase plateau aboutissant à une diminution de la réplication de SINV au jour 10 après infection (médiane = 5,3 log10(TCID50/ml) ; test statistique de Wilcoxon-Mann-Whitney entre jour 8 et jour 10 ; *p-value* = 2,2e-3). On note également que SINV se réplique beaucoup plus chez les mutants *Dcr2*$^{[L811fsX]}$ par rapport aux *w*^{1118} avec environ 2 à 3 log10(TCID50/ml)

supplémentaires aux jours 4 à 10 après infection (**Figure 55.** ; test statistique de Wilcoxon-Mann-Whitney ; *p-value* < 0,05).

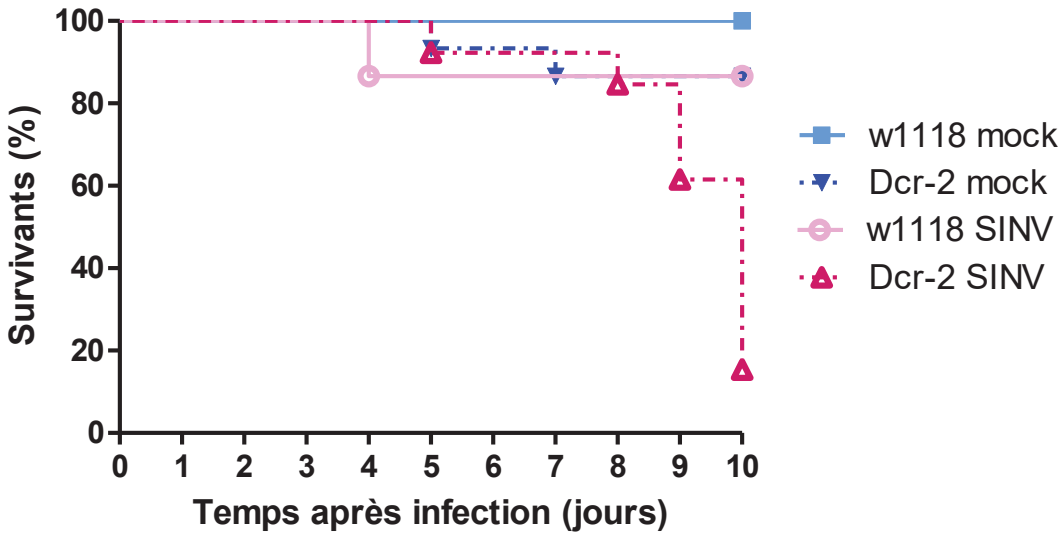


Figure 54. Mortalité des mutants *D. melanogaster Dcr-2*$^{[L811fsX]}$ induite par SINV. Les drosophiles *w*^{1118} (traits pleins) et *Dcr-2*$^{[L811fsX]}$ (traits en pointillés) ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. Les drosophiles vivantes ont été quantifiées tout au long de la cinétique de réplication (du jour de l'infection à 10 jours après infection). Les courbes en bleu représentent la condition mock (non infectée), les courbes en rose représentent la condition infectée par SINV. n = 3 expériences indépendantes (15 drosophiles par condition).

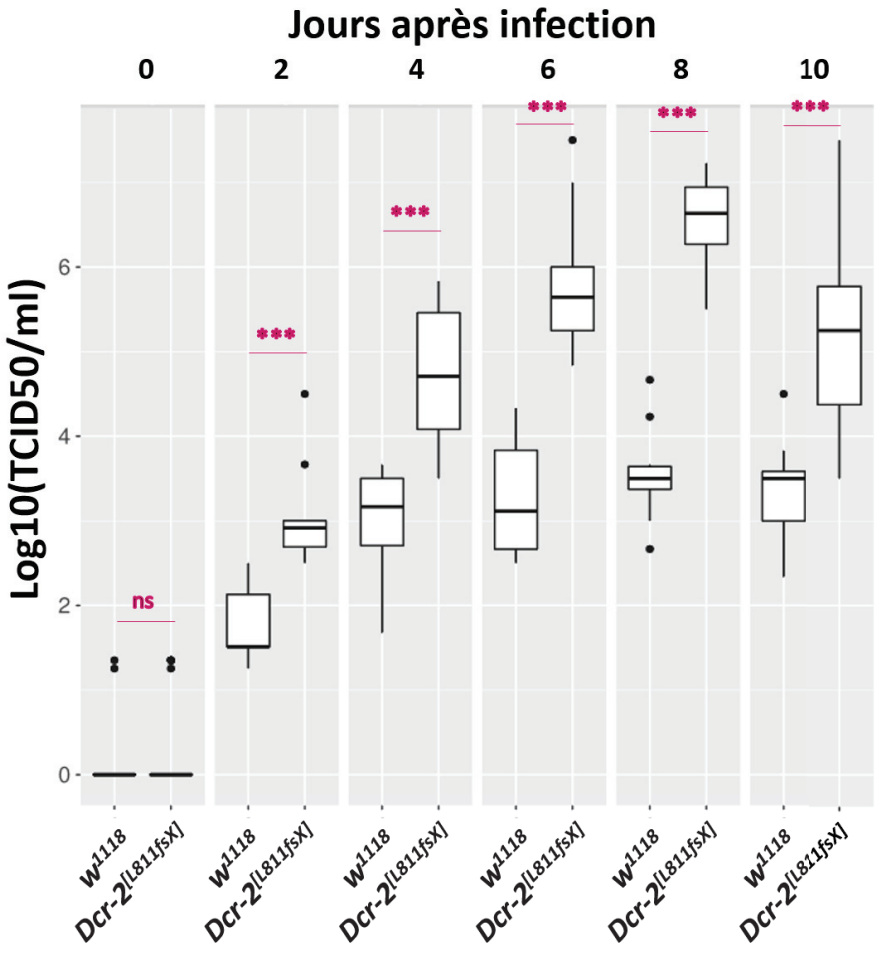


Figure 55. Réplication de SINV dans les mutants *Dcr2*$^{[L811fsX]}$ et les drosophiles *w*^{1118}. Les drosophiles *w*^{1118} et *Dcr-2*$^{[L811fsX]}$ ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de drosophiles infectées, tout au long de la cinétique de réplication (du jour de l'infection à 10 jours après infection). n = 3 expériences indépendantes (5 drosophiles par condition). Test statistique de Wilcoxon-Mann-Whitney ; ns : non significatif ; ***P < 0,001.

Comme pour les autres lignées de drosophiles, les mutants *Dcr-2*$^{[L811fsX]}$ ont été disséqués afin d'analyser les tissus séparemment, incluant les paires d'ovaires d'un côté et les carcasses de l'autre. Les titres viraux de SINV au niveau des carcasses des mutants *Dcr- 2*$^{[L811fsX]}$ sont comparables à ceux déterminés en drosophiles entières, incluant une phase exponentielle de réplication de SINV jusqu'à 6 jours après infection (médiane au jour 2 = 2,7 log10(TCID50/ml) ; médiane au jour 6 = 5,3 log10(TCID50/ml) puis une diminution de la réplication virale au jour 10 après infection (médiane = 3,7 log10(TCID50/ml) (**Figure 51.)**. De plus, comme dans les drosophiles entières, tout au long de la cinétique de réplication virale, les titres de SINV sont plus élevés dans les carcasses des mutants *Dcr- 2*$^{[L811fsX]}$ par rapport aux drosophiles *w*^{1118} et aux mutants *CG4572*^{c05963} (**Figure 51.** ; test statistique de Wilcoxon-Mann-Whitney ; *p-value* < 0,05). Par ailleurs, il y a très peu de détection de SINV dans les ovaires des *w*^{1118} par rapport

aux mutants *Dcr2*$^{[L811fsX]}$ qui atteignent des titres viraux supérieurs aux titres de SINV retrouvés dans les ovaires des *w*^{1118} (**Figure 51.** ; médiane au jour 6 = 3,6 log10(TCID50/ml) ; médiane au jour 10 = 2,9 log10(TCID50/ml ; test statistique de Wilcoxon-Mann-Whitney, *p-value* < 0,05), pouvant même atteindre des titres viraux comparables à ceux obtenus dans les carcasses des *w*^{1118} infectées.

Par la suite, des expériences d'immunofluorescence ont été réalisées dans le but de vérifier la présence de la capside de SINV dans les ovaires des mutants *Dcr2*$^{[L811fsX]}$. Cependant, SINV n'a pas été détecté au sein des structures ovariennes mais seulement en périphérie des ovaires des *Dcr2*$^{[L811fsX]}$, à proximité des cellules folliculaires (**Figure 56.**). Ces résultats suggèrent une absence de réplication virale de SINV au sein des ovaires des drosophiles *w*^{1118} et *Dcr2*$^{[L811fsX]}$.

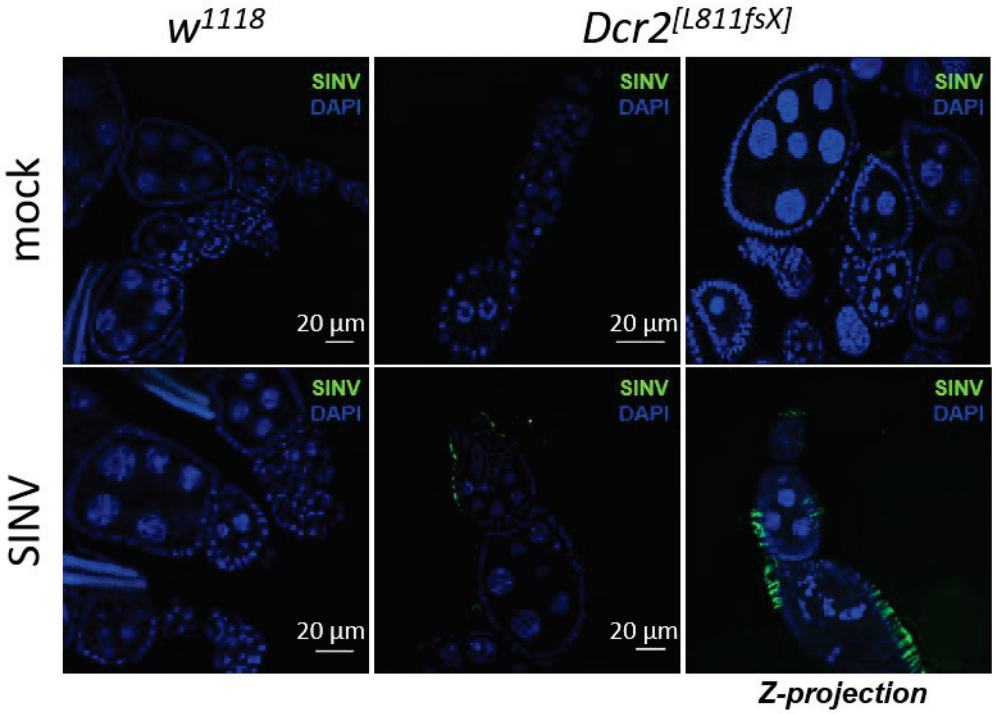


Figure 56. Détection de SINV seulement en périphérie des ovaires des mutants *Dcr-2*$^{[L811fsX]}$. Les drosophiles *w*^{1118} et *Dcr-2*$^{[L811fsX]}$ ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl puis les paires d'ovaires ont été disséquées au jour 6 après infection. La technique d'immunofluorescence a été réalisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488). La projection en Z correspond aux mutants *Dcr-2*$^{[L811fsX]}$ mock (haut) et infectés (bas). L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 10 paires d'ovaires.

Le tropisme de SINV chez *D. melanogaster w*^{1118} et *Dcr-2*$^{[L811fsX]}$ a été étudié plus en détail, incluant leur comparaison, en réalisant des coupes de drosophiles et un immunomarquage de la capside de SINV. Au niveau des drosophiles contrôles *w*^{1118}, le marquage de la capside de

SINV dans le thorax **(Figures 57.)** et l'abdomen **(Figures 58.)** est sous forme d'agrégats fluorescents intenses, particulièrement au niveau de la zone d'infection, alors que pour les mutants *Dcr-2*$^{[L811fsX]}$ le marquage est beaucoup plus diffus et moins intense. SINV semble ainsi se répliquer de façon ubiquitaire chez *D. melanogaster* et plus particulièrement au niveau du corps gras des drosophiles (**Figures 59.**)

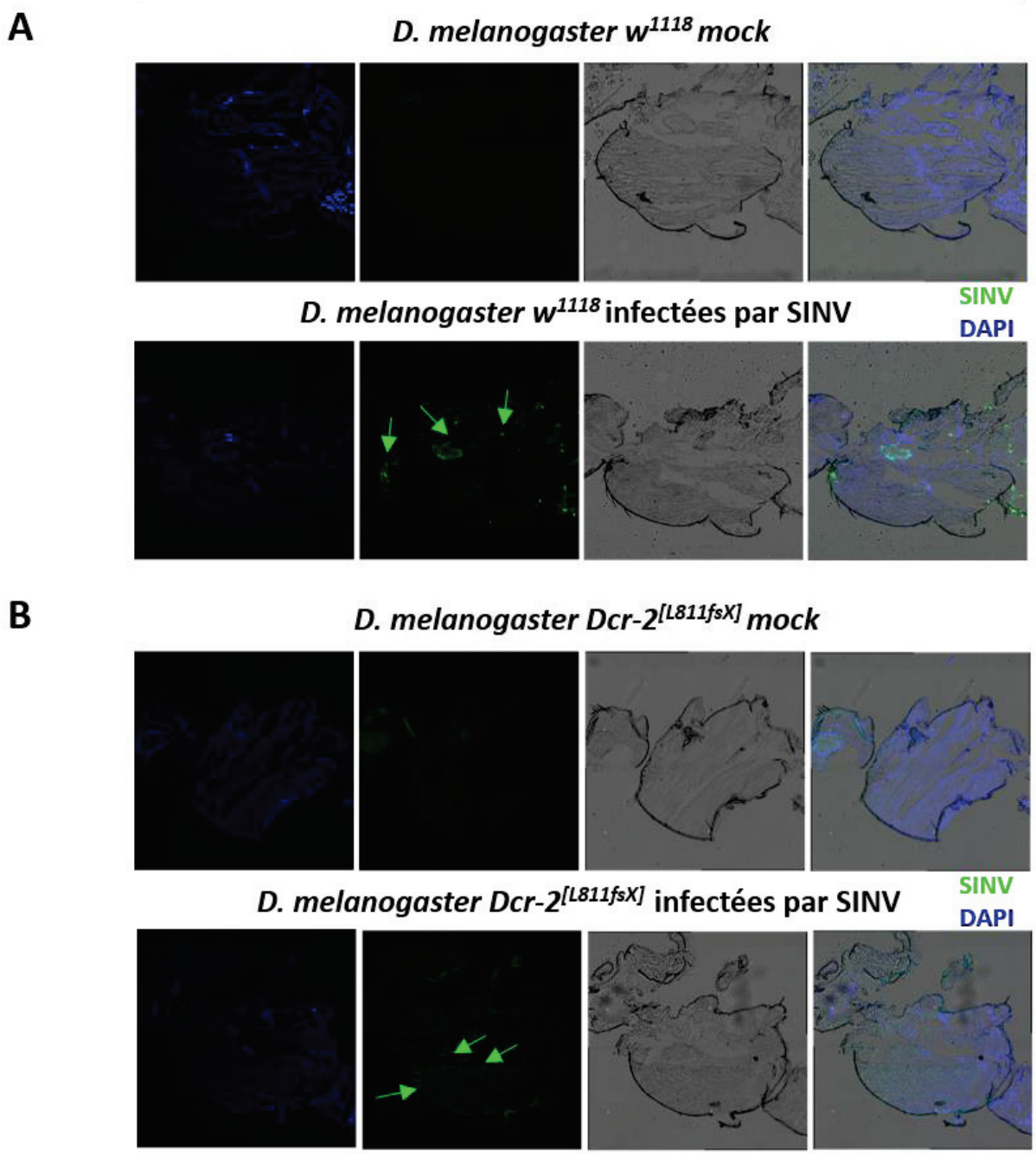


Figure 57. Détection de SINV dans le thorax des drosophiles *w*^{1118} et *Dcr-2*$^{[L811fsX]}$. Les drosophiles ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. La technique d'immunofluorescence a été réalisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488) au jour 6 après l'infection des drosophiles *w*^{1118} **(A)** et des mutants *Dcr-2*$^{[L811fsX]}$ **(B)**. L'ADN a

été marqué en bleu (DAPI). Les flèches vertes indiquent les zones de détection de la capside de SINV. Les résultats sont représentatifs d'au moins 2 expériences indépendantes incluant chacune 3 drosophiles.

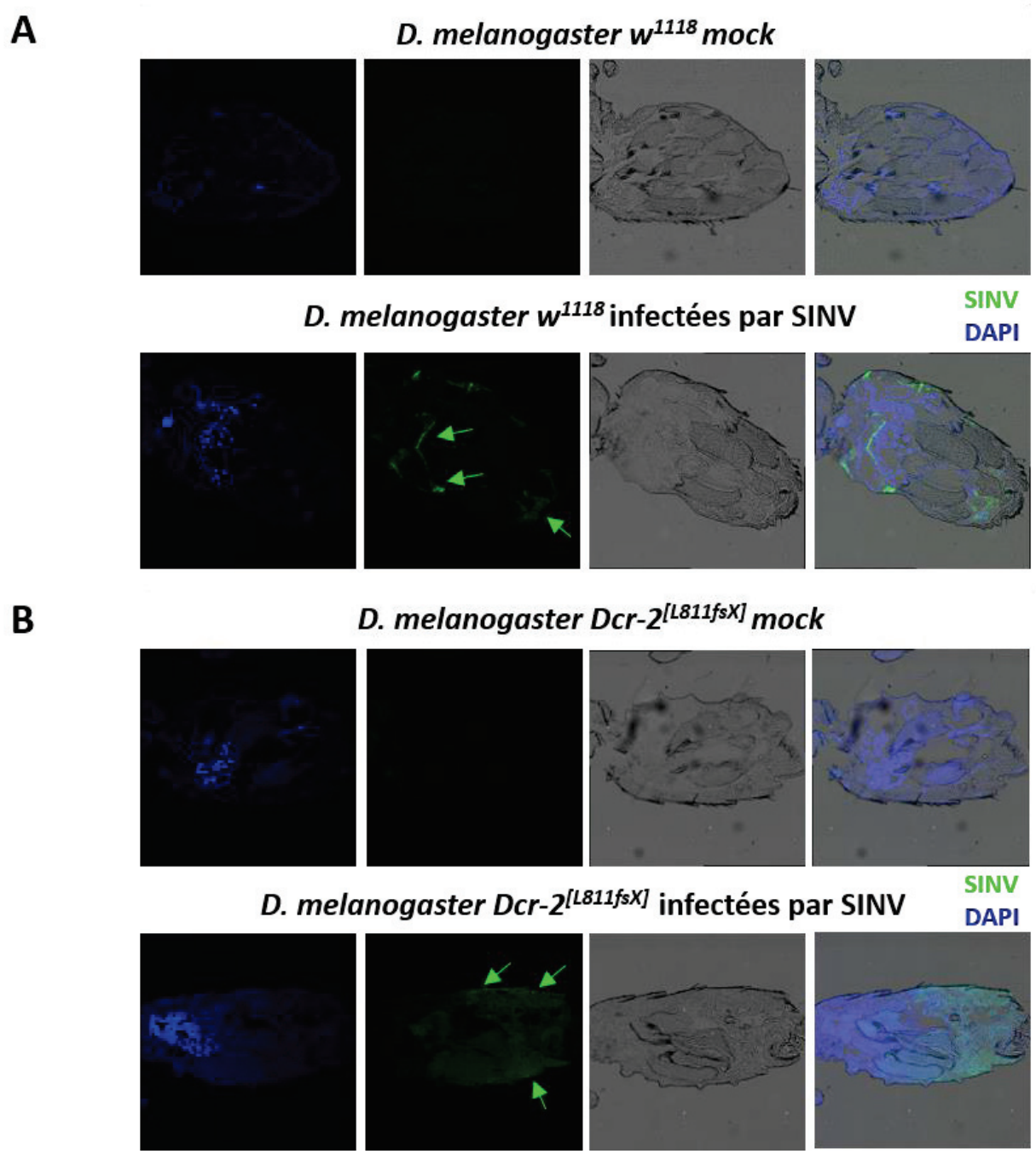


Figure 58. Détection de SINV dans l'abdomen des drosophiles *w*^{1118} et *Dcr-2*$^{[L811fsX]}$. Les drosophiles ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. La technique d'immunofluorescence a été réalisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488) au jour 6 après l'infection des drosophiles *w*^{1118} **(A)** et des mutants *Dcr-2*$^{[L811fsX]}$ **(B)**. L'ADN a été marqué en bleu (DAPI). Les flèches vertes indiquent les zones de détection de la capside de SINV. Les résultats sont représentatifs d'au moins 2 expériences indépendantes incluant chacune 3 drosophiles.

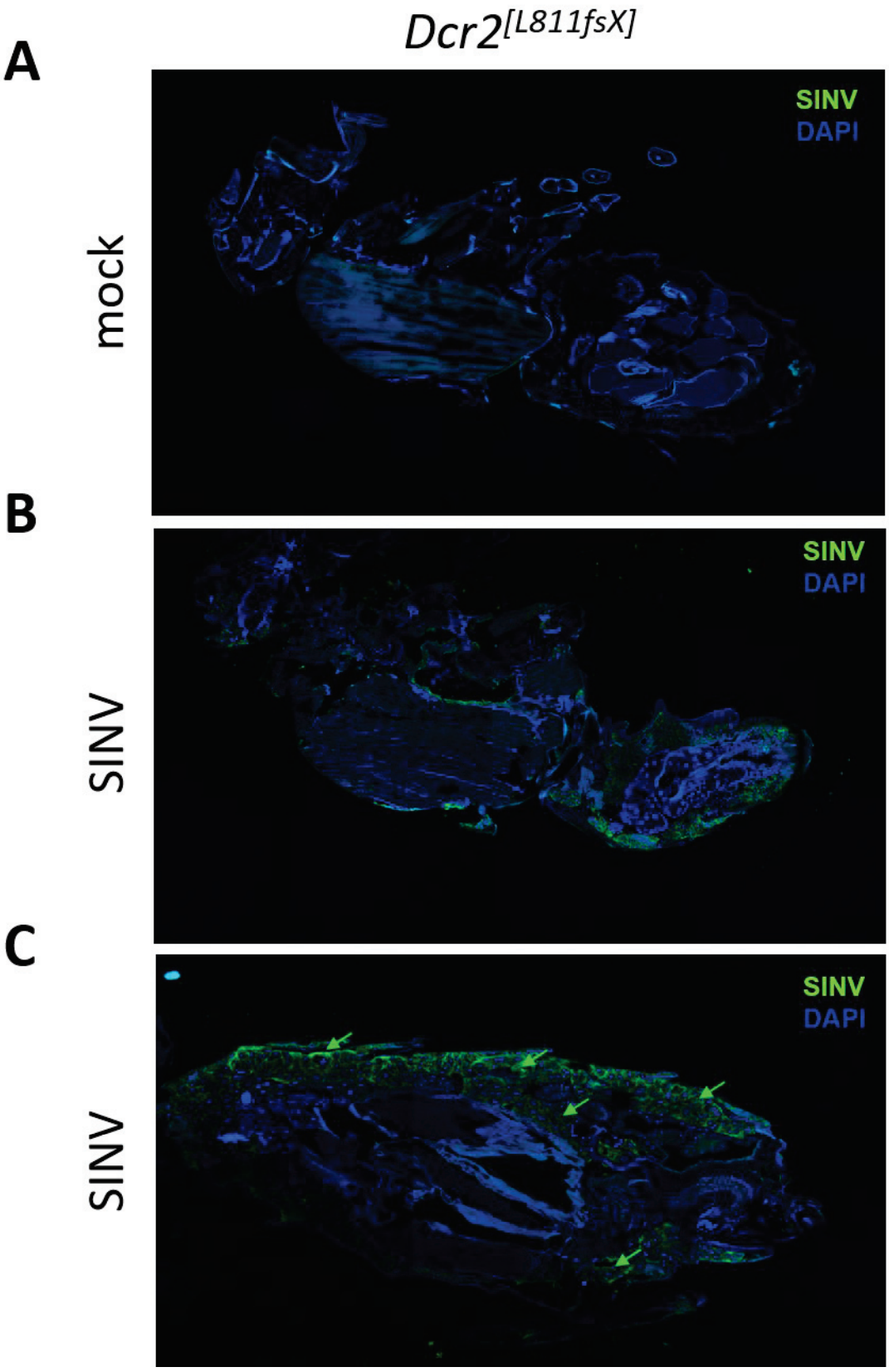


Figure 59. Détection de SINV dans le corps gras des drosophiles *Dcr-2*$^{[L811fsX]}$. Les drosophiles ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl. **(A)** Drosophile contrôle non infectée (mock). **(B)** Drosophile infectée par SINV. **(C)** Abdomen de drosophile infectée par SINV ; les flèches vertes

indiquent le corps gras positif pour SINV. La technique d'immunofluorescence a été réalisée en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488) au jour 6 après l'infection. L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 3 expériences indépendantes incluant chacune 3 drosophiles.

L'ensemble de ces résulats indique une réplication plus importante de SINV dans les mutants *Dcr-2*$^{[L811fsX]}$, associée à une mortalité de ces drosophiles infectées aux temps tardifs de réplication. Par conséquent, ces résultats attendus permettent une première validation de ces mutants et ont ensuite été suivis par l'analyse de la production de petits ARN interférents ciblant SINV.

b) Modulation de la production des petits ARN interférents ciblant le virus Sindbis chez *D. melanogaster Dcr-2*$^{[L811fsX]}$

A l'opposé des drosophiles *w*^{1118}, les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV, révèlent seulement en faible quantité les brins sens des petits ARN interférents de 21 nucléotides de longueur, indiquant qu'aucun siARN double brin n'est produit lors de l'infection dans les carcasses et les ovaires (**Figure 60 A.**). De plus, chez les mutants *Dcr-2*$^{[L811fsX]}$, la distribution des petits ARN interférents de 21 nucléotides le long du génome viral révèle seulement le brin sens et inclut notamment un point chaud de production à environ 7600 nucléotides (**Figure 60 B.**). Cette région génomique correspond à la jonction entre l'ARN génomique et sous-génomique de SINV. Cet aspect sera précisé dans la partie discussion. Par ailleurs, lors de l'infection par SINV, nous avons détecté en faible quantité presque uniquement les brins sens des petits ARN interférents de 23 à 30 nucléotides dans les carcasses et les ovaires des mutants *Dcr-2*$^{[L811fsX]}$ (**Figure 60 A.**).

Dans cette partie, l'analyse des petits ARN interférents valide l'absence de production de siARN fonctionnels chez les mutants *Dcr-2*$^{[L811fsX]}$. De plus, nous n'avons détecté aucun petit ARN interférent ciblant les virus responsables d'infections chroniques, tels que les virus NV, DAV et DCV chez *D. melanogaster w*^{1118} et *Dcr-2*$^{[L811fsX]}$ (**Figure S16.**). Ces résultats permettent de valider le modèle de *D. melanogaster Dcr-2*$^{[L811fsX]}$. Les lignées *w*^{1118} et *Dcr-2*$^{[L811fsX]}$ ont ensuite été comparées pour déterminer l'implication de la production des siARN dans les modulations de l'activité des ET et de la production des piARN dirigés contre les ET lors de l'infection par SINV.

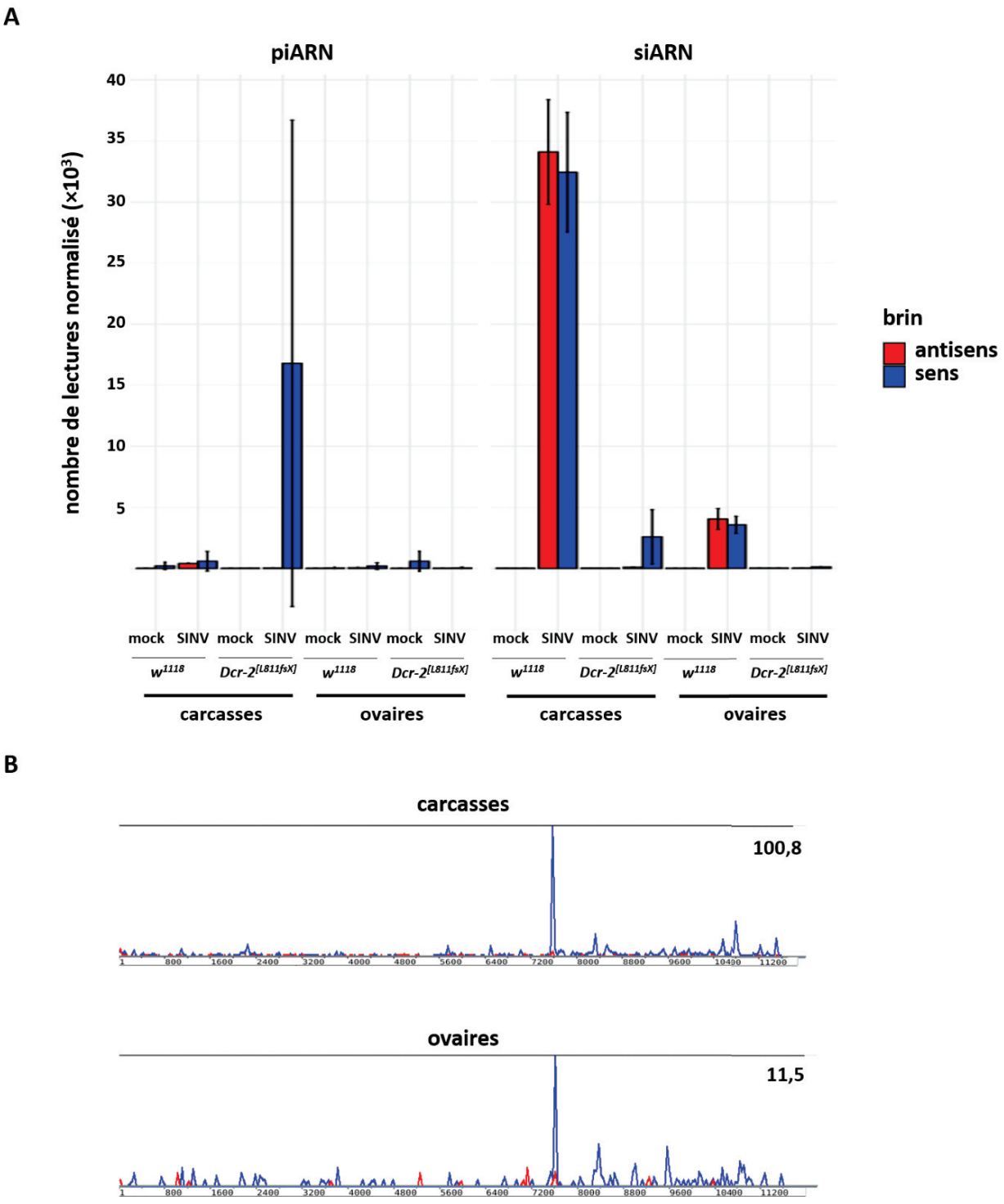


Figure 60. Absence de production des brins antisens des petits ARN interférents ciblant SINV chez *les mutants Dcr-2$^{[L811fsX]}$*. (A) Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles sauvages *w^{1118}* et des mutants *Dcr-2$^{[L811fsX]}$*, normalisé par les comptes des lectures des miARN. **(B)** Distribution des petits ARN interférents de 21 nulcéotides ciblant SINV, le long du génome viral, incluant les brins sens (bleu) et antisens (rouge) chez les mutants *Dcr-2$^{[L811fsX]}$*. La valeur numérique en haut à droite correspond au nombre de lectures pour le pic le plus haut. n = 2 expériences indépendantes (50 tissus par échantillon).

c) Modulation de la production des petits ARN interférents ciblant les ET chez *D. melanogaster Dcr-2$^{[L811fsX]}$*

Nous avons ensuite déterminé l'impact de l'invalidation de la voie des siARN (mutants *Dcr- 2$^{[L811fsX]}$*) sur la modulation des quantités de transcrits d'ET et des petits ARN interférents ciblant les ET, précédemment observée chez les *w^{1118}* (résultats de la partie I.2.C.b « Modulation de l'abondance des transcrits d'ET et de la production des piARN ciblant les ET chez *D. melanogaster* » reportés ici pour la comparaison avec les mutants *D. melanogaster Dcr- 2$^{[L811fsX]}$*).

Carcasses :

Suite à l'infection par SINV, la modulation de l'activité des ET n'est pas la même chez les drosophiles sauvages *w^{1118}* par rapport aux mutants *Dcr-2$^{[L811fsX]}$*, incluant une diminution des quantités de transcrits d'ET pour les drosophiles *w^{1118}* (**Figure 61A.** ; points noirs ; moyenne log2 *fold change* = -0,32 ; *p-value* = 1,8e-24) alors qu'une augmentation a été observée pour les mutants *Dcr-2$^{[L811fsX]}$* (**Figure 61A.** ; points rouges ; moyenne log2 *fold change* = 1,85 ; *p-value* = 4,2e-33). De plus, chez les mutants *Dcr-2$^{[L811fsX]}$*, l'infection par SINV conduisant à cette augmentation des quantités de transcrits d'ET (**Figure 61 A.**) est associée à une diminution des piARN ciblant les ET (**Figure 61 B.** ; points rouges ; moyenne log2 *fold change* = -0,50 ; *p-value* = 2,1e-32) et à une répression du gène *piwi* (**Figure 61 C.**, log2 *fold change* = -1,46 ; *p-value* = 0,027) et du gène *ninaC* (**Figure 61 C.**, log2 *fold change* = -1,30 ; *p-value* = 2,0e-5), impliqué dans la voie de l'import de l'ARN double brin. Cependant, nous avons observé une surexpression du gène *egh* (**Figure 61 C.**, log2 *fold change* = -1,54 ; *p-value* = 3,2e-11), impliqué également dans la voie de l'import de l'ARN double brin. On retrouve également une surexpression des gènes *dcr2* dans les mutants *Dcr- 2$^{[L811fsX]}$* suite à l'infection par SINV (**Figure 61 C.**, log2 *fold change* = 3,77 ; *p-value* = 2,6e-61), et *ago2* (**Figure 61 C.**, log2 *fold change* = 1,06 ; *p-value* = 2,7e-5) indiquant une stimulation de la transcription des effecteurs de la voie des siARN lorsque leur production est invalidée. Ces résultats sont opposés à ceux obtenus chez *w^{1118}* incluant une augmentation significative des piARN (**Figure 61 B**. ; points noirs ; moyenne log2 *fold change* = 0,61 ; *p-value* = 1,1e-35) et aucune modulation significative des quantités de transcrits des gènes codant les principaux effecteurs de la voie des piARN (**Figure 61 C.** ; évaluée par DESeq2 ; *p-value* > 0,05).

De plus, en condition non infectée, la comparaison entre les mutants *Dcr-2$^{[L811fsX]}$* et les drosophiles sauvages *w^{1118}*, révèle une diminution significative des quantités de transcrits d'ET (**Figure S17 A** ; moyenne log2 *fold change* = -0,29 ; *p-value* = 1,6e-4), associée à une augmentation des piARN ciblant ces ET (**Figure S17 B.** ; moyenne log2 *fold change* = 0,59 ; *p-value* = 4,4e-22) et excluant une modulation des quantités de transcrits des gènes codant les effecteurs principaux de cette voie d'ARN interférence (**Figure 61 C.** ; évaluée par DESeq2 ; *p-*

value > 0,05). Dans cette comparaison, en condition non infectieuse, l'expression du gène *dcr2* est significativement réprimée (**Figure 61 C.** ; log2 *fold change* = -0,88 ; *p-value* = 0,011)*,* tandis que l'expression du gène *ninaC* est stimulée (**Figure 61 C.** ; log2 *fold change* = 1,06 *; p-value =* 4,7e-3*).*

Par ailleurs, en condition infectée par SINV, la comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ et les drosophiles sauvages *w*^{1118}, révèle une augmentation des quantités de transcrits d'ET (**Figure S17 C.** ; moyenne log2 *fold change* = 1,92 ; *p-value* = 2,2e-24), associée à une diminution des piARN ciblant ces ET (**Figure S17 D.** ; moyenne log2 *fold change* = -0,49 ; *p-value* = 3,6e-10). Cette même comparaison révèle également une importante répression des gènes *piwi* (**Figure 61 C.**, log2 *fold change* = -2,13 ; *p-value* = 7,0e-4) et *aub* (**Figure 61 C.**, log2 *fold change* = -2,35 ; *p-value* = 2,7e-3), associée à une forte induction de la transcription de *dcr2* (**Figure 61 C.**, log2 *fold change* = 2,79 ; *p-value* = 1,4e-23) -qui cependant, ne conduit pas à une protéine Dcr-2 fonctionnelle puisque la séquence codante de cette dernière est mutée- et enfin une surexpression des gène *ago2* (**Figure 61 C.**, log2 *fold change* = 1,46 ; *p-value* = 5,3e-8) et *egh* (**Figure 61 C.**, log2 *fold change* = 1,73 ; *p-value* = 5,4e-12).

Ovaires :

Comme détaillé dans la partie I.2 (Infection par le virus Sindbis chez *D. melanogaster*), lors de l'infection par SINV des drosophiles sauvages *w*^{1118}, les modulations des quantités de transcrits des ET (**Figure 61 D.** ; points noirs ; moyenne log2 *fold change* = -0,10 ; *p-value* = 5,3e-14) et des petits ARN interférents ciblant les ET sont très faibles dans les ovaires et notamment en ce qui concerne les piARN (**Figure 61 E-F.** ; points noirs ; moyenne log2 *fold change* = - 0,23 ; *p-value* = 4,4e-37). De la même façon, les quantités de transcrits d'ET ne sont quasiment pas altérées chez les mutants *Dcr- 2*$^{[L811fsX]}$, au niveau des ovaires (**Figure 61 D.** ; points rouges ; moyenne log2 *fold change* = -0,05 ; *p-value* = 0,042), indiquant que les ET ne sont pas (ou peu) contrôlés par les siARN dans cet organe. Par ailleurs, l'importante stimulation de l'expression du gène *dcr2* (**Figure 61 F.** ; log2 *fold change* = 2,13 ; *p-value* = 4,9e-28) a été identifiée dans les ovaires lors de l'infection par SINV des mutants *Dcr-2*$^{[L811fsX]}$, suggérant une stimulation de cette voie suite à l'infection mais n'aboutissant pas à une production de siARN en raison de l'invalidation du gène codant Dcr-2. L'augmentation de la production des piARN chez les mutants *Dcr-2*$^{[L811fsX]}$ lors de l'infection par SINV, n'incluant pas d'impact sur l'activité des ET, reste à élucider (**Figure 61 E.** ; points rouges ; moyenne log2 *fold change* = 0,25 ; *p-value* = 1,4e-20).

De plus, en condition non infectée et infectée par SINV, la comparaison entre les mutants *Dcr- 2*$^{[L811fsX]}$ et les drosophiles sauvages *w*^{1118}, ne révèle pas de modulation des quantités de transcrits d'ET (moyenne log2 *fold change* en condition non infectée = -0,05 ; *p-value* = 0,8 ; moyenne log2 *fold change* en condition infectée par SINV = 0,0086 ; *p-value* = 0,5) (**Figure S18 A-B**), mais induit une diminution significative des piARN (moyenne log2 *fold change* en

condition mock = -0,79 ; *p-value* = 1,4e-21 ; moyenne log2 *fold change* en condition infectée par SINV = -0,35 ; *p-value* = 1,1e-5) (**Figure S18 C-D**). Par la suite, nous avons notamment observé une surexpression significative mais modérée du gène *piwi* (log2 *fold change* en condition mock = 0,27 ; *p-value* = 9,4e-4 ; log2 *fold change* en condition infectée par SINV = 0,33 ; *p-value* = 3,4e-7) dans les mutants $Dcr\text{-}2^{[L811fsX]}$ par rapport aux drosophiles sauvages w^{1118} (**Figure 61 F.**). Ces résultats pourraient expliquer l'augmentation des petits ARN interférents de 23 à 30 nucléotides lors de l'infection chez les mutants $Dcr\text{-}2^{[L811fsX]}$. En condition non infectée, notons également une surexpression du gène *egh* (**Figure 61 F.**, log2 *fold change* = 0,35 ; *p-value* = 3,0e-6) dans les mutants $Dcr\text{-}2^{[L811fsX]}$ par rapport aux drosophiles sauvages.

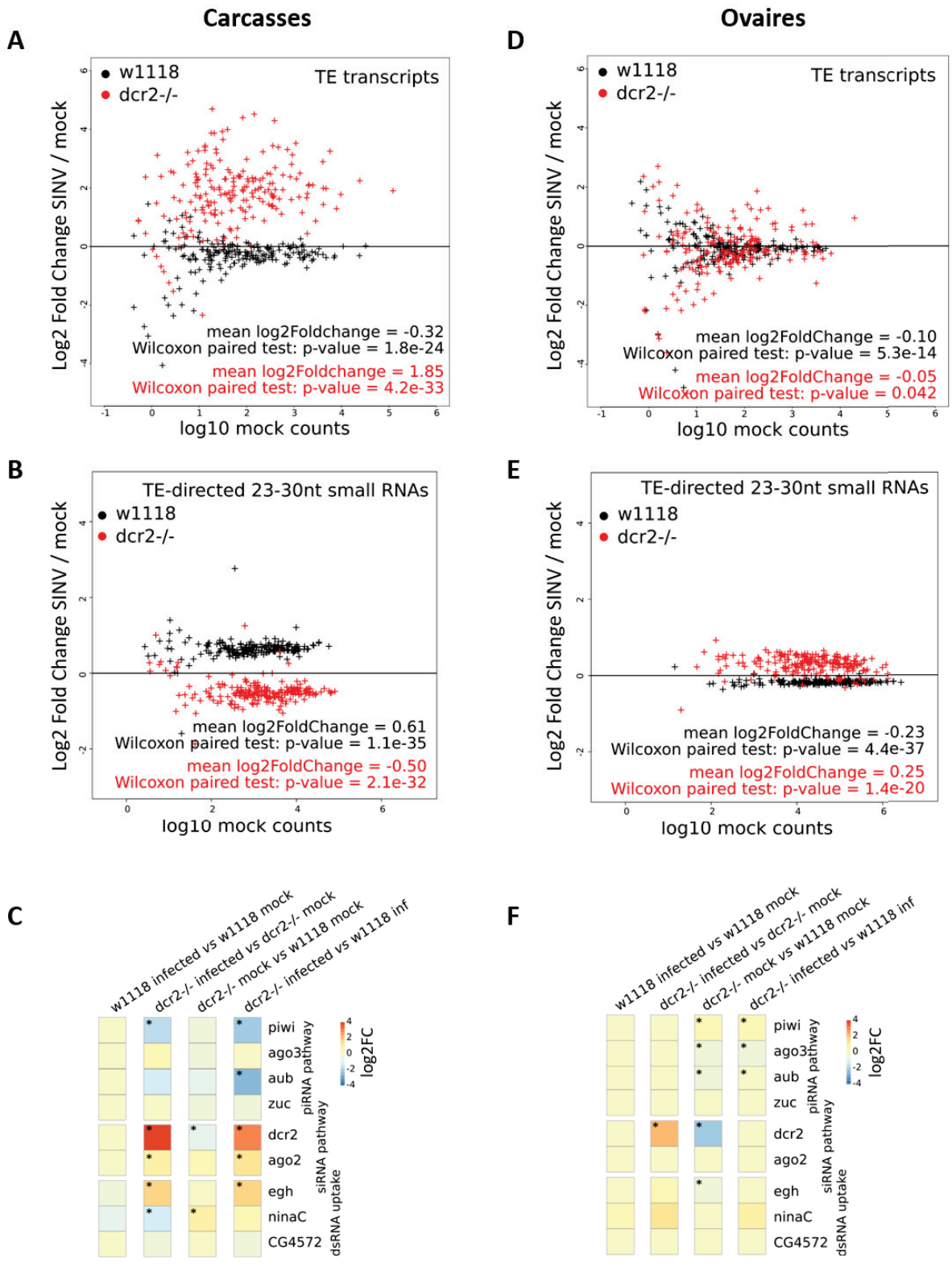


Figure 61. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez *D. melanogaster Dcr-2*$^{[L811fsX]}$, lors de l'infection par SINV, dans les carcasses (A-C) et dans les ovaires (D-F). Chaque point représente le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par SINV et la condition non infectée (mock), pour les transcrits d'ET **(A, D)**, pour les petits ARN interférents de 23-30 nucléotides (piARN) **(B, E)**. La significativité de la modulation est établie en comparant les comptes de la

condition infectée par rapport aux comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). **(C-F)** Modulation des taux de transcrits pour un ensemble de gènes impliqués dans la voie des piARN (*piwi*, *ago3*, *aub*, *zuc*), la voie des siARN (*dcr2*, *ago2*) et la voie de l'import de l'ARN double brin (*egh*, *ninaC*, *CG4572*) telle qu'évaluée par DESeq2, dans les carcasses (C) et dans les ovaires (F). Les changements significatifs (*P<0,05*) sont indiqués par des étoiles. n = 2 expériences indépendantes.

L'ensemble des résultats en carcasse révèle un impact de l'invalidation de la production des siARN sur les modulations des quantités de piARN ciblant les ET et des quantités de transcrits des ET observés chez les drosophiles sauvages w^{1118}. A l'inverse, au niveau des ovaires, l'invalidation de la production de siARN n'a pas d'effet sur les quantités de transcrits des ET mais révèle cependant une modulation différente des piARN ciblant les ET par rapport aux drosophiles w^{1118}. Par la suite, afin de préciser les résultats précédents, l'analyse des petits ARN interférents dirigés contre les ET a été appronfondie.

C. Analyse des petits ARN interférents ciblant les ET chez *D. melanogaster*

L'analyse du nombre de lectures des brins sens et antisens des petits ARN interférents de 23-30 nucléotides ciblant les ET révèle une proportion équivalente de brin sens et antisens dans les carcasses, et une proportion plus importante de brin antisens, dans les ovaires, pour les drosophiles w^{1118} et les mutants $CG4572^{c05963}$ et $Dcr\text{-}2^{[L811fsX]}$ (**Figure 62 A.**). Ces résultats suggèrent une activation de la voie primaire et secondaire des piARN, représentant environ 70% de la totalité des petits ARN interférents de 23-30 nucléotides dans les ovaires et environ 5% dans les carcasses (**Figure S19.**). De plus, la signature caractéristique des piARN (ciblant les ET) impliqués dans la boucle d'amplification est, comme attendu, détectée dans les ovaires mais aussi dans les carcasses w^{1118}, $CG4572^{c05963}$ et $Dcr\text{-}2^{[L811fsX]}$, pour certaines familles d'ET, telles que *tirant*, *doc*, *Doc6*, *accord*, *412*, *gypsy* et *zam* (**Figure S20.** ; *doc*, *412* et *accord* montrés). Cependant, une trop faible production des piARN ciblant une famille d'ET ne permet pas la détection de cette signature. Par ailleurs, aucune différence de signature n'a été détectée entre les conditions non infectées et infectées par SINV (**Figure S20.**).

De plus, l'abondance du brin sens et antisens pour les petits ARN interférents de 21 nucléotides ciblant les ET est équivalente chez les drosophiles w^{1118} (**Figure 62 B.**), représentant environ 5% de la totalité des petits ARN interférents de 21 nucléotides dans les carcasses, tandis qu'ils représentent environ 20% dans les ovaires (**Figure S19.**). Chez les mutants $CG4572^{c05963}$ et $Dcr\text{-}2^{[L811fsX]}$, les petits ARN interférents de 21 nulcléotides représentent environ 2% de l'ensemble des petits ARN interférents de 21 nucléotides produits dans les carcasses (**Figure S19.**). A l'inverse des drosophiles w^{1118} et $CG4572^{c05963}$, nous n'avons détecté dans les carcasses des mutants $Dcr\text{-}2^{[L811fsX]}$ quasiment que les brins sens des petits

ARN interférents de 21 nulcéotides, ce qui confirme que ce ne sont pas des siARN fonctionnels (**Figure 62 B.**). Par ailleurs, de façon surprenante, de la même façon que pour les drosophiles w^{1118} et $CG4572^{c05963}$, nous avons observé une porportion équivalente des brins sens et antisens dans les ovaires des mutants $Dcr\text{-}2^{l(L811fsX}$ (**Figure 62 B.**), représentant néanmoins une plus faible proportion (12 %) de la totalité des petits ARN interférents de 21 nucléotides produits dans les ovaires de ces mutants par rapport aux drosophiles w^{1118} (**Figures S19.**).

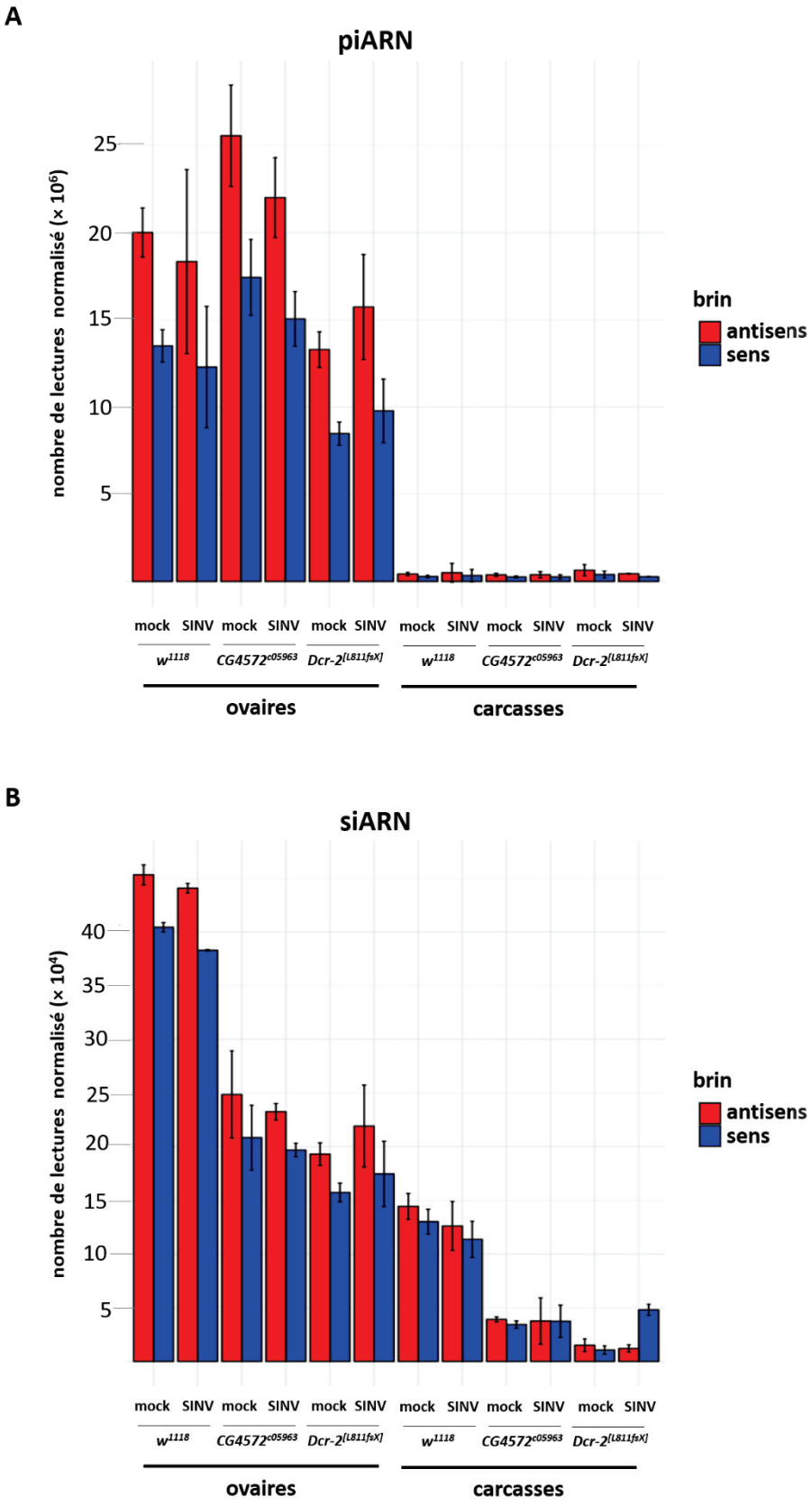


Figure 62. Analyse des siARN et des piARN ciblant les ET chez *D. melanogaster*. Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents de 23-30 nucléotides **(A)** et de 21 nucléotides de longueur **(B)**, ciblant les ET, dans les carcasses et les ovaires des drosophiles sauvages w^{1118} et des mutants

CG4572^{c05963} et *Dcr- 2*$^{[L811fsX]}$, normalisé par les comptes des lectures des miARN. n = 2 expériences indépendantes (50 tissus par échantillon).

D. Résumé

Nous avons démontré que l'infection par SINV affecte les quantités de transcrits d'ET *via* la modulation des quantités de petits ARN interférents dirigés contre les ET, au niveau des carcasses (**Figure 63.**) et des ovaires (**Figure 64.**) des drosophiles.

Carcasses :

Dans les carcasses, la comparaison entre les drosophiles *w*^{1118} infectées par SINV et les drosophiles *w*^{1118} non infectées (mock) révèle une diminution des quantités de transcrits d'ET (ARN), associée à une augmentation des piARN et siARN lors de l'infection par SINV (**Figure 63 A.**). De plus, il n'y a pas de modulation significative (évaluée par DESeq2 ; *p-value* > 0,05) de l'expression des gènes principaux impliqués dans les voies suivantes : piARN, siARN, *uptake* de l'ARN double brin, lors de l'infection par SINV par rapport à la condition mock (**Figure 63 A.**).

La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ mock et les drosophiles *w*^{1118} mock indique que dans un contexte non infectieux, la perte de siARN fonctionnels (*Dcr-2*$^{[L811fsX]}$) entraîne une diminution des quantités de transcrits d'ET, associée à une augmentation des piARN (**Figure 63 B.**). L'expression du gène *dcr2* est significativement réprimée (évaluée par DESeq2 ; *p-value* < 0,05), tandis que l'expression du gène *ninaC*, impliquée dans la voie de l'import de l'ARN double brin est significativement stimulée (évaluée par DESeq2 ; *p-value* < 0,05) chez les mutants *Dcr-2*$^{[L811fsX]}$ par rapport aux drosophiles *w*^{1118} (mock) (**Figure 63 B.**).

La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV et les drosophiles *w*^{1118} infectées par SINV indique que, dans un contexte infectieux, la perte de siARN fonctionnels (*Dcr- 2*$^{[L811fsX]}$) entraîne une augmentation des quantités de transcrits d'ET, associée à une diminution des piARN (**Figure 63 C.**). Les gènes *dcr2*, *ago2* (siARN) et *egh* (voie de l'import de l'ARN double brin) sont significativement surexprimés, tandis que *piwi* et *aub* (piARN) sont significativement réprimés chez les mutants *Dcr-2*$^{[L811fsX]}$ par rapport aux drosophiles *w*^{1118} (**Figure 63 C.** ; DESeq2 ; *p-value* < 0,05).

Enfin, la comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV et les mutants *Dcr- 2*$^{[L811fsX]}$ mock révèle une augmentation des quantités de transcrits d'ET, associée à une diminution des piARN lors de l'infection par SINV (**Figure 63 D.**). De plus, les gènes *dcr2*, *ago2* et *egh* sont significativement surexprimés, tandis que *piwi* et *ninaC* sont significativement réprimés lors de l'infection par SINV par rapport à la condition mock (**Figure 63 D.** ; DESeq2 ; *p-value* < 0,05).

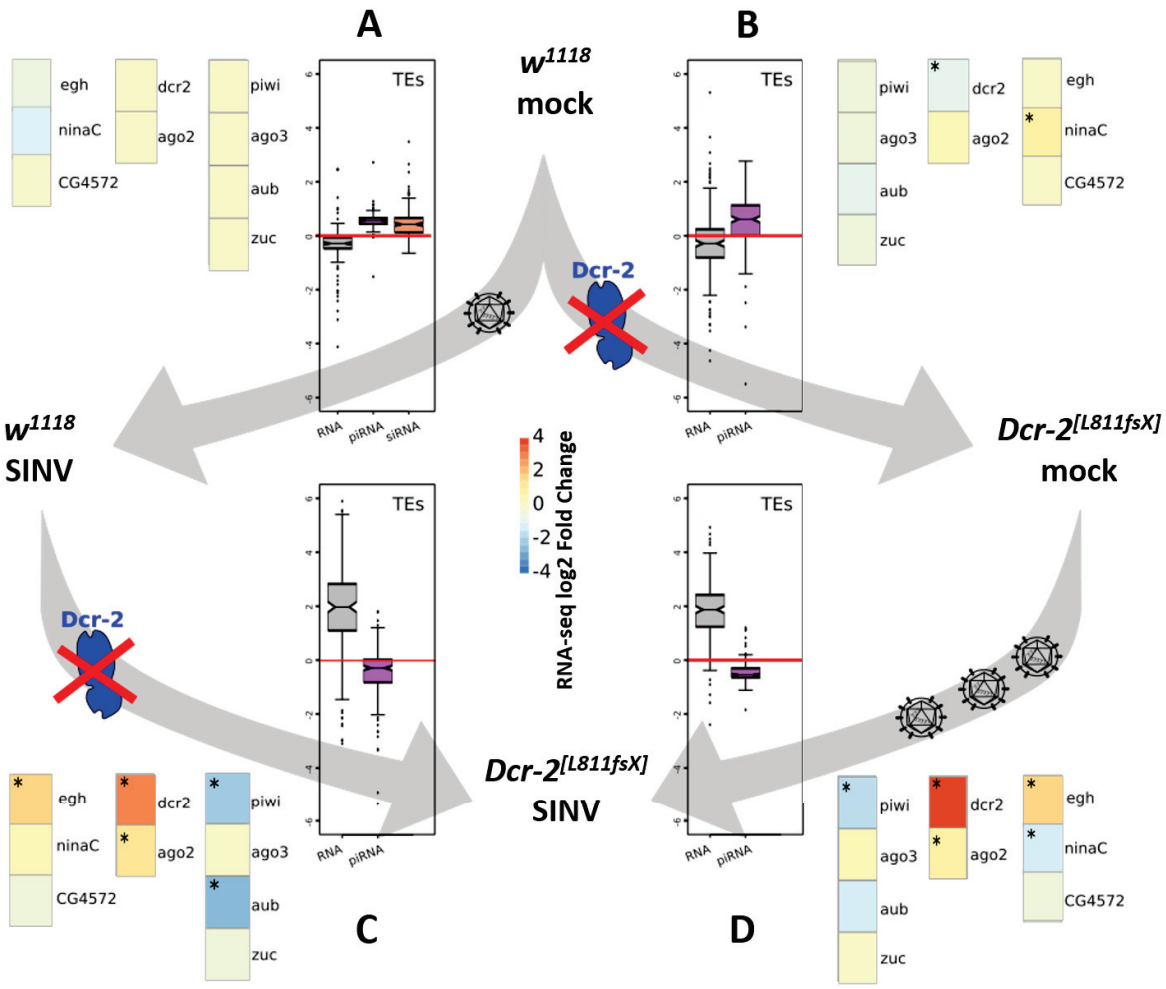


Figure 63. Résumé de la modulation de l'activité des ET et des petits ARN interférents ciblant les ET dans les carcasses de *D. melanogaster*. Pour les différentes comparaisons (drosophiles infectées par rapport aux mock et mutants par rapport aux drosophiles sauvage), les *boxplots* représentent les modulations en log2 *fold change* des quantités de transcrits des ET et des petits ARN interférents. Les *heatmaps* représentent le log2 *fold change* (entre les différentes conditions) des comptes des transcrits des gènes des voies piARN (*piwi*, *ago3*, *aub*, *zuc*) ; siARN (*dcr2*, *ago*2) et de l'import de l'ARN double brin (*egh*, *ninaC*, *CG4572*). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. Les modulations des taux de transcrits des gènes sont déterminées par DESeq2 et sont indiquées par des étoiles lorsqu'elles sont significatives (**P<0,05*). **(A)** Comparaison entre *w*^{1118} infectée par SINV *versus* *w*^{1118}non infectée (mock). **(B)** Comparaison entre *Dcr-2*$^{[L811fsX]}$ mock *versus* *w*^{1118} mock. **(C)** Comparaison entre *Dcr-2*$^{[L811fsX]}$ infectée par SINV *versus* *w*^{1118} infectée par SINV. **(D)** Comparaison entre *Dcr- 2*$^{[L811fsX]}$ infectée par SINV *versus* *Dcr-2*$^{[L811fsX]}$ mock. n = 2 expériences indépendantes.

Ovaires :

Dans les ovaires, la comparaison entre les drosophiles w^{1118} infectées par SINV et les drosophiles w^{1118} non infectées (mock) révèle une faible diminution des quantités de transcrits d'ET (RNA), et des petits ARN interférents (piARN et siARN) ciblant les ET lors de l'infection par SINV (**Figure 64 A.**). Il n'y a pas de modulation significative de l'expression des gènes impliqués dans les voies suivantes : piARN, siARN, *uptake* de l'ARN double brin, lors de l'infection par SINV par rapport à la condition mock (**Figure 64 A.** ; évaluée par DESeq2 ; *p-value* > 0,05).

La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ mock et les drosophiles w^{1118} mock montre que dans un contexte non infectieux, la perte de siARN fonctionnels (*Dcr-2*$^{[L811fsX]}$) n'impacte pas les quantités de transcrits d'ET (RNA), alors qu'elle induit la diminution des piARN (**Figure 64 B**.) De plus, l'expression du gène *dcr2* est significativement réprimée, tandis que les expressions de *piwi*, *aub*, *ago3* et *egh* sont significativement stimulées chez les mutants *Dcr- 2*$^{[L811fsX]}$ par rapport aux drosophiles w^{1118} (**Figure 64 B**. ; évaluées par DESeq2 ; *p-value* < 0,05).

La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV et les drosophiles w^{1118} infectées par SINV indique que dans un contexte infectieux, la perte de siARN fonctionnels (*Dcr-2*$^{[L811fsX]}$) n'impacte pas les quantités de transcrits d'ET (RNA), alors qu'elle induit la diminution des piARN (**Figure 64 C**.). Les expressions des gènes *piwi*, *aub*, *ago3* et *CG4572* sont significativement stimulées chez les mutants *Dcr-2*$^{[L811fsX]}$ par rapport aux drosophiles w^{1118} (**Figure 64 C**. ; évaluées par DESeq2 ; *p-value* < 0,05).

La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV et les mutants *Dcr-2*$^{[L811fsX]}$ mock révèle une légère diminution des quantités de transcrits d'ET (RNA), associée à une augmentation des piARN lors de l'infection virale (**Figure 64 D.**). Les gènes *dcr2* et *egh* sont significativement surexprimés chez les mutants *Dcr-2*$^{[L811fsX]}$ infectés par SINV par rapport à la condition mock (**Figure 64 D**. ; évalués par DESeq2 ; *p-value* < 0,05).

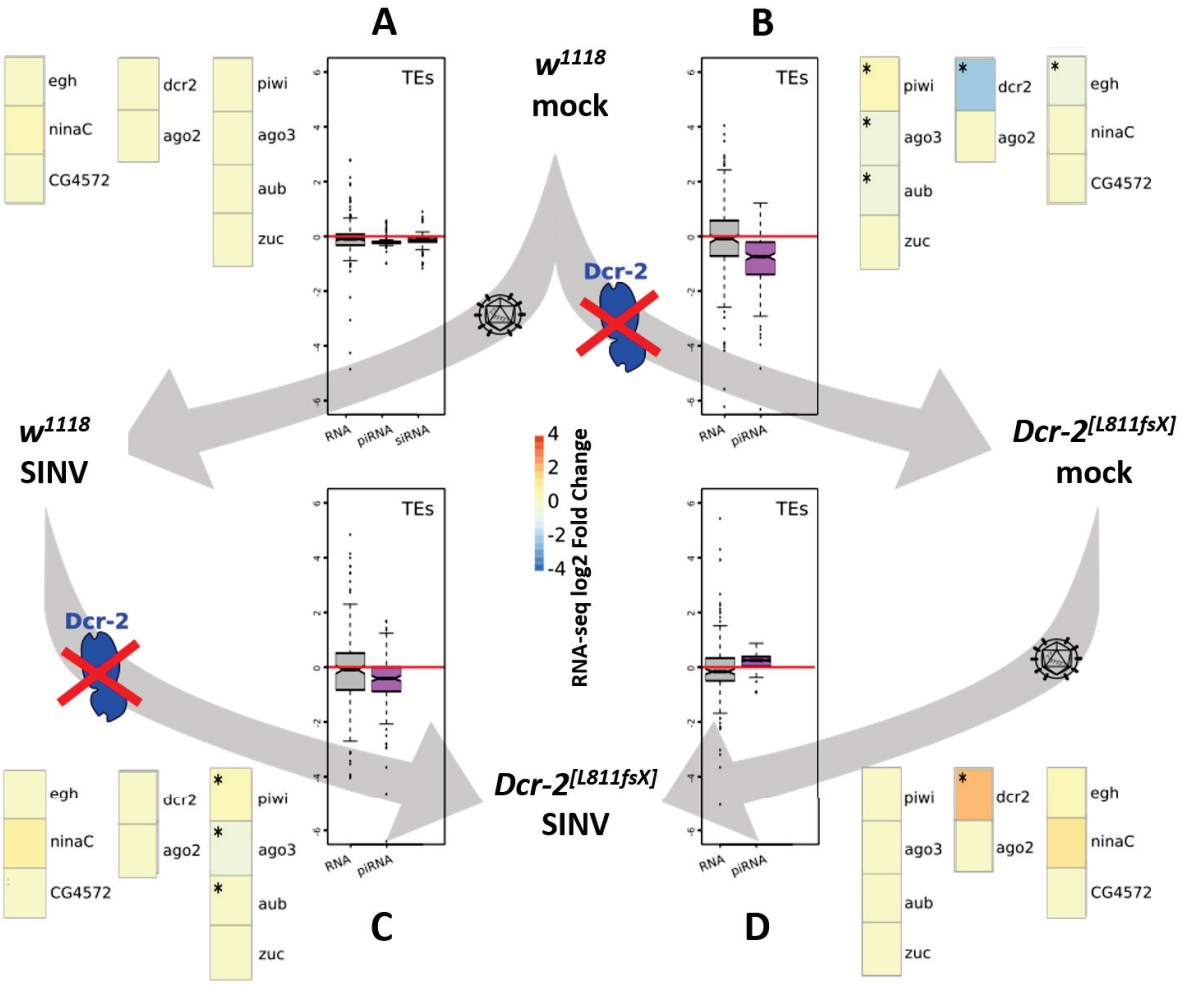


Figure 64. Résumé de la modulation de l'activité des ET et des petits ARN interférents ciblant les ET dans les ovaires de *D. melanogaster*. Pour les différentes comparaisons (drosophiles infectées par rapport aux mock et mutants par rapport aux drosophiles sauvage), les *boxplots* représentent les modulations en log2 *fold change* des quantités de transcrits des ET et des petits ARN interférents. Les *heatmaps* représentent le log2 *fold change* (entre les différentes conditions) des comptes des transcrits des gènes des voies piARN (*piwi*, *ago3*, *aub*, *zuc*) ; siARN (*dcr2*, *ago*2) et de l'import de l'ARN double brin (*egh*, *ninaC*, *CG4572*). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. Les modulations des taux de transcrits des gènes sont déterminées par DESeq2 et sont indiquées par des étoiles lorsqu'elles sont significatives (**P<0,05*). **(A)** Comparaison entre w^{1118} infectée par SINV *versus* w^{1118}non infectée (mock). **(B)** Comparaison entre *Dcr-2*$^{[L811fsX]}$ mock *versus* w^{1118} mock. **(C)** Comparaison entre *Dcr-2*$^{[L811fsX]}$ infectée par SINV *versus* w^{1118} infectée par SINV. **(D)** Comparaison entre *Dcr- 2*$^{[L811fsX]}$ infectée par SINV *versus* *Dcr-2*$^{[L811fsX]}$ mock. n = 2 expériences indépendantes.

Discussion (I)

Comme présenté dans l'introduction, l'expression des ET est étroitement contrôlée chez les animaux, incluant une régulation des taux de mutation, pouvant être bénéfiques ou délétères pour l'organisme. Deux voies d'ARNi sont impliquées dans le contrôle des ET. En effet, la voie des piARN permet le contrôle des ET principalement au niveau des gonades (Iwasaki et al., 2015), tandis que la voie des endo-siARN contrôle les ET au niveau somatique (Czech et al., 2008). Par ailleurs, la voie des siARN est également l'une des principales voies de l'immunité antivirale chez les insectes (Bronkhorst and van Rij, 2014; Gammon and Mello, 2015). Dans la première partie du travail expérimental, nous avons révélé un lien entre ces deux voies d'ARNi, impliquant le contrôle de l'activité des ET et la production de siARN viraux.

Quatre lignées de drosophiles ont été utilisées au cours de mes travaux de thèse, incluant *D. melanogaster* w^{1118}, $Dcr2^{[L811fsX]}$, $CG4572^{c05963}$ et *D. simulans* Makindu. SINV se réplique dans les carcasses de *D. melanogaster* et *D. simulans*, incluant des titres viraux moins importants pour les drosophiles Makindu (*D. simulans*) (**Figure 40.**). Alors qu'aucune différence de réplication virale n'a été observée entre les mutants $CG4572^{c05963}$ et les drosophiles w^{1118}, les mutants $Dcr2^{[L811fsX]}$ ont révélé une réplication virale supérieure (**Figure 51.**). Cependant, au temps tardif de réplication virale (jour 10 après infection), nous avons détecté une diminution des titres viraux dans les mutants $Dcr2^{[L811fsX]}$ (**Figure 55.** ; test statistique de Wilcoxon-Mann-Whitney entre jour 8 et jour 10 ; *p-value* = 2,2e-3), suggérant une stimulation d'autres voies de l'immunité antivirale, telles que les voies Jak/STAT ou Toll (Dostert et al., 2005; Zambon et al., 2005). Nos analyses de RNA-seq révèlent notamment une surexpression de certains gènes impliqués dans la voie Jak/STAT (*domeless* : log2 *fold change* = 1,42 ; *p-value* = 7,1e-5 ; *hopscotch* : log2 *fold change* = 2,49 ; *p-value* = 2,2e-18) et également dans la voie Toll (*toll* : log2 *fold change* = 2,96 ; *p-value* = 3,87e-33 ; *PGRP-SA* : log2 *fold change* = 1,83 ; *p-value* = 1,07e-3) chez les mutants $Dcr2^{[L811fsX]}$ infectés par SINV par rapport à la condition non infectée (données non montrées). Ces mêmes résultats sont retrouvés chez les mutants $Dcr2^{[L811fsX]}$ infectés par SINV par rapport aux drosophiles w^{1118} infectées (données non montrées). Dans ces mêmes mutants, des titres viraux importants ont été obtenus dans les ovaires (**Figures 51 B.**). Cependant, l'immunofluorescence n'a pas permis de détecter la capside de SINV au sein même des ovaires (**Figure 56.**). La détection de la capside de SINV en périphérie des ovaires des mutants $Dcr2^{[L811fsX]}$ (**Figure 56.**) peut techniquement s'expliquer par le très haut niveau de réplication de SINV dans ces mutants, qui malgré des lavages à répétition, laisserait SINV attaché aux tissus conjonctifs des ovaires. L'immuno-marquage d'une protéine non structurale de SINV pourrait permettre de confirmer la potentielle absence de réplication virale dans les ovaires. Dans les autres lignées (w^{1118}, $CG4572^{c05963}$ et *D. simulans* Makindu), la titration des paires d'ovaires a révélé des titres viraux très faibles, incluant un faible pourcentage de drosophiles positives pour SINV. Ainsi, SINV ne semble pas se répliquer dans les ovaires chez *D. melanogaster* (**Figure 56. et S11.**) et *D. simulans* (**Figure**

42.). Par ailleurs, aucune réplication d'arbovirus dans les ovaires n'a été rapportée chez *Drosophila*, ce qui laisse penser que ces virus ne se transmettent pas verticalement chez un insecte ne pouvant être infecté naturellement. Ainsi, il serait intéressant d'étudier plus précisément la transmission verticale potentielle de SINV *via* les oeufs de drosophile en analysant les générations descendantes des drosophiles infectées expérimentalement et en déchorionnant ou non les œufs de ces descendants pour ensuite titrer le virus dans les drosophiles adultes sous-jacentes. En effet, le virus ne semble pas se répliquer au sein même des ovaires mais il se peut qu'il reste attaché au chorion des œufs de drosophile, ne pouvant pas aboutir à une transmission verticale (Habayeb et al., 2009). Pour déterminer s'il y a une transmission verticale de SINV chez la drosophile, on pourrait également réaliser des immuno-marquages pour la capside de SINV, au niveau des œufs de drosophile infectées et de leurs descendants.

Concernant la technique de titration virale choisie, établie à partir de la méthode de Reed et Muench (Reed and Muench, 1938) et exprimée en TCID50/ml, la titration de drosophile entière et de carcasse au jour de l'infection (jour 0) ne permet pas la détection du virus dans tous les échantillons en raison du seuil de détection de cette technique, fixé à 1,33 log10(TCID50/ml). Cependant, la réplication virale à partir du 2ème jour après infection a permis de valider l'infection des drosophiles.

SINV semble se répliquer de façon ubiquitaire chez *D. melanogaster* et plus particulièrement au niveau du corps gras des mutants *Dcr2*$^{[L811fsX]}$ (**Figures 59.**) où des piARN contrôlant l'activité des ET ont été décrits (Jones et al., 2016). De plus, ces résultats sont en accord avec ceux obtenus chez *D. simulans* Makindu même si la réplication virale semble plus restreinte au thorax, proche de la zone d'injection du virus. Par ailleurs, Galiana-Arnoux et collaborateurs ont détecté par immunofluorescence la glycoprotéine E1 de SINV au niveau des cellules du corps gras chez *D. melanogaster*, après infection intrathoracique (Galiana-Arnoux et al., 2006). Ces mêmes auteurs ont observé les particules de SINV par microscopie électronique au niveau des cellules musculaires entourant l'intestin moyen de la drosophile.

Chez *w*^{1118} et Makindu, la réponse de type ARNi est principalement médiée par les siARN (**Figures 36. et 44.**). Nous avons injecté une dose infectieuse beaucoup plus importante chez Makindu par rapport à *w*^{1118}, permettant l'obtention de titres viraux comparables dans les carcasses (à 6 jours après infection) qui cependant ne semblent pas induire une réponse de type siARN similaire. En effet, le nombre de lectures des brins de siARN est beaucoup plus faible chez Makindu, incluant une baisse de 99% (**Figures 36 A. et 44 A.**). De plus, ces siARN viraux ont une proportion équivalente de brin sens et anti-sens, incluant des points chauds et des points froids de production le long du génome viral (Dietrich et al., 2017). Par ailleurs, lors de l'infection, la détection de siARN ciblant SINV dans les ovaires des drosophiles *w*^{1118} (**Figures 36.**) et des drosophiles Makindu (**Figure 44.**) pourrait être expliquée par le phénomène de diffusion systémique engendré lors d'une infection virale chez *Drosophila* (Saleh et al., 2009). Comme attendu, en comparaison avec les *w*^{1118}, les mutants *CG4572*^{c05963}

présentent une faible production de siARN viraux (**Figure 52 A.**) puisqu'il a été montré que le gène *CG4572* est impliqué dans la voie de l'import de l'ARN double brin, nécessaire à l'induction de la production systémique des siARN viraux chez la drosophile (Saleh et al., 2009). Par ailleurs, l'analyse des petits ARN interférents de 21 nucléotides spécifiques de SINV dans les mutants *Dcr2*$^{[L811fsX]}$, présentant une mutation au niveau de la séquence codante de la protéine Dcr-2, révèle seulement en faible quantité le brin sens ne permettant pas la formation de siARN fonctionnels (**Figure 60.**). De plus, chez ces mutants, la distribution des petits ARN interférents viraux de 21 nucléotides révèle un pic correspondant au brin sens, à environ 7600 nucléotides du génome de SINV. Cette région génomique contient un codon stop qui interrompt la séquence de l'ARNm génomique 49S, codant les protéines non structurales et incluant 7000 nucléotides (Abu Bakar and Ng, 2018; Strauss et al., 1983). C'est également ici que débute la séquence de l'ARNm subgénomique 26S, codant les protéines structurales et incluant 4000 nucléotides (Ou et al., 1982) ; (**Figure 60 B.**). Ce pic représente quasiment la seule détection dans les mutants *Dcr2*$^{[L811fsX]}$ mais des analyses complémentaires sont nécessaires pour élucider sa présence. Cependant, nous avons détecté une proportion équivalente de brins sens et antisens des siARN dirigés contre les ET dans les ovaires des mutants *Dcr2*$^{[L811fsX]}$ (**Figure 62 B.**). En effet, il a été suggéré que Dcr-1 peut prendre le relais pour la production d'endo-siARN, de façon réduite par rapport aux *w*^{1118}, mais pouvant expliquer la détection de ces petits ARN interférents, incluant les brins sens et antisens de 21 nucléotides au niveau des ovaires (Lee et al., 2004; Liu et al., 2006). De plus, Beek et colloborateurs ont également observé une forte réduction des quantités de siARN ciblant les ET chez les mutants *Dcr2*$^{[L811fsX]}$, excluant une absence totale d'endo-siARN (Beek et al., 2018). Dans l'ensemble des lignées de drosophile utilisées, la faible proportion de petits ARN interférents de 23 à 30 nucléotides de longueur ne présente pas de signature spécifique des piARN, incluant un chevauchement de 10 nucléotides entre le brin sens et le brin antisens des piARN (« signature ping-pong »). Ces résultats sont en accord avec la littérature (Petit et al., 2016).

Les siARN sont produits de façon systémique dans l'ensemble du corps de la drosophile, tandis que les piARN dirigés contre les ET sont considérés comme étant spécifiques de la lignée germinale chez *Drosophila* (Lewis et al., 2018). Cependant, plusieurs études ont récemment révélé des piARN somatiques (Mirkovic-Hösle and Förstemann, 2014), incluant dans certains cas une production de piARN secondaires impliqués dans la boucle d'amplification en ping-pong, telle que dans des neurones spécifiques du cerveau de drosophile (Perrat et al., 2013), ou bien une production de piARN primaires contrôlant notamment la mobilisation des ET notamment dans le corps gras (Jones et al., 2016). En effet, ces piARN somatiques semblent spécifiques de certaines cellules ou tissus, puisqu'ils sont notamment retrouvés au niveau des têtes de drosophiles (Beek et al., 2018; Perrat et al., 2013) et au niveau des cellules du corps gras de l'abdomen (Jones et al., 2016) mais ne sont pas détectés dans le thorax (Lewis et al., 2018). Le corps gras, présent au niveau de l'abdomen et

du thorax (Padmanabha and Baker, 2014), joue notamment un rôle dans l'immunité chez la drosophile puisqu'il est le lieu de production primaire des peptides antimicrobiens (AMP), permettant l'induction de l'immunité humorale systémique (Charroux and Royet, 2010; Dionne, 2014). Dans les carcasses de *D. melanogaster* w^{1118} et *D. simulans* Makindu, nous avons observé une diminution des quantités de transcrits d'ET associée à une augmentation des productions de piARN et siARN ciblant les ET (**Figures 37. et 45.**). Pour déterminer si l'impact de l'infection sur les ET est spécifique des piARN retrouvés dans le cerveau des drosophiles, nous avons disséqué et analysé les têtes des drosophiles Makindu. Comme mentionné dans les résultats, la capside de SINV a été détectée au niveau des coupes de têtes des Makindu (**Figure 43 C.**). De plus, SINV semble se répliquer au niveau des têtes des Makindu, puisque les particules virales infectieuses (**Figure S21 A.**) et copies du génome (**Figure S21 B.**) de SINV y ont été détectées, au jour 6 après infection. Par la suite, nous avons analysé l'expression des mêmes ET que précédemment (**Figure 46.**), incluant sept familles d'ET, en RT-qPCR, au niveau des têtes de Makindu en comparaison avec le reste de la carcasse de drosophile (exemptée de la tête et la paire d'ovaires). De façon surprenante, nous n'avons pas observé de modulation de l'expression des ET suite à l'infection par SINV dans les têtes (**Figure S22 A.**) et les carcasses (**Figure S22 B.**) ; (n = 3 ; Test de Wilcoxon-Mann-Whitney ; *p-value* > 0,05 pour chacune des sept familles d'ET analysées en RT-qPCR). Ces résultats suggèrent un impact global de l'infection par SINV sur la transcription des ET plutôt qu'un impact spécifique sur certaines familles, en fonction de la prise en charge de ces ET par le mécanisme de régulation génomique, et notamment impliquant la boucle d'amplification en ping-pong. En outre, les travaux de thèse de M. Petit (2016), révèlent une absence de modulation significative au niveau de l'expression de 24 ET de différentes classes (Rétrotransposons à LTR et sans LTR, transposons ADN), par RT-PCR quantitative, lors de l'infection de *D. melanogaster* par les virus DCV et DAV. De plus, lors de ces mêmes travaux, une grande variabilité d'expression a été observée pour certaines familles d'ET entre les différents réplicats. Par conséquent, l'analyse globale du transcriptome (RNA-seq) semble être plus adaptée pour une analyse globale des transcrits d'ET lors d'une infection virale.

Par ailleurs, afin de comparer de façon plus spécifique les espèces *D. melanogaster* et *D. simulans*, nous avons notamment éliminé l'infection persistante de NV dans les Makindu (**Figure S6.**) pour ensuite étudier la modulation de l'expression des mêmes sept familles d'ET que précédemment, en réalisant des expériences de RT-qPCR. Ces résultats ont été comparés dans l'optique de confirmer l'absence d'impact de l'infection par NV sur la modulation de l'activité des ET engendrée lors de l'infection par SINV. De la même façon que pour *D. melanogaster*, les œufs de drosophiles Makindu ont été javellisés, permettant de déchorionner les œufs et d'éliminer NV (**Figure S6.**). Ensuite, les femelles adultes ont été infectées par SINV par injections intrathoraciques (14900 PFU dans un volume de 207 nl). Nous n'avons pas observé de différence significative de réplication de SINV entre les drosophiles Makindu infectées par NV et celles dépourvues d'infection virale, en utilisant la technique de

titration virale en TCID50/ml (**Figure S23 A.**), ni de différence de charge génomique virale en RT-PCR quantitative (**Figure S23 B et C.**). De la même façon que pour les drosophiles Makindu non javellisées, aucune détection de la capside de SINV n'a été faite dans les ovaires, par immunofluorescence (**Figure S24.**). L'ensemble de ces résultats suggère que NV n'impacte pas la réplication de SINV chez les drosophiles Makindu. L'analyse par RT-qPCR de l'expression des sept familles d'ET n'a révélé aucune modulation significative de l'expression de ces ET, au niveau des carcasses (**Figure S25 A.**) et au niveau des ovaires (**Figures S25 B.**). Ceci suggère que ce ne sont pas nécessairement les mêmes familles d'ET qui sont impactées à chaque événement d'infection par SINV. Ainsi, l'hypothèse exposant que la modulation de l'expression des ET soit globale lors de l'infection par SINV, est renforcée. En effet, il semblerait que les différentes familles d'ET soient régulées aléatoirement dans l'organisme de chaque individu. On peut penser à un processus stochastique qui pourrait dépendre de la distribution spatiale des transcrits d'ET dans la cellule (Adrion et al., 2017).

Nous avons choisi d'étudier la modulation de l'activité des ET et de la production des petits ARN interférents à 6 jours après l'infection par SINV, révélant une importante réplication de SINV et une induction de la voie des siARN. Néanmoins, l'analyse de ces modulations à différents temps après infection, en utilisant notamment les techniques de séquençage à haut débit, permettrait de suivre l'évolution de l'impact de l'infection virale sur les ET en fonction de la mise en place des voies d'ARNi chez la drosophile.

Nous avons démontré que l'infection par SINV affecte les quantités de transcrits d'ET via la modulation des quantités de petits ARN interférents dirigés contre les ET, au niveau des carcasses (**Figure 63.**) et des ovaires des drosophiles (**Figure 64.**). La modulation des transcrits d'ET ne dépend pas de la classe d'ET puisque les transposons ADN, ainsi que les rétrotransposons à LTR et sans LTR sont impactés chez *D. simulans* et *D. melanogaster* lors de l'infection virale (données non montrées). Alors que les voies des piARN et des siARN sont décrites de façon indépendante dans la littérature, incluant des effecteurs spécifiques à chacune de ces voies, nous proposons que le lien entre ces deux voies d'ARN interférence dans les carcasses soit dû en partie au fait que les siARN et les piARN dirigés contre les ET partagent des substrats d'ARN (**Figure 65 A-B.**). Cette hypothèse est notamment renforcée par l'importante corrélation positive entre les siARN et piARN ciblant les ET, identifiée dans toutes les conditions (**Figure S26.**). De plus, il a été précédemment montré que les ET produisent des transcrits sens et antisens, permettant la formation d'ARN double brin et également la production de paires mixtes de piARN secondaires et siARN (Olovnikov et al., 2013; Shpiz et al., 2014).

Carcasses :

Nos résultats révèlent dans les mutants *CG4572*^{c05963} une importante diminution par rapport aux drosophiles w^{1118} de l'effet de l'infection par SINV sur les quantités de petits ARN interférents (**Figure 53 A-B.**), suggérant l'implication de la machinerie de l'import d'ARN double brin dans la modulation de la production des petits ARN interférents ciblant les ET lors de l'infection virale. De plus, suite à l'infection (*Dcr-2*$^{[L811fsX]}$ SINV *versus Dcr-2*$^{[L811fsX]}$ mock) ou dans un contexte infectieux (*Dcr-2*$^{[L811fsX]}$ SINV *versus w*^{1118} SINV), l'invalidation de la production des siARN engendre une augmentation des quantités de transcrits des ET (**Figure 61 A. et S17 C.**) et une diminution des piARN dirigés contre les ET (**Figure 61 B. et S17 D.**), confirmant l'implication des siARN ciblant les ET dans le contrôle de l'activité des ET. Dans le modèle *Dcr-2*$^{[L811fsX]}$, on peut également proposer que l'importante réplication virale, induisant une grande quantité d'ARN viraux, sature la machinerie des piARN aboutissant à une diminution de la production des piARN ciblant les ET et à une augmentation de l'activité des ET sous-jacente ; d'autant plus que nous avons détecté une quantité non négligeable de piARN sens ciblant SINV dans les carcasses des mutants *Dcr-2*$^{[L811fsX]}$ infectés (**Figure 60 A.**). De plus, lors de l'infection des mutants *Dcr-2*$^{[L811fsX]}$, nous avons observé une surexpression du gène *dcr2*, et à l'inverse une répression du gène *piwi* (**Figure 61 C.**). Ces résultats suggèrent que lorsque la voie des siARN est déficiente, des inducteurs, jusqu'à présent non identifiés, seraient capables de stimuler la voie des siARN et de réprimer la voie des piARN. Par ailleurs, aucune fonctionnalité de la protéine Dcr-2 ne peut être induite par la suite puisque la séquence codant cette protéine est mutée. Une telle régulation aurait été sélectionnée pour permettre en priorité le contrôle des infections par la machinerie d'ARN interférence, lorsque cela se produit, représentant une menace plus grande que celle induite par l'activité des ET, dans les carcasses.

L'ensemble de mes travaux de thèse concernant la première partie des résultats expérimentaux nous a conduit à établir un modèle de l'impact de l'infection par SINV sur l'activité des ET (**Figure 65.**). En condition non infectée, les transcrits d'ET peuvent être des substrats pour les voies des siARN et des piARN, dont les principaux effecteurs sont respectivement Dcr-2 et PIWI (**Figure 65 A.**). Lors de l'infection virale, les intermédiaires viraux d'ARN double brin sont diffusés de façon systémique, impliquant la voie de l'import de l'ARN double brin, ce qui amplifie la production de siARN viraux (Saleh et al., 2009) (**Figure 65 B.**). Nous proposons que les intermédiaires d'ARN des ET ou les transcrits d'ET eux-mêmes soient également absorbés de façon systémique par la machinerie de l'import d'ARN double brin puis pris en charge par Dcr-2, engendrant la production de siARN dirigés contre les ET (**Figure 65 B.**). Ces substrats d'ARN peuvent également être utilisés par la voie des piARN, augmentant la production de piARN ciblant les ET. La production accrue des deux classes de petits ARN interférents a un impact négatif sur les quantités de transcrits d'ET (**Figure 65 B.**).

Ovaires :

SINV ne semble pas se répliquer dans les ovaires chez *D. melanogaster* (**Figure 56. et S11.**) et *D. simulans* (**Figure 42.**). De plus, nos résultats montrent que les faibles modulations des quantités de petits ARN interférents ciblant les ET observée suite à l'infection virale sont quasiment les mêmes dans les lignées w^{1118} et $CG4572^{c05963}$ (**Figure 53.**), suggérant que la voie de l'import de l'ARN bouble brin ne détermine pas les quantités de petits ARN interférents dirigés contre les ET dans les ovaires. En contexte infectieux ou non infectieux, l'invalidation de la production des siARN entraîne une diminution de la production des piARN ciblant les ET mais n'aboutit pas à une modulation des quantités de transcrits d'ET (**Figure S18**). De plus, dans les mutants $Dcr\text{-}2^{[L811fsX]}$, l'infection virale induit une légère augmentation des quantités de piARN ciblant les ET (**Figure 61 E.**) mais n'impacte toujours pas l'activité des ET (**Figure 61 D.**). Ces résultats suggèrent que les transcrits d'ET sont principalement contrôlé par les piARN dirigés contre les ET au niveau des ovaires (**Figure 65 C-D.**). En effet, dans cet organe, lorsque la voie des piARN est pleinement fonctionnelle, la voie des siARN est seulement utilisée comme un système de secours (Barckmann et al., 2018). Par ailleurs, lors de l'infection par SINV, il n'y a quasiment pas de modulation des quantités de transcrits des ET dans les ovaires chez w^{1118} (**Figure 37 D.**) et Makindu (**Figure 45 D.**). Cependant, les quantités de petits ARN interférents ciblant les ET sont modulées, et cela de façon différente entre w^{1118} et Makindu, incluant une faible diminution des deux classes de petits ARN interférents chez w^{1118} (**Figure 37 E-F.**) et une production en miroir des petits ARN interférents dans Makindu (**Figure 45 E-F.**). Ces résultats suggèrent l'existence d'une forte pression de sélection sur le maintien de la stabilité du transcriptome des ET dans les ovaires, atteint cependant par des mécanismes distincts chez *D. melanogaster* et *D. simulans*. Ces résultats laissent également penser qu'en raison d'une absence de modulation de l'activité des ET dans les ovaires, cela ne peut pas avoir d'impact évolutif. Or, les quantités des piARN ciblant les ET sont diminuées de façon importante dans les ovaires des Makindu. On peut ainsi penser que le contrôle des ET sera affecté dans les générations suivantes issues de ces ovaires, incluant des quantités différentes de piARN transférés maternellement. Par conséquent, chez *D. simulans* Makindu, il serait intéressant d'explorer les impacts de la modulation des quantités de piARN ovariens maternels sur l'activité des ET des générations de drosophiles suivantes. Par ailleurs, une autre perspective intéressante inclurait l'analyse de l'impact de l'infection virale sur les traits d'histoire de vie des drosophiles (longévité, nombre d'œufs par drosophile, nombre de descendants adultes etc.), en parallèle des modulations au niveau des taux de transcrits des ET observées lors de l'infection virale. En effet, la *fitness* est également une composante du processus qui permet de quantifier l'effet de la sélection naturelle mise en jeu dans l'évolution des génomes.

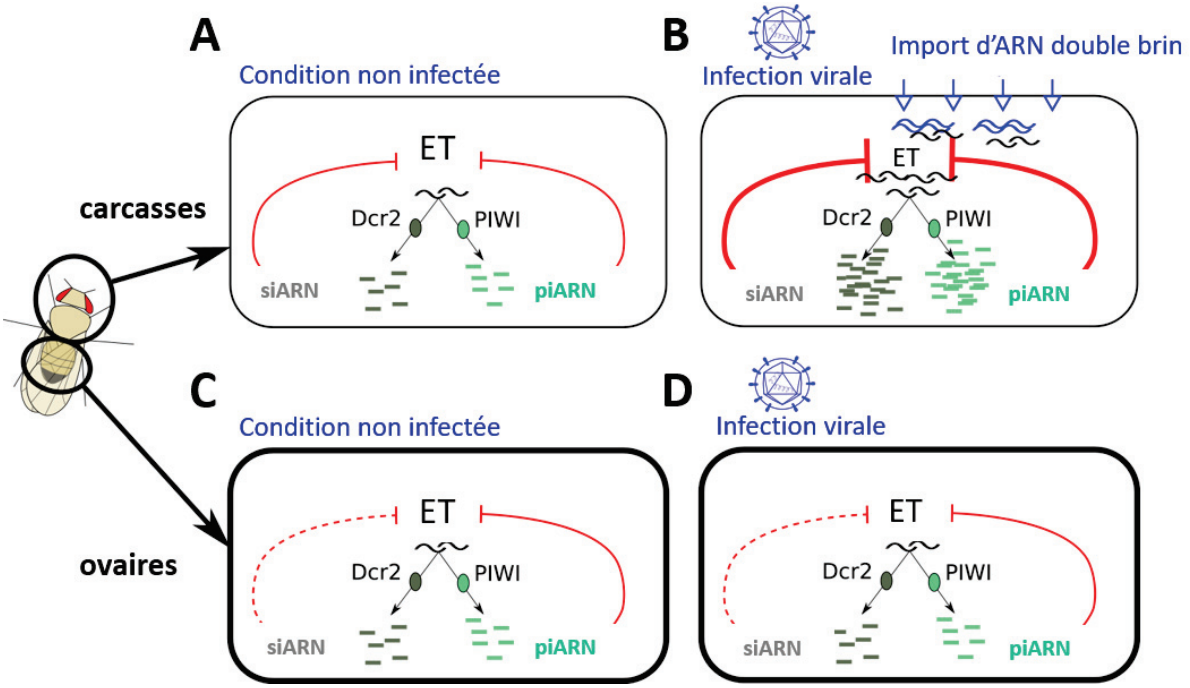


Figure 65. Modèle de l'impact de l'infection par SINV sur les ET. Les mécanismes de modulation de la transcription des ET et de la production des petits ARN interférents dirigés contre les ET et les virus ont été étudiés de façon indépendante dans les carcasses (A-B) et les ovaires (C-D) chez *D. melanogaster*. Nous avons analysé ces modulations lors de l'infection par SINV (B-D) en comparaison avec la condition non infectée (A-C). **(A)** En condition non infectée, les transcrits d'ET sont contrôlés par les piARN et les siARN, dont les principaux effecteurs sont respectivement Dicer-2 et PIWI. **(B)** L'infection virale stimule la voie de l'import d'ARN double brin aboutissant à une augmentation des mécanismes de contrôle de la transcription des ET, *via* la production de piARN et siARN (représentée par les traits pleins rouges), et induisant une diminution des quantités de transcrits des ET. **(C-D)** Dans les deux conditions (non infectée et infectée par SINV), les ET sont contrôlés principalement par les piARN, comme l'indiquent les traits rouges pleins. Dans cet organe, SINV ne semble pas se répliquer et il n'y a quasiment pas d'impact sur la transcription des ET lors de l'infection des drosophiles. Les ovaires semblent protégés des effets de l'infection, se mettant en place dans le reste du corps de la drosophile, comme illustré par les traits noirs épais.

Ces résultats représentent la première démonstration de l'impact d'une infection virale sur l'activité des ET, impliquant une interférence entre les voies d'ARN interférence (siARN et piARN). Il serait particulièrement intéressant par la suite de déterminer si le taux de transposition des ET est impacté par ces modulations de quantité de transcrits d'ET. L'analyse de la transposition des ET peut être réalisée en utilisant des analyses de *pool-seq* à haut débit, incluant le séquençage d'échantillons d'ADN issus de différents embryons de drosophiles, et permettant de calculer le taux de transposition pour chaque famille d'ET (Barckmann et al., 2018). De plus, il serait particulièrement pertinent d'utiliser ces analyses de *DNA-seq* dans le but d'étudier la transposition des ET en somatique, comme au niveau du cerveau des drosophiles, puisque nous avons observé un effet de l'infection virale plus marqué sur l'activité des ET en carcasses par rapport aux ovaires. Par ailleurs, l'analyse de la transposition

des ET peut être réalisée en utilisant des techniques de *Southern Blot* ou de *Transposon Display*, qui semblent cependant moins adaptées car il s'agit d'étude famille-spécifique, et on ne sait pas a priori quelles familles sont pertinentes (Van den Broeck et al., 1998). Si les expériences ci-dessus concluent à une modulation des taux de transposition associée à ces modulations de quantités de transcrits, alors cela suggère que les infections virales ont aussi un impact sur la vitesse de diversification génétique des organismes.

DONNEES SUPPLEMENTAIRES (I)

Données supplémentaires (I)

1. Infection par le virus Sindbis de *D. melanogaster*

Infections persistantes chez *D. melanogaster*

L'ensemble des infections chroniques a été éliminé en déchorionnant les œufs de drosophile par un traitement à l'eau de javel. Nous avons détecté par RT-PCR le virus DAV seulement chez les mutants *Dcr-2*$^{[L811fsX]}$. Aucun virus testé n'a été détecté chez les drosophiles w^{1118} et les mutants *CG4572*^{c05963} par RT-PCR. Cependant, afin d'uniformiser les protocoles et dans le but d'éliminer tous les pathogènes présents, nous avons traité les œufs issus de toutes les drosophiles *D. melanogaster* (w^{1118} , *Dcr-2*$^{[L811fsX]}$, *CG4572*^{c05963}), selon le protocole décrit par Merkling et van Rij, (2015). Nous avons validé le traitement par l'absence de virus détecté en réalisant des expériences de RT-PCR. Pour l'ensemble des expériences, les témoins positifs et négatifs correspondent à des échantillons d'ARN issus de drosophile.

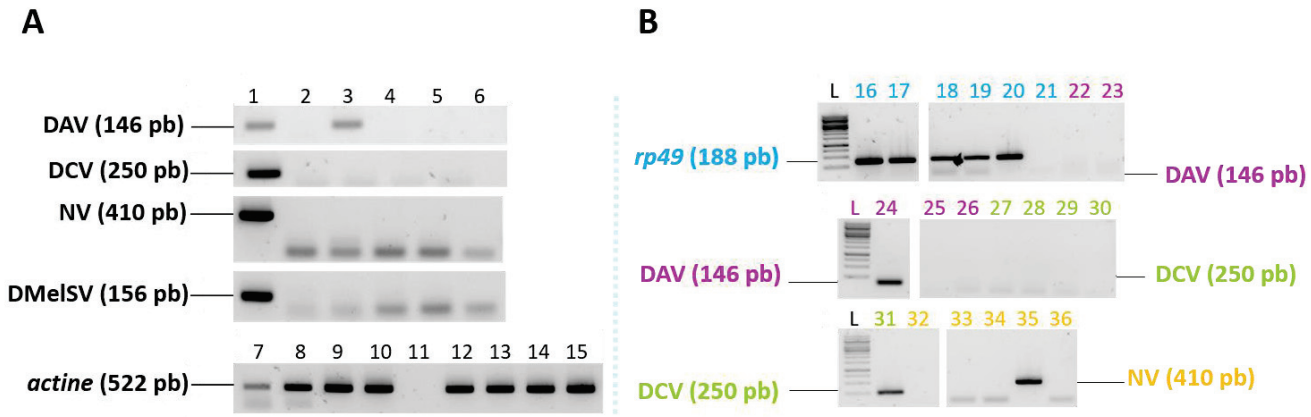


Figure S1. Absence de détection des virus naturels chez *D. melanogaster* après javellisation des œufs de drosophiles. Détection des virus par RT-PCR avant et après javellisation des œufs de drosophile, en utilisant des amorces spécifiques des virus suivants : NV, DAV, DCV, DMelSV. **(A) *D. melanogaster* w^{1118} et *Dcr-2*$^{[L811fsX]}$.** Détection des virus (1-6), *actine* (7-15). Contrôle positif respectif de DAV, DCV, NV, DMelSV (1), w^{1118} (2), *Dcr- 2* $^{[L811fsX]}$ (3), w^{1118} après javellisation (4), *Dcr-2*$^{[L811fsX]}$ après javellisation (5), témoin négatif eau (6). Témoin d'ARN par détection du gène de ménage *actine*, pour les témoins positifs DCV, DAV, NV, DMelSV respectivement (7-10), témoin négatif eau (11), w^{1118} (12), *Dcr-2*$^{[L811fsX]}$ (13), w^{1118} après javellisation (14), *Dcr-2*$^{[L811fsX]}$ après javellisation (15). Résultats représentatifs de 3 expériences (n = 3). **(B) *D. melanogaster CG4572*^{c05963}.** Témoin d'ARN par détection du gène de ménage *rp49* (16-21, en bleu) pour les mutants *CG4572*^{c05963}(16), *CG4572*^{c05963} après javellisation (17), témoin positif DAV (18), témoin positif DCV (19), témoin négatif sans virus (20), PCR- (eau) (21). Détection du DAV (22-26, en violet) pour les mutants *CG4572*^{c05963} (22), *CG4572*^{c05963} après javellisation (23), témoin positif DAV (24), témoin négatif DAV (25), PCR- (eau) (26). Détection de DCV (27-31, en vert) pour les mutants *CG4572*^{c05963} (27), *CG4572*^{c05963} après javellisation (28), témoin négatif DCV (29), PCR- (eau) (30),

témoin positif DCV (31). Détection du NV (32-36, en orange) pour les mutants *CG4572*^{c05963} (32), *CG4572*^{c05963} après javellisation (33), témoin négatif NV (34), témoin positif NV (35), PCR- (eau) (36). Résultats représentatifs de 2 expériences, incluant 10 drosophiles par échantillon.

D. melanogaster et *Wolbachia*

Les bactéries endosymbiotiques *Wolbachia* n'ont pas été détectées chez *D. melanogaster* *w*^{1118} et *Dcr2* $^{[L811fsX]}$. Cependant, *Wolbachia* a été détectée chez les mutants *CG4572*^{c05963}. Afin d'uniformiser les traitements, l'ensemble des drosophiles (*w*^{1118}, *Dcr-2*$^{[L811fsX]}$ et *CG4572*^{c05963}) a été traité à la tétracycline pendant deux générations, afin d'éliminer toutes les bactéries endosymbiotiques (Merkling and van Rij, 2015). La validation de l'élimination de *Wolbachia* a été effectuée par RT-PCR (**Figure S2.**). Pour l'ensemble des expériences, les témoins positifs et négatifs correspondent à des échantillons d'ADN issus de drosophile.

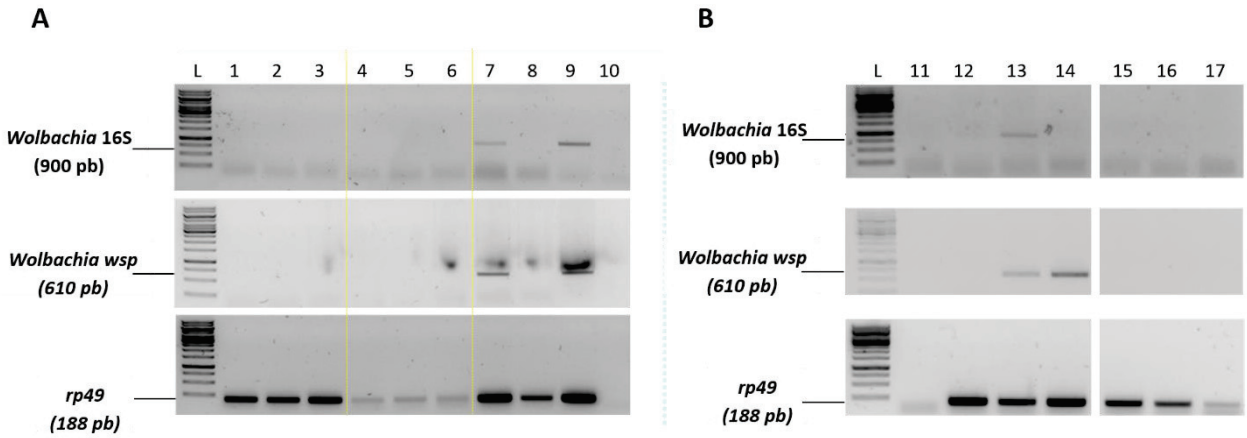


Figure S2. Absence de détection de *Wolbachia* après traitement à la tétracycline chez *D. melanogaster*. Les PCR ont été réalisées en utilisant des amorces spécifiques pour la détection des gènes 16s rRNA (*W. pipientis*) et pour la détection du gène *wsp* (codant pour une protéine de surface retrouvée chez un grand nombre d'espèces de *Wolbachia*). Témoin d'ADN par détection du gène de ménage *rp49*. **(A) *D. melanogaster w*^{1118} et *Dcr-2*$^{[L811fsX]}$.** *w*^{1118} (1-3), mutants *Dcr-2*$^{[L811fsX]}$ (4-6), témoin positif (7 et 9), témoin négatif (8), PCR- (eau) (10). **(B) *D. melanogaster CG4572*^{c05963}.** PCR- (eau) (11), témoin négatif (12), témoins positifs (13-14), *CG4572*$^{c05963\ -/-}$ (15-17). Résultats représentatifs de 2 expériences indépendantes incluant chacune 3 drosophiles analysées individuellement.

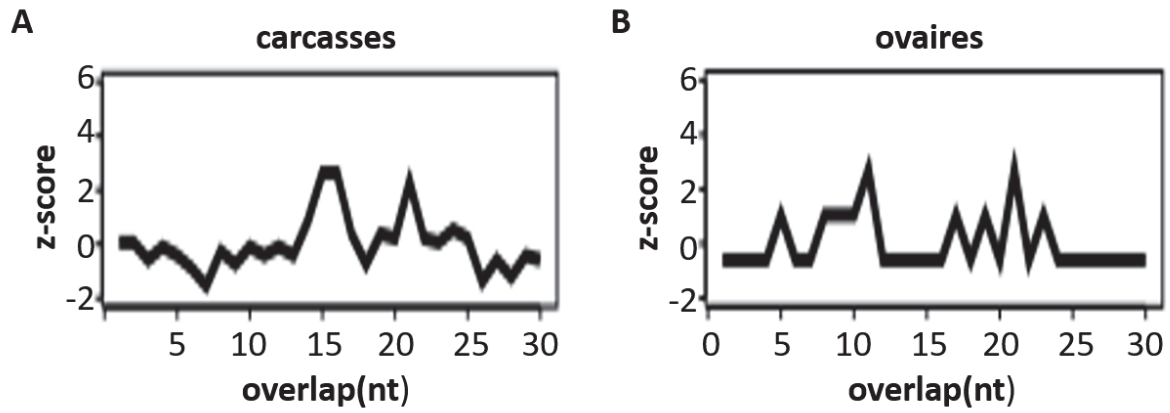


Figure S3. Absence de signature ping-pong des piARN ciblant SINV chez *D. melanogaster* w^{1118}. Analyse bioinformatique des piARN viraux dans les carcasses **(A)** et dans les ovaires **(B)** à 6 jours après infection par SINV. Le z-score permet de standardiser les mesures entre les différentes conditions ; il correspond à la transformation de la variable « nombre de paires de lectures alignées ». La signature « ping-pong » correspond au chevauchement (*overlap*) de 10 nucléotides (nt) entre le brin sens et antisens des piARN et est caractérisée par un pic à 10 nucléotides. Ce pic est significatif si le z-score est supérieur à 2. n = 2 expériences indépendantes.

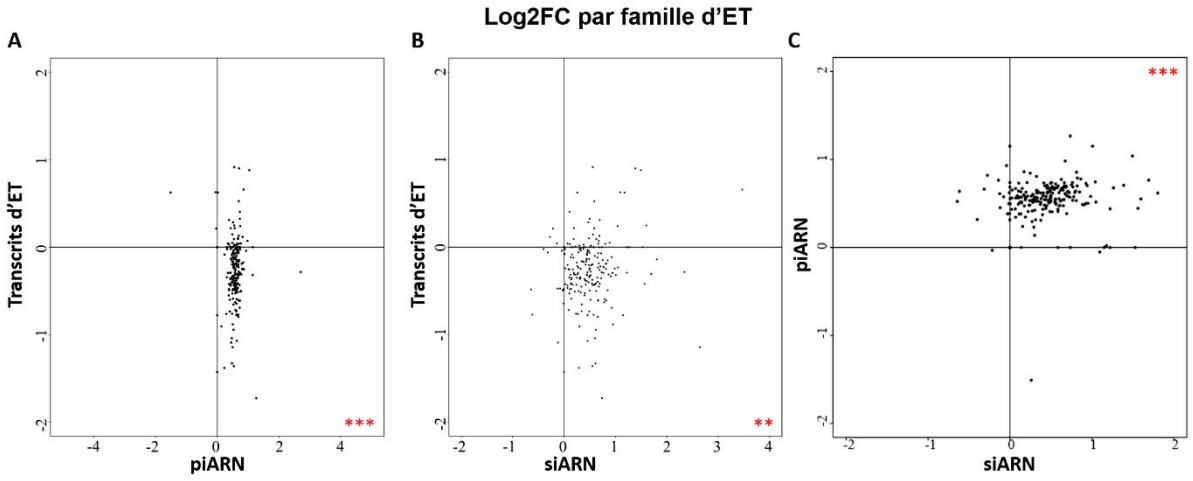


Figure S4. La diminution de l'abondance des transcrits d'ET est dépendante de l'augmentation des productions de piARN et siARN ciblant les ET, lors de l'infection par SINV, dans les carcasses w^{1118}. Chaque point représente le log2 du ratio des comptes par famille d'ET entre la condition infectée par SINV et la condition non infectée (mock). **(A)** Dépendance entre les transcrits d'ET et les piARN. **(B)** Dépendance entre les transcrits d'ET et les siARN. **(C)** Dépendance entre les piARN et les siARN. Le test exact de Fisher a été utilisé pour les analyses statistiques (*Fisher exact test ; *P < 0,05 ; **P < 0,01 ; ***P < 0,001*). n = 2 expériences indépendantes.

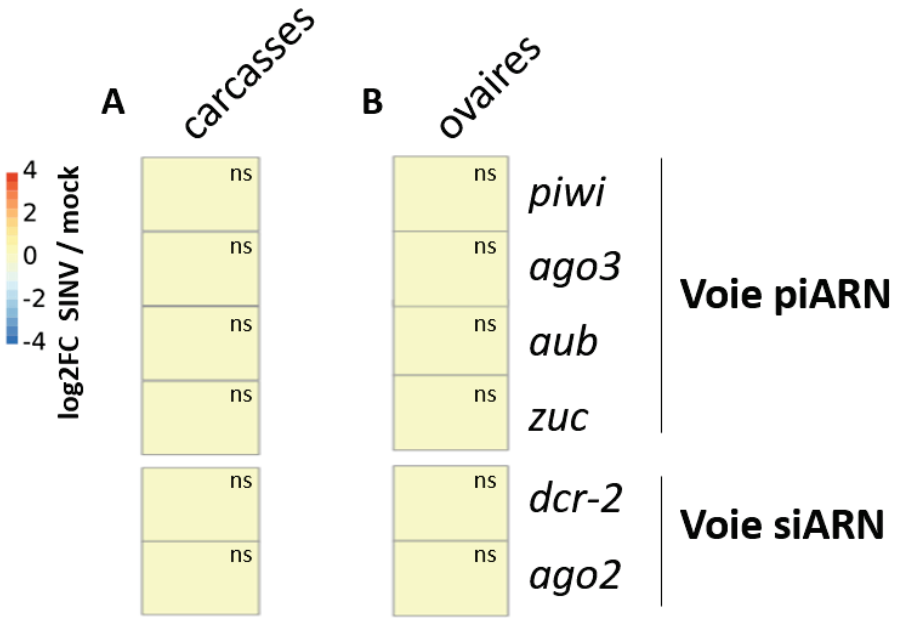


Figure S5. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN et siARN chez *D. melanoagster* w^{1118}, lors de l'infection par SINV. Les *heatmaps* représentent le log2 du ratio des comptes des gènes, entre la condition infectée par SINV et la condition non infectée (mock) dans les carcasses **(A)** et dans les ovaires **(B)**. Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock par DESeq2 (*test wald*). *ns : non significatif ; p-value* : **P<0,05 ; **P<0,01 ; ***P<0,001*. Légende : *piwi* (*P-element induced wimpy testis*), *ago3* (*argonaute 3*), *aub* (*aubergine*), *zuc* (*zucchini*), *dcr-2* (*dicer-2*), *ago2* (*argonaute 2*). n = 2 expériences indépendantes (30 tissus par échantillon de RNA-seq).

2. Infection par le virus Sindbis chez *D. simulans*

Infections persistantes chez *D. simulans*

Nous avons détecté par RT-PCR le virus Nora (NV) chez les drosophiles Makindu. Les infections chroniques virales, telles que celle induite par NV n'ont pas été éliminées pour les expériences de RNA-seq et Small RNA-seq des Makindu. Cependant, nous avons traité les drosophiles Makindu à l'eau de javel, selon le protocole décrit par Merkling et van Rij, (2015), pour effectuer des analyses complémentaires de RT-PCR quantitatives (**Figure S6.**).

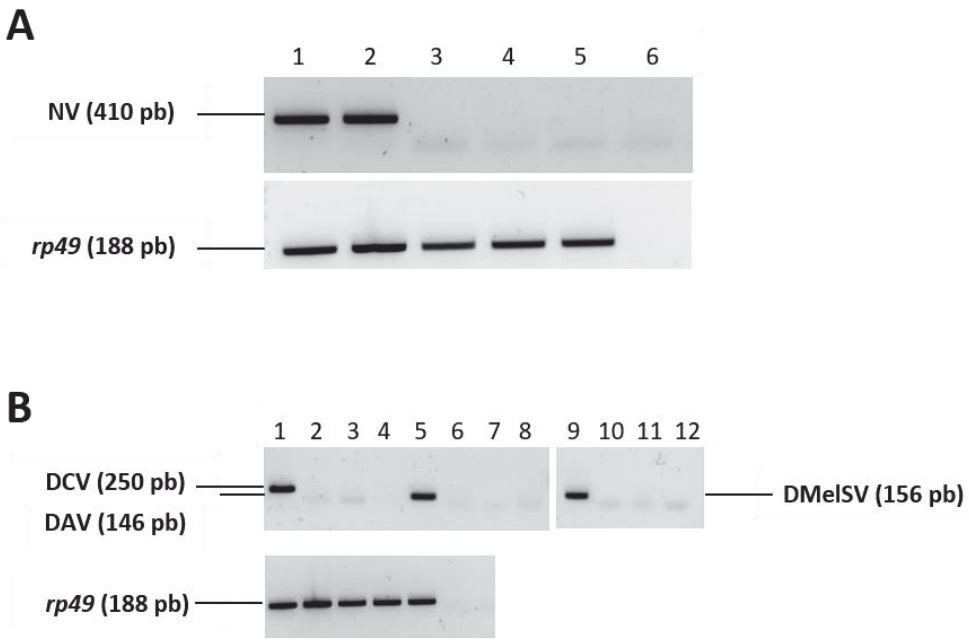


Figure S6. Détection de NV chez (*D. simulans*) Makindu. Détection des virus par RT-PCR avant et après javellisation des œufs de drosophiles, en utilisant des amorces spécifiques des virus suivant : NV, DCV, DAV, DMelSV. Témoin d'ARN par détection du gène de ménage *rp49*. **(A) Détection de NV / *rp49* (1-6)**, pour les échantillons d'ARN de drosophiles Makindu non traitées (1-2), Makindu après javellisation (3-5), et PCR- (eau) (6). **(B) Détection de DCV (1-4), DAV (5-8), DMelSV (9-12).** Témoins positifs DCV (1), DAV (5), DMelSV (9). Makindu non traitées (2 ; 6 ; 10), Makindu après javellisation (3 ;7 ; 11), PCR- (4 ; 8 ; 12). (Panel du bas) Détection du gène de ménage *rp49*, pour le T+ DCV (1), DAV (2), DmelSV (3), Makindu non traitées (4), Makindu après javellisation (5), PCR- (6). Résultats représentatifs de 2 expériences, incluant 10 drosophiles par échantillon.

D. simulans and *Wolbachia*

La bactérie endosymbiotique *Wolbachia* n'a pas été détectée chez *D. simulans* Makindu avant le traitement. Afin d'uniformiser les traitements, les drosophiles ont été traitées à la tétracycline, permettant l'élimination de toutes les bactéries endosymbiotiques (Merkling and van Rij, 2015). L'absence de Wolbachia a été validée par RT-PCR (**Figure S7.**).

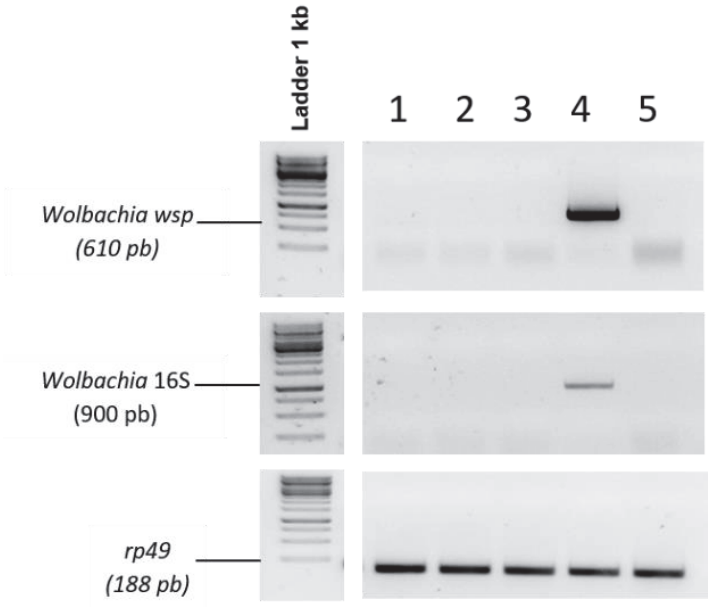


Figure S7. Absence de détection de *Wolbachia* après traitement à la tétracycline chez *D. simulans*. Les PCR ont été réalisées en utilisant des amorces spécifiques pour la détection des gènes 16s rRNA (*W. pipientis*) et pour la détection du gène *wsp* (codant une protéine de surface retrouvée chez un grand nombre d'espèces de *Wolbachia*). Témoin d'ADN par détection du gène de ménage *rp49*. Makindu (1-3), témoin positif (4), témoin négatif (5). Résultats représentatifs de 3 expériences indépendantes incluant chacune 3 drosophiles analysées individuellement.

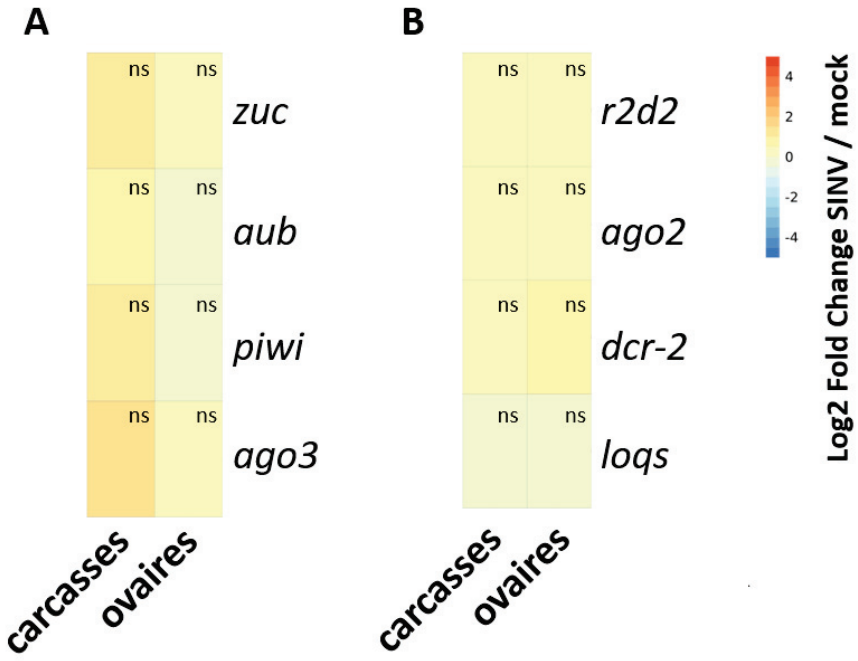


Figure S8. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN (A) et siARN (B) chez Makindu, lors de l'infection par SINV. Les *heatmaps* représentent le log2 du ratio des comptes des gènes, entre la condition infectée par SINV et la condition non infectée (mock). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la

condition mock par DESeq2 (*test wald*). *ns : non significatif ;* (**p-value* < 0,05). Légende : *zuc* (*zucchini*), *aub* (*aubergine*), *piwi* (*P-element induced wimpy testis*), *ago3* (*argonaute 3*), *r2d2* (*r2d2*), *ago2* (*argonaute 2*), *dcr-2* (*dicer-2*), *loqs* (*loquacious*). n = 2 expériences indépendantes (30 tissus par échantillon de RNA-seq).

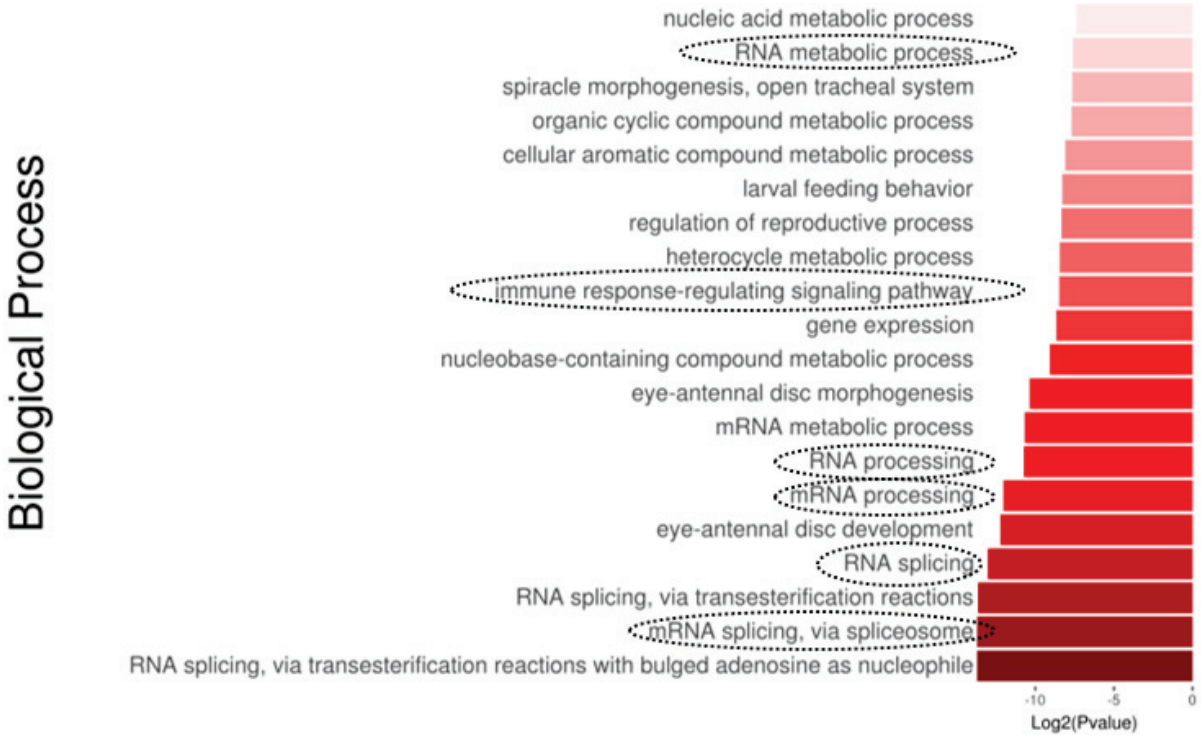


Figure S9. Parmi les gènes différentiellement exprimés lors de l'infection par SINV, on trouve un enrichissement en gènes associés au métabolisme de l'ARN. L'analyse de Gene Ontology (GO) a été réalisée pour déterminer s'il existe un enrichissement en certaines fonctions parmi les gènes différentiellement exprimés lors de l'infection par SINV par rapport à la condition non infectée. Les résulats sont exprimés en log2 de la *p-value* pour l'ensemble des gènes classés dans un processus biologique donné.

3. Mécanismes impliqués dans la dérégulation des ET

Validation du génotype chez *D. melanogaster CG4572^{c05963}*

Le stock de mutants *D. melanogaster CG4572^{c05963}* est maintenu hétérozygote, assurant la fertilité des drosophiles. Pour cela, le chromosome balanceur TM6B est présent dans le génome de certaines drosophiles. Nous avons sélectionné les mutants homozygotes pour réaliser la suite des expérimentations (**Figure S10.**).

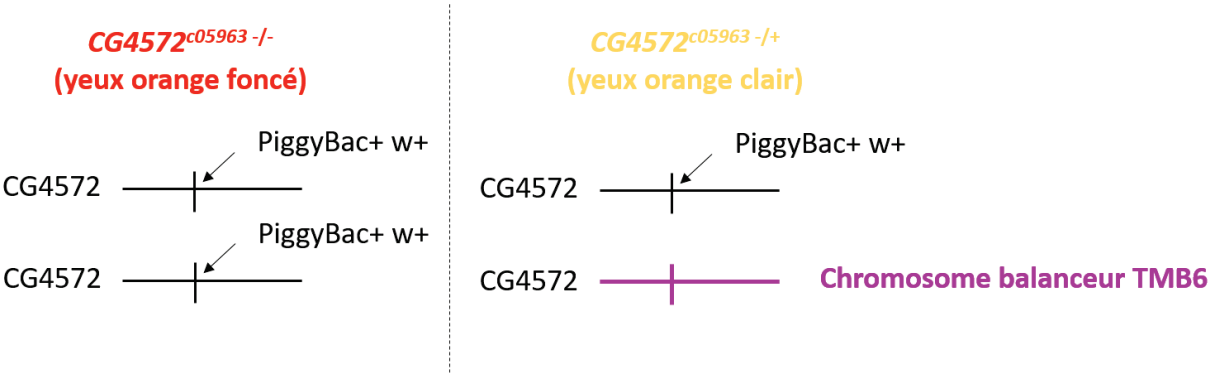


Figure S10. Validation du génotype *CG4572*$^{c05963\ -/-}$ par la sélection du phénotype associé. Les mutants homozygotes pour l'insertion du transposon *PiggyBac* contenant le marqueur *white* ont été sélectionnés en fonction de la couleur des yeux des drosophiles et de l'absence de soie au niveau des épaules du thorax des drosophiles.

Infection de *D. melanogaster CG4572*^{c05963} par SINV

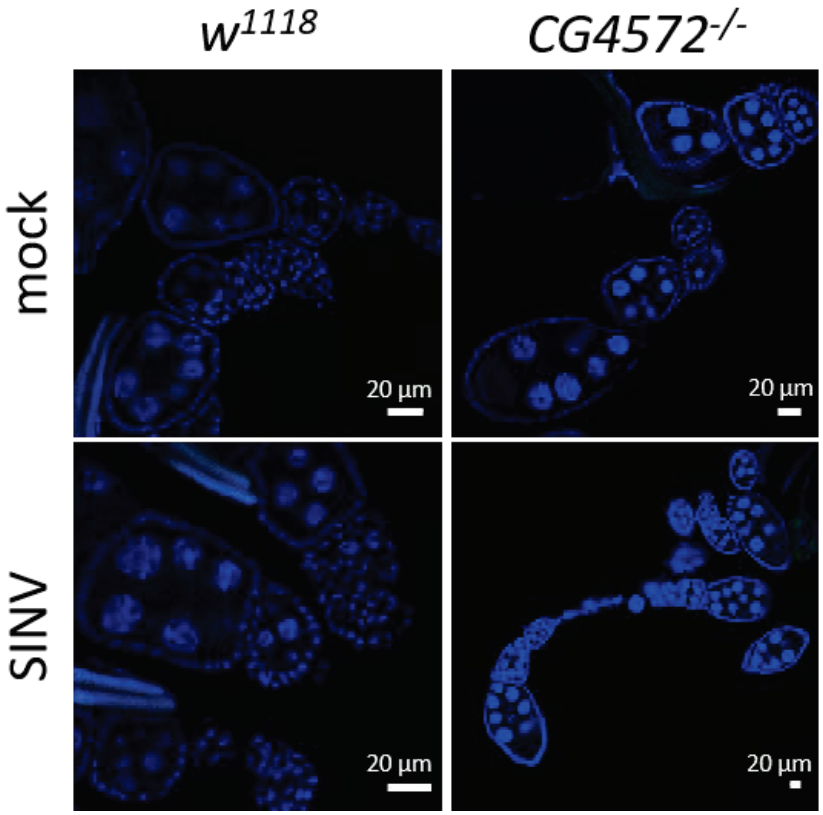


Figure S11. Absence de détection de SINV dans les ovaires des mutants *D. melanogaster CG4572*^{c05963}. Les drosophiles ont été infectées avec une dose de 2300 PFU de SINV dans un volume d'injection de 32,2 nl puis les paires d'ovaires ont été disséquées au jour 6 après infection. La technique d'immunofluorescence a été réalisée

en utilisant un anticorps anti-capside de SINV et un anticorps secondaire fluorescent (A488). L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 10 paires d'ovaires.

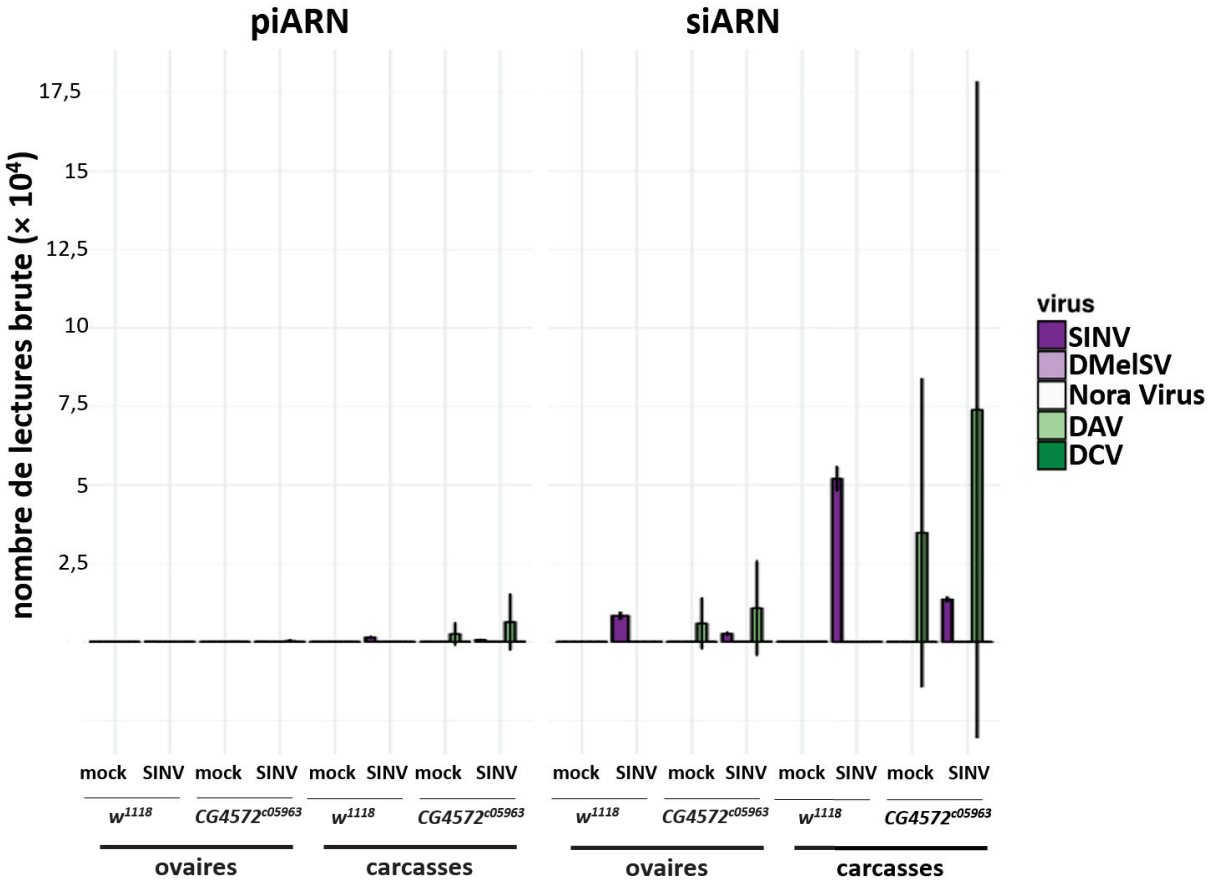


Figure S12. Production de siARN ciblant SINV et DAV chez *D. melanogaster CG4572^{c05963}*. Nombre de lectures brute des brins des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles sauvages *w^{1118}* et des mutants *CG4572^{c05963}*. n = 2 expériences indépendantes (50 tissus par échantillon).

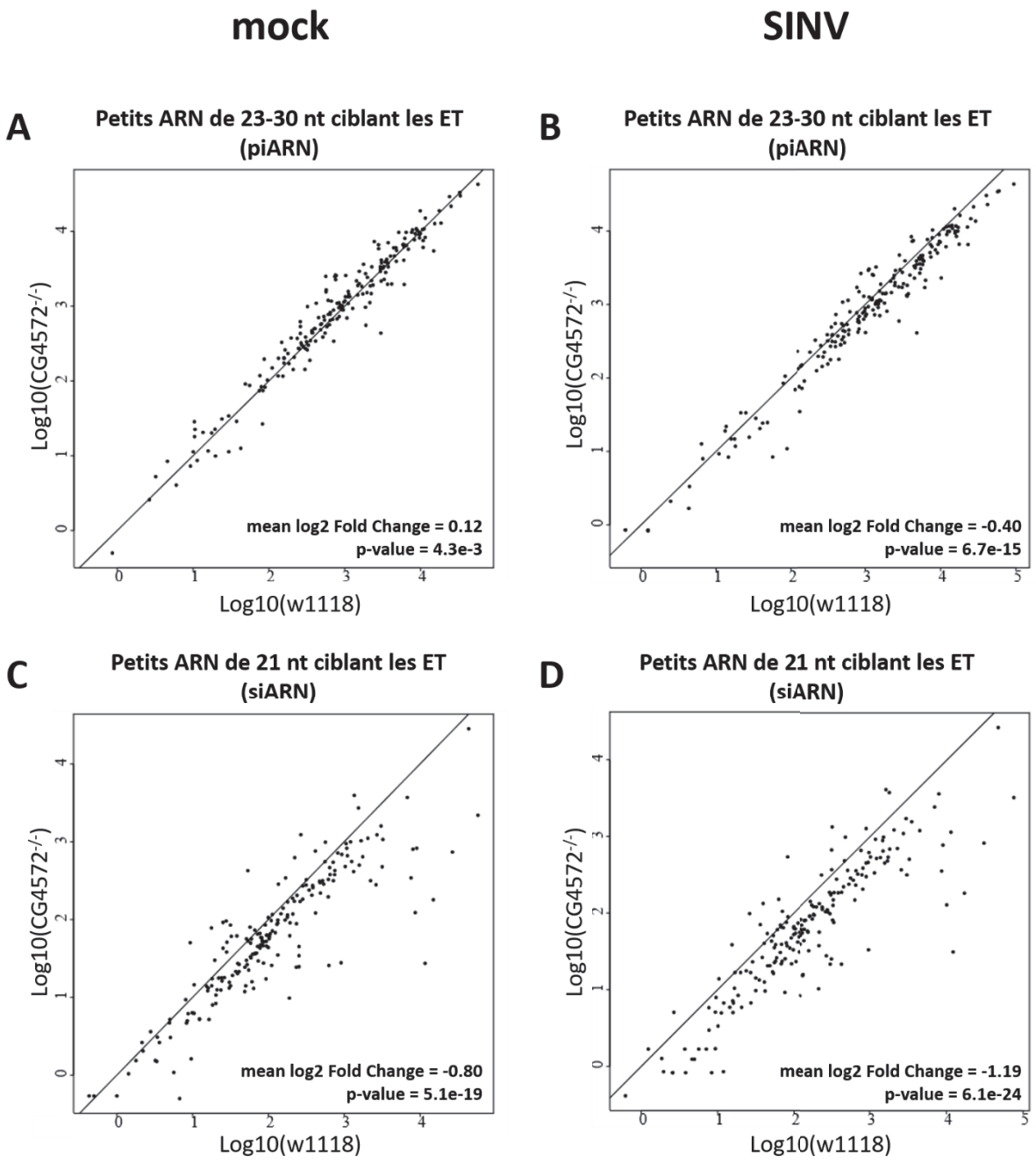


Figure S13. Comparaison des quantités de petits ARN interférents ciblant les ET entre les *w*^{1118} et les mutants *CG4572*^{c05963}, dans les carcasses. Chaque point représente une famille d'ET. La droite est y = x. La comparaison entre les mutants *CG4572*^{c05963} (CG4572$^{-/-}$) et les drosophiles sauvages *w*^{1118} a été analysée pour la production des piARN (23-30 nucléotides (nt) de long) (A-B) et des siARN (21 nucléotides de long) (C-D) en condition non infectée (mock) (A-C) et lors de l'infection par SINV (B-D). **(A) piARN (mock)**. Les comptes des mutants *CG4572*^{c05963} sont significativement supérieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = 0,12 ; *p-value* = 4,3e-3). **(B) piARN (SINV)**. Les comptes des mutants *CG4572*^{c05963} sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = - 0,40 ; *p-value* = 6,7e-15). **(C) siARN (mock)**. Les comptes des mutants *CG4572*^{c05963} sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,80 ; *p-value* = 5,1e-19). **(D) siARN (SINV)**. Les comptes des mutants *CG4572*^{c05963} sont significativement inférieurs à ceux des

w^{1118} (moyenne du log2 *fold change* = -1,19 ; *p-value* = 6,1e-24). Test statistique utilisé : *Wilcoxon paired test* ; n = 2 expériences indépendantes.

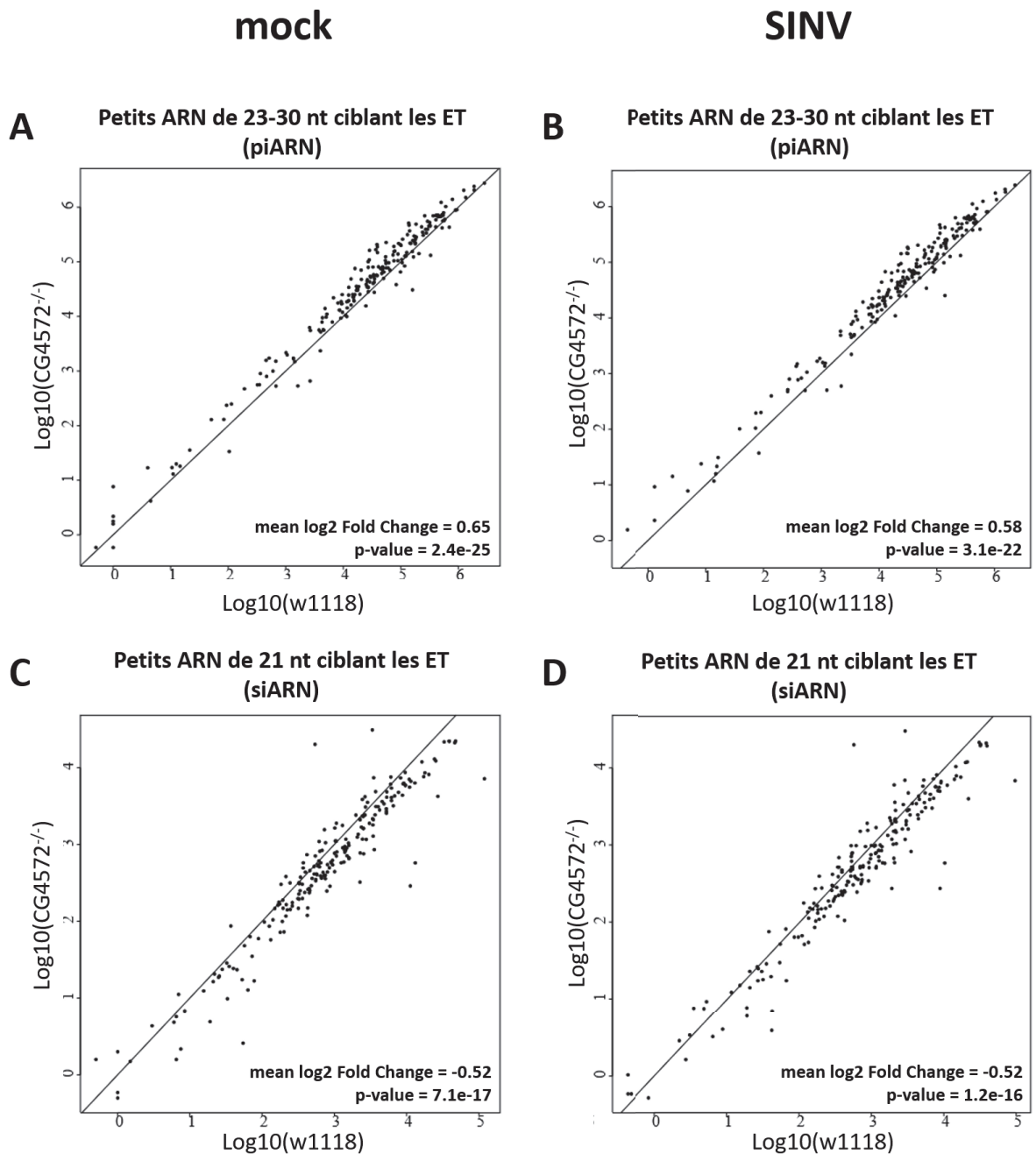


Figure S14. Comparaison des quantités de petits ARN interférents ciblant les ET entre les *w*^{1118} et les mutants *CG4572*^{c05963}, dans les ovaires. Chaque point représente une famille d'ET. La droite est y = x. La comparaison entre les mutants *CG4572*^{c05963} (CG4572$^{-/-}$) et les drosophiles sauvages *w*^{1118} a été analysée pour la production des piARN (23-30 nucléotides (nt) de long) (A-B) et des siARN (21 nucléotides de long) (C-D) en condition non infectée (mock) (A-C) et lors de l'infection par SINV (B-D). **(A) piARN (mock)**. Les comptes des mutants *CG4572*^{c05963} sont significativement supérieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = 0,65 ; *p-value* = 2,4e-25). **(B) piARN (SINV)**. Les comptes des mutants *CG4572*^{c05963} sont significativement supérieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = 0,58 ; *p-value* = 3,1e-22). **(C) siARN (mock)**. Les comptes des mutants *CG4572*^{c05963} sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,52 ; *p-value* =

7,1e-17). **(D) siARN (SINV)**. Les comptes des mutants *CG4572*^{c05963} sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,52 ; *p-value* = 1,2e-16). Test statistique utilisé : *Wilcoxon paired test* ; n = 2 expériences indépendantes.

Validation du génotype chez *D. melanogaster Dcr-2*$^{[L811fsX]}$

Nous avons utilisé les mutants *D. melanogaster Dcr-2*$^{[L811fsX]}$ présentant le génotype suivant : *y[d2] w[1118], P{ry+t7.2 = ey–FLP.N}2 ; Dcr-2[L811fX] (Tomari et al., 2007)*. Ces mutants homozygotes *Dcr-2*$^{[L811fsX]}$ présentent une mutation au niveau de la Leucine 811. Suite à un décalage du cadre de lecture, la Leucine 811 est remplacée par 19 nouveaux acides aminés ce qui introduit prématurément un codon stop au niveau du gène codant cette protéine. Par conséquent, Dcr-2 n'est plus fonctionnelle dans ces mutants. Ces derniers ne sont plus capables de produire des siARN double brin et ainsi de détecter spécifiquement l'ARNm à cliver. En effet, Dcr-2 est également un composant du complexe RISC (RNA-induced silencing complex) ce qui l'implique dans le clivage de l'ARNm par les siARN (*Young Sik Lee et al*., 2004, *Xiang. Liu et al., 2006*, *J.T. Marques et al*., 2010, *Do-HwanLim et al*., 2011). A partir de ces mutants *Dcr-2*$^{[L811fsX]}$, nous avons séquencé l'ADN et cela nous a permis de valider leur génotype.

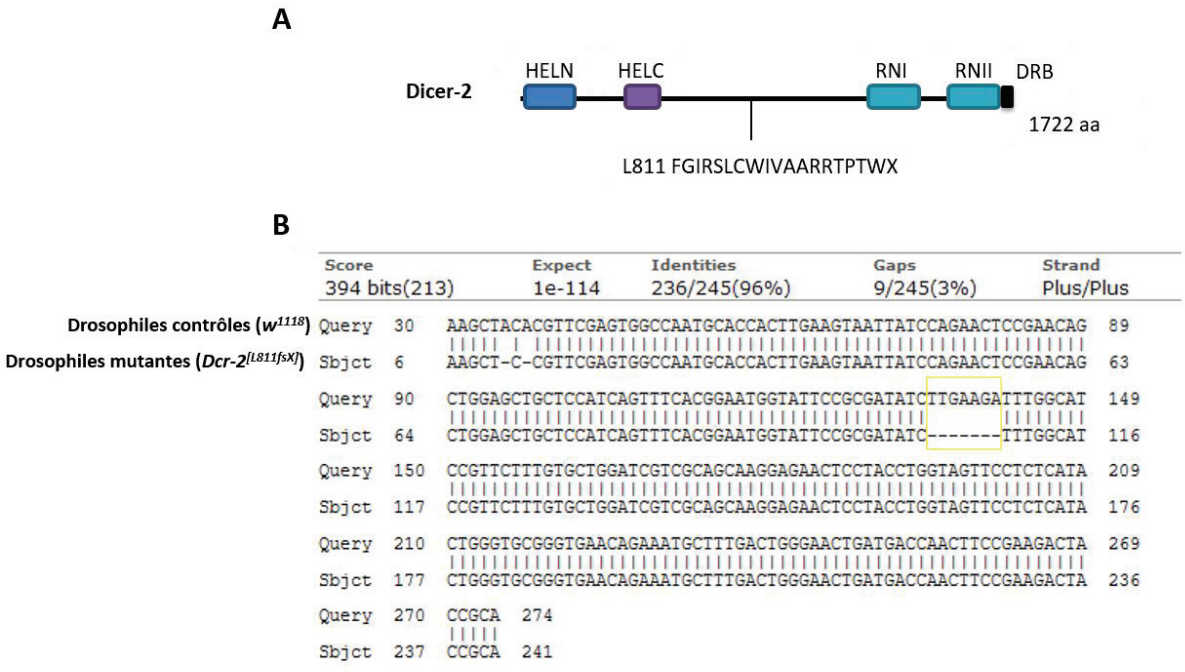


Figure S15. Génotype des mutants *Dcr-2*$^{[L811fsX]}$. (A) Schéma de la localisation de la mutation L811 sur la protéine Dicer-2. La protéine Dcr-2 contient les domaines DEAD/DEAH box-hélicase (HELN) et C-hélicase (HELC), les domaines RNase III (RNI et RNII) répétés en tandem et le domaine C-terminal de liaison à l'ARN double brin (DRB). D'après (Lee et al., 2004; Poirier et al., 2018). **(B)** Validation du génotype des mutants *Dcr-2*$^{[L811fsX]}$. Délétion des 7 nucléotides (en jaune) dont les nucléotides TTG codant la Leucine 811 chez les mutants *Dcr-2*$^{[L811fsX]}$ par rapport

aux drosophiles contrôles w^{1118}. Décalage du cadre de lecture chez les mutants *Dcr-2*$^{[L811fsX]}$ permettant la traduction de 19 nouveaux acides aminés : FGIRSLCWIVAARRTPTWX.

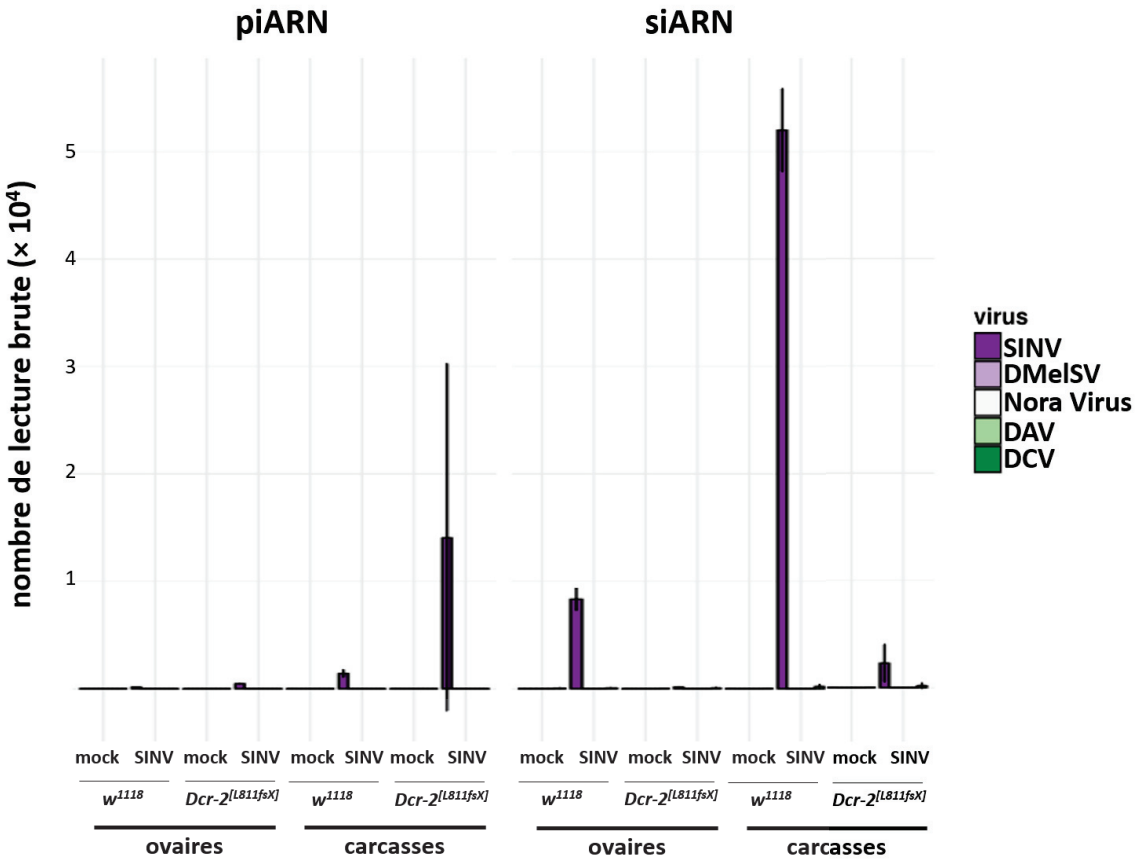


Figure S16. Absence de production de petits ARN interférents ciblant les virus naturels persistants chez *D. melanogaster Dcr-2*$^{[L811fsX]}$. Nombre de lectures brute des brins des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles sauvages w^{1118} et des mutants *Dcr-2*$^{[L811fsX]}$. n = 2 expériences indépendantes (50 tissus par échantillon).

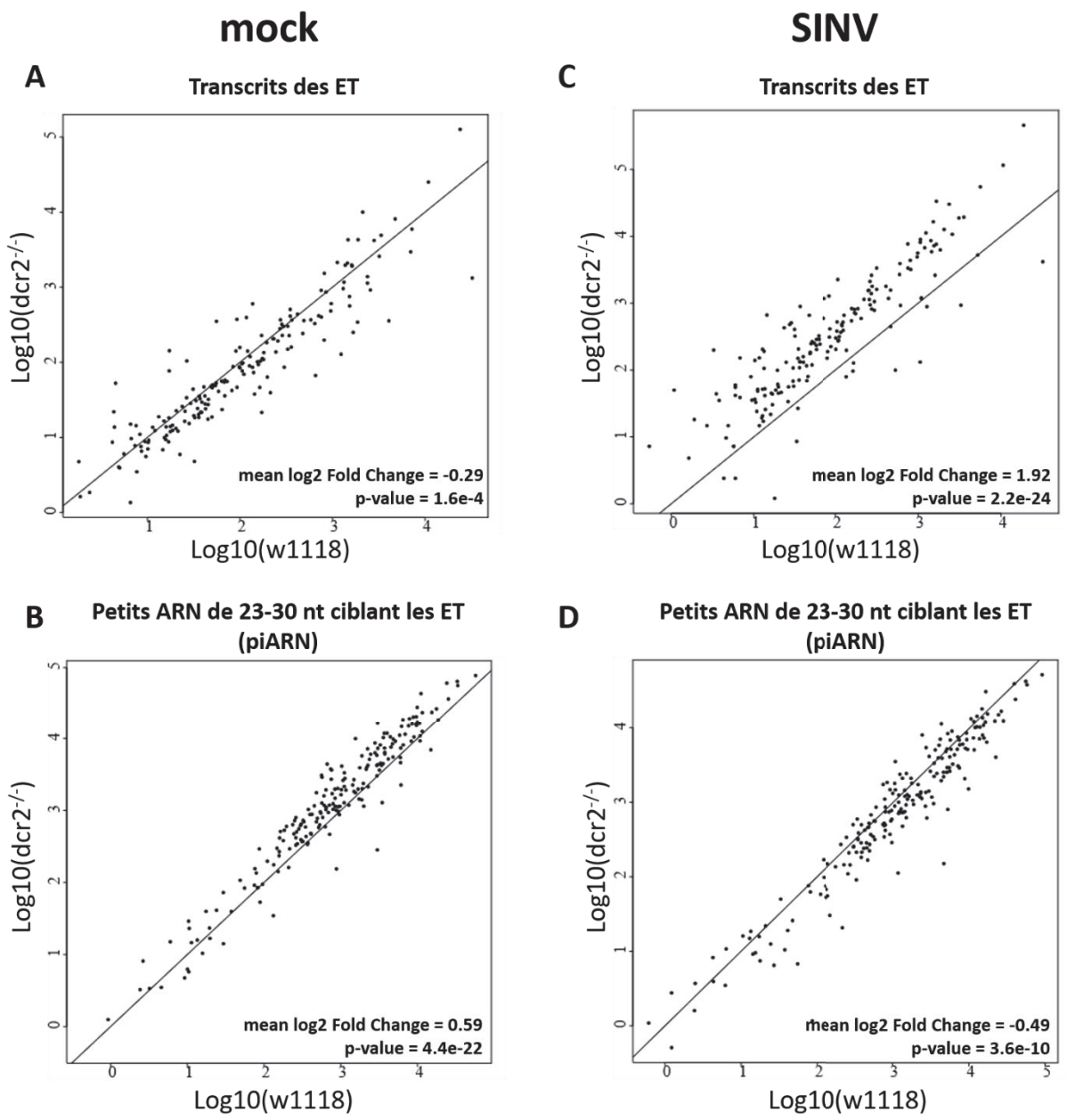


Figure S17. Comparaison des quantités de transcrits des ET et des piARN ciblant les ET entre les *w*^{1118} et les mutants *Dcr-2*$^{[L811fsX]}$, dans les carcasses. Chaque point représente une famille d'ET. La droite est y = x. La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ (dcr2$^{-/-}$) et les drosophiles sauvages *w*^{1118} a été analysée pour les quantités de transcrits d'ET (A-B) et la production des piARN (23-30 nucléotides (nt) de long) (C-D) en condition non infectée (mock) (A-B) et lors de l'infection par SINV (C-D). **(A) Transcrits des ET (mock)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* (DESeq2) = -0,29 ; *p-value* = 1,6e-4). **(B) piARN (mock)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ sont significativement supérieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* (DESeq2) = 1,92 ; *p-value* = 2,22e-24). **(C) Transcrits des ET (SINV)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ sont significativement supérieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = 0,59 ; *p-value* = 4,4e-22). **(D) piARN (SINV)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,49 ; *p-value* = 3,6e-10). Test statistique utilisé : *Wilcoxon paired test* ; n = 2 expériences indépendantes.

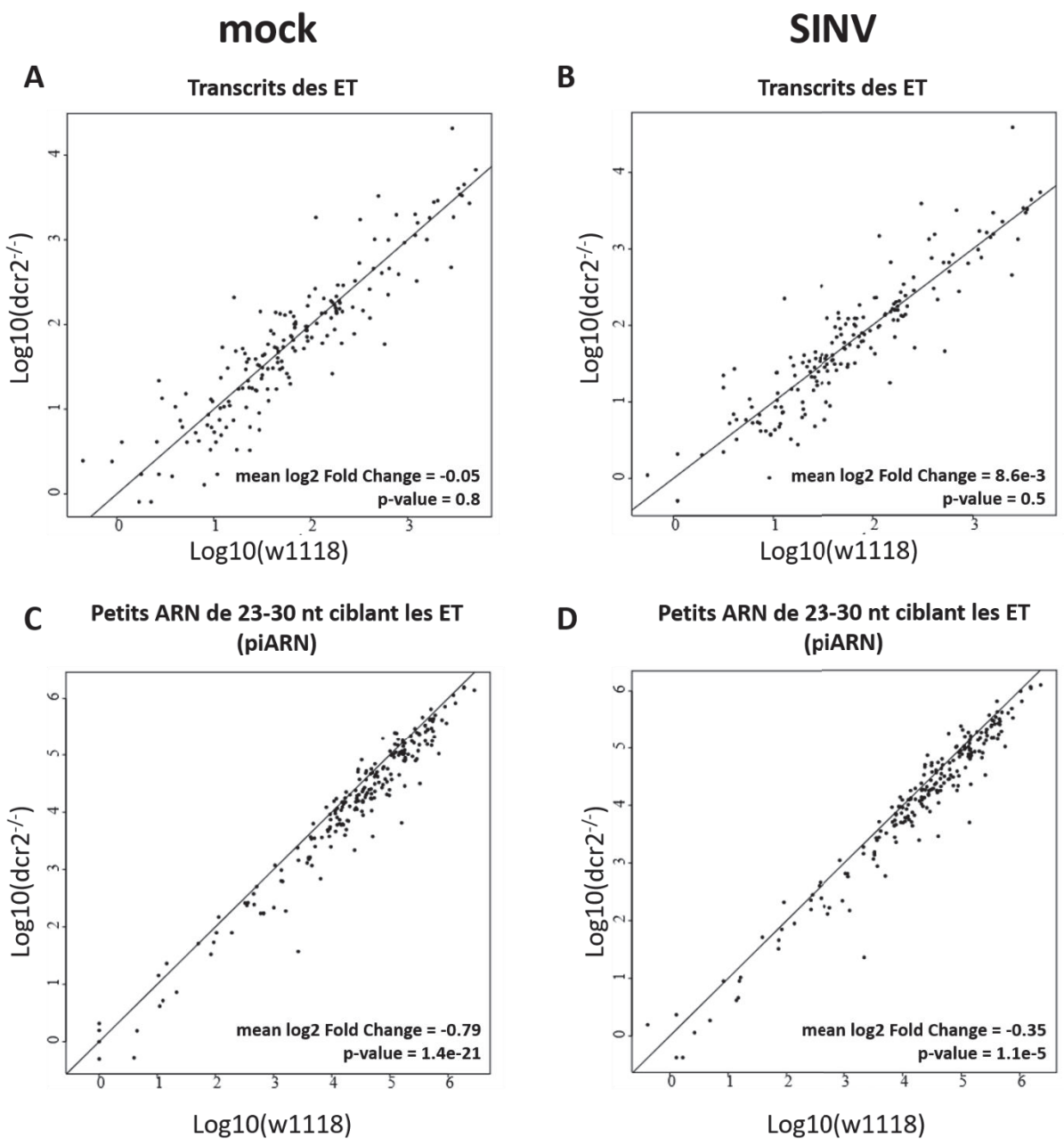


Figure S18. Comparaison des quantités de transcrits des ET et des piARN ciblant les ET entre les *w*^{1118} et les mutants *Dcr-2*$^{[L811fsX]}$, dans les ovaires. Chaque point représente une famille d'ET. La droite est y = x. La comparaison entre les mutants *Dcr-2*$^{[L811fsX]}$ (dcr2$^{-/-}$) et les drosophiles sauvages *w*^{1118} a été analysée pour les quantités de transcrits d'ET (A-B) et la production des piARN (23-30 nucléotides (nt) de long) (C-D) en condition non infectée (mock) (A-C) et lors de l'infection par SINV (B-D). **(A) Transcrits des ET (mock)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ ne sont pas significativement différents de ceux des *w*^{1118} (moyenne du log2 *fold change* (DESeq2) = -0,05 ; *p-value* = 0,8). **(B) Transcrits des ET (SINV)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ ne sont pas significativement différents de ceux des *w*^{1118} (moyenne du log2 *fold change* (DESeq2) = 8,6e-3 ; *p-value* = 0,5). **(C) piARN (mock)**. Les comptes des mutants *Dcr-2*$^{[L811fsX]}$ sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,79 ; *p-value* = 1,4e-21). **(D) piARN (SINV)**. Les comptes des mutants *Dcr- 2*$^{[L811fsX]}$ sont significativement inférieurs à ceux des *w*^{1118} (moyenne du log2 *fold change* = -0,35 ; *p-value* = 1,1e-5). Test statistique utilisé : *Wilcoxon paired test* ; n = 2 expériences indépendantes.

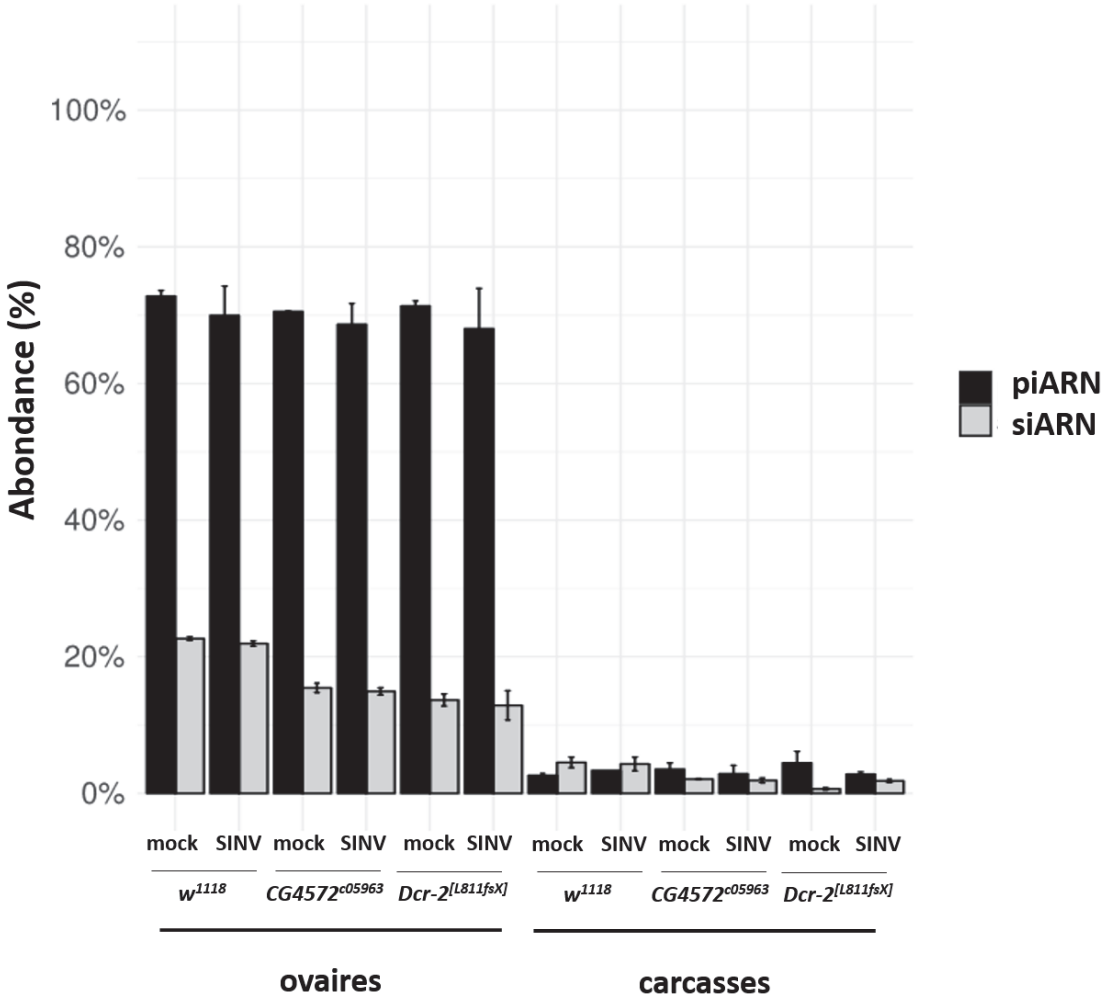


Figure S19. Pourcentage des petits ARN interférents ciblant les ET chez *D. melanogaster*. Abondance des petits ARN interférents (ciblant les ET) de 23-30 nucléotides (piARN)/piARN totaux et de 21 nulcéotides (siARN)/siARN totaux, ciblant les ET, dans les carcasses et les ovaires des drosophiles sauvages w^{1118} et des mutants $CG4572^{c05963}$ et $Dcr\text{-}2^{[L811fsX]}$. n = 2 expériences indépendantes.

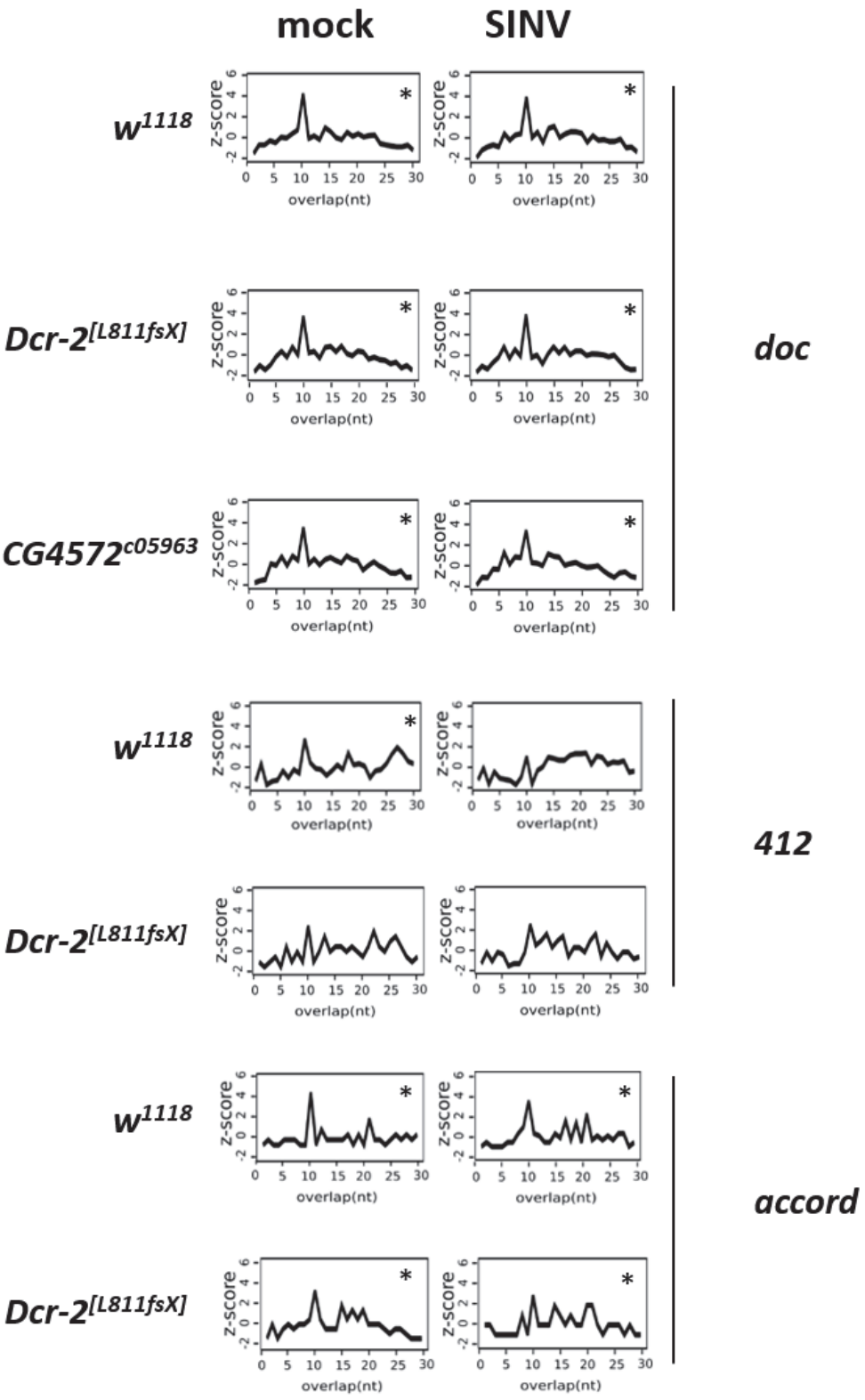


Figure S20. Signature ping-pong des piARN ciblant les ET dans les carcasses de *D.* melanogaster. Le z-score permet de standardiser les mesures entre les différentes conditions ; il correspond à la transformation de la variable « nombre de paires de lectures alignées ». Le chevauchement (*overlap*) de 10 nucléotides (nt) entre le brin sens et antisens des piARN est visible au niveau du pic à 10 nt. Ce pic est significatif si le z-score est supérieur à 2, représenté par une étoile. n = 2 expériences indépendantes.

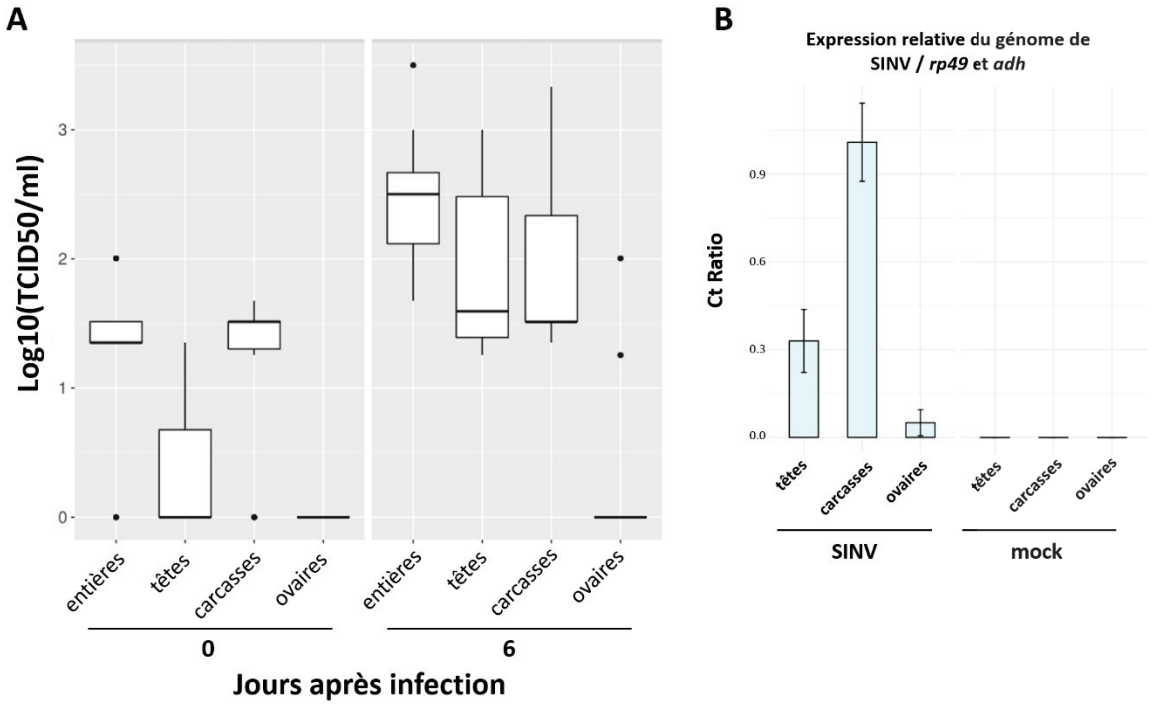


Figure S21. Réplication de SINV dans les têtes de drosophiles chez Makindu. Les drosophiles ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl. **(A)** Les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de carcasses (mouche entière sans la tête et les ovaires), de paire d'ovaires individuelle, et de groupe de 5 têtes issues de drosophiles infectées (J6 après infection). n = 3 expériences indépendantes (10 drosophiles par condition / 6 pools de 5 têtes). **(B)** L'expression relative du génome de SINV par rapport aux gènes de ménage *rp49* et *adh* a été déterminée par RT-qPCR pour les échantillons infectés par SINV et les échantillons non infectés (mock). n = 2 expériences indépendantes (30 carcasses / paires d'ovaires par échantillon d'ARN).

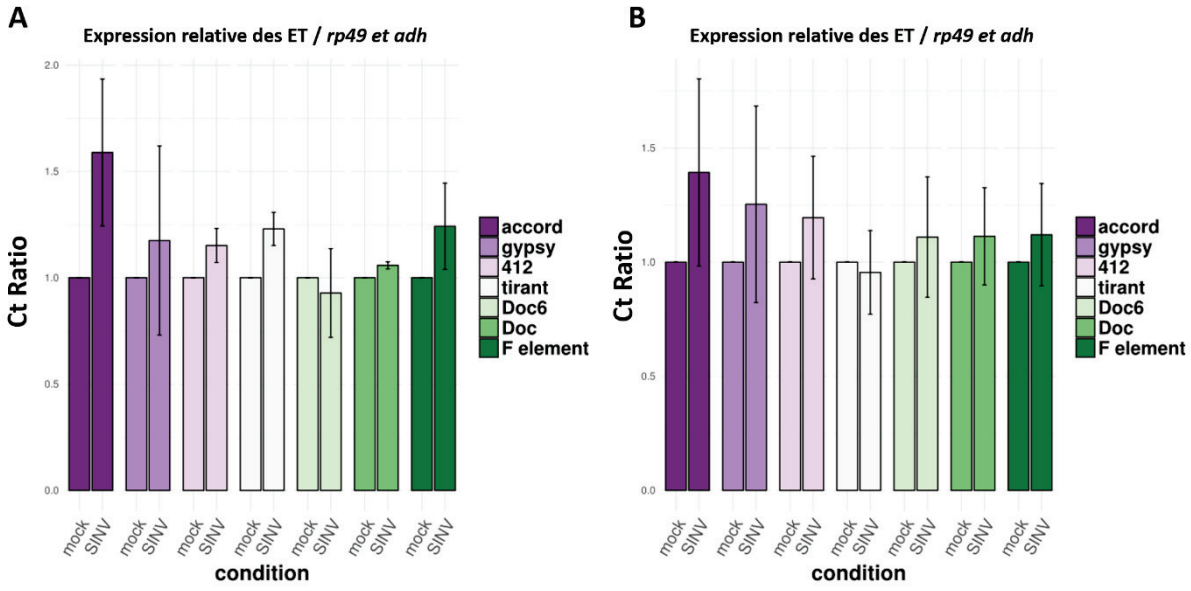


Figure S22. Absence de modulation significative de l'expression des ET lors de l'infection par SINV dans les têtes et le reste des carcasses des Makindu. Les drosophiles ont été infectées avec une dose de 14900 PFU de

SINV dans un volume d'injection de 207 nl. Les niveaux d'expression des familles des ET *accord, gypsy, 412, tirant, Doc6, Doc* et *F element* ont été déterminés par RT-PCR quantitative dans les têtes **(A)** et dans les carcasses (corps de la drosophile sans la tête et les ovaires) **(B)** et sont représentés en expression relative par rapport aux gènes de ménage *rp49* et *adh*. n = 3 expériences indépendantes (30 carcasses / têtes par échantillon d'ARN). Test de Wilcoxon-Mann-Whitney, non significatif pour l'ensemble des familles d'ET analysées ; (A) *p-value* des ET dans les têtes : *accord* = 0,4 ; *gypsy* = 0,2 ; *412* = 0,5 ; *tirant* = 0,06 ; *Doc6* = 0,7 ; *Doc* = 0,7 ; *F element* = 0,8. (B) *p-value* des ET en carcasses : *accord* = 0,2 ; *gypsy* = 0,7 ; *412* = 0,7 ; *tirant* = 1 ; *Doc6* = 0,7 ; *Doc* = 1 ; *F element* = 0,8.

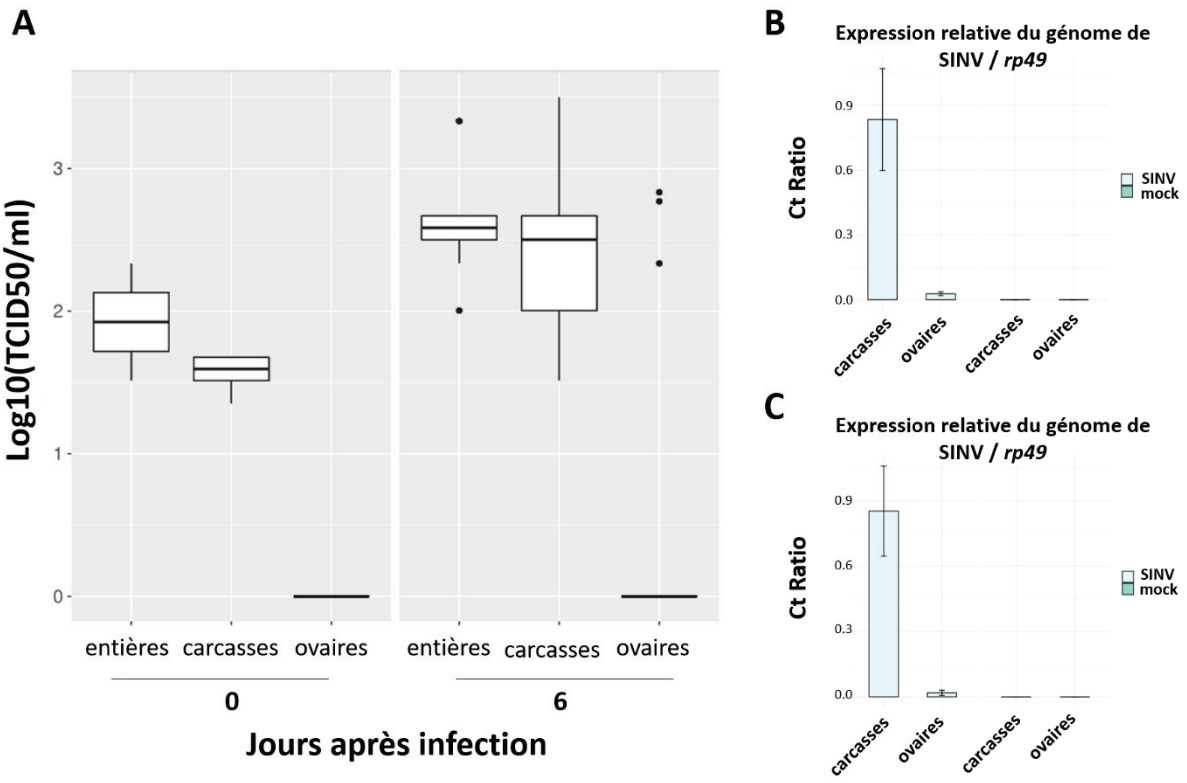


Figure S23. Réplication de SINV chez Makindu après javellisation des œufs. Les drosophiles ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl. **(A)** A partir des drosophiles privées d'infections virales chroniques, les titres viraux de SINV ont été déterminés en utilisant la technique de quantification de particules virales en TCID50/ml à partir des broyats individuels de drosophiles entières, de carcasses et de paires d'ovaires issues de drosophiles infectées, au jour de l'infection et à 6 jours après infection. n = 3 expériences indépendantes (10 drosophiles par condition). L'expression relative du génome de SINV par rapport au gène de ménage *rp49* a été déterminée par RT-qPCR pour les drosophiles Makindu issues d'œufs non javellisés **(B)** et pour celles issues d'oeufs javellisés **(C)**. n = 3 expériences indépendantes (30 carcasses / paires d'ovaires par échantillon d'ARN).

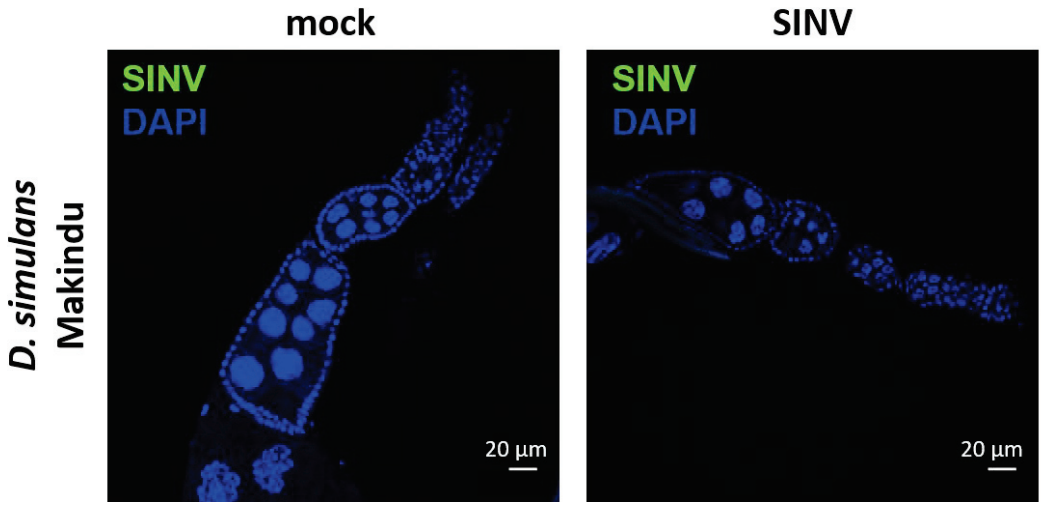


Figure S24. Absence de détection de SINV dans les ovaires chez Makindu après javellisation des œufs. Les drosophiles ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl. Les paires d'ovaires ont été disséquées au jour 6 après infection. La technique d'immunofluorescence a été utilisée en utilisant un anticorps anti-capside de SINV et un conjugué fluorescent (A488). L'ADN a été marqué en bleu (DAPI). Les résultats sont représentatifs de 2 expériences indépendantes incluant chacune 10 paires d'ovaires.

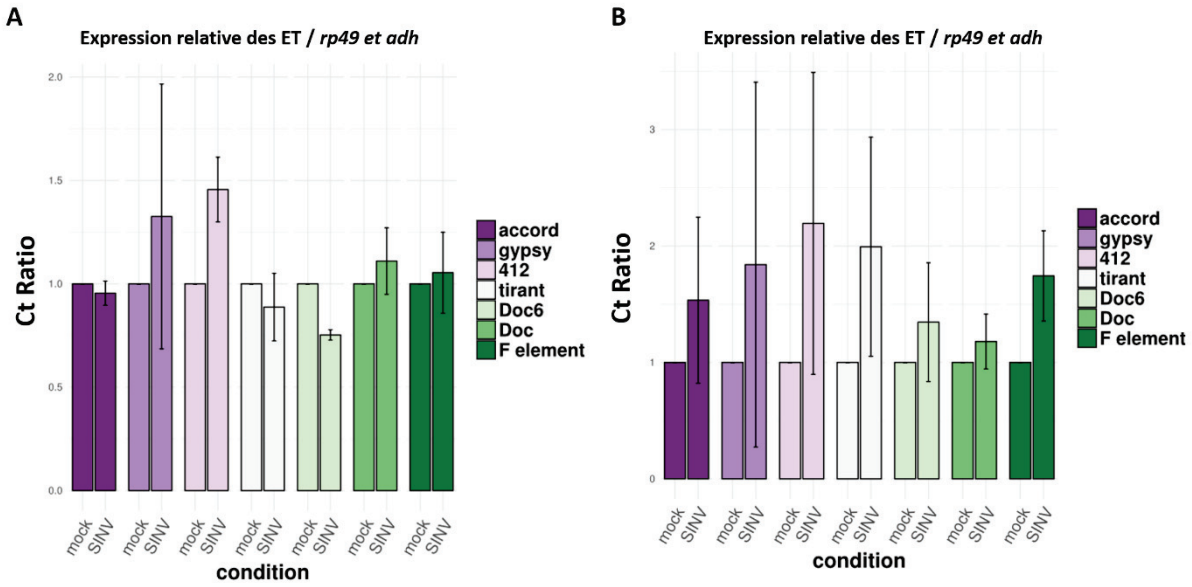


Figure S25. Absence de modulation significative de l'expression des ET chez Makindu, après élimination des infections persistantes, lors de l'infection par SINV. Après élimination des infections persistantes, les drosophiles ont été infectées avec une dose de 14900 PFU de SINV dans un volume d'injection de 207 nl. Les niveaux d'expression des familles des ET *accord, gypsy, 412, tirant, Doc6, Doc* et *F element* ont été déterminés par RT-PCR quantitative dans les carcasses **(A)** et dans les ovaires **(B)** et sont représentés en expression relative par rapport aux gènes de ménage *rp49* et *adh*. n = 3 expériences indépendantes (12 carcasses / paires d'ovaires par échantillon d'ARN). Test de *Wilcoxon-Mann-Whitney*, non significatif pour l'ensemble des familles d'ET analysé ; (A) *p-value* des ET dans les carcasses : *accord* = 0,1 ; *gypsy* = 0,7 ; *412* = 0,1 ; *tirant* = 0,1 ; *Doc6* = 0,2 ; *Doc* = 0,4 ; *F element* = 0,1. (B) *p-value* des ET en ovaire : *accord* = 1 ; *gypsy* = 0,7 ; *412* = 0,1 ; *tirant* = 0,7 ; *Doc6* = 0,1 ; *Doc* = 0,7 ; *F element* = 0,1.

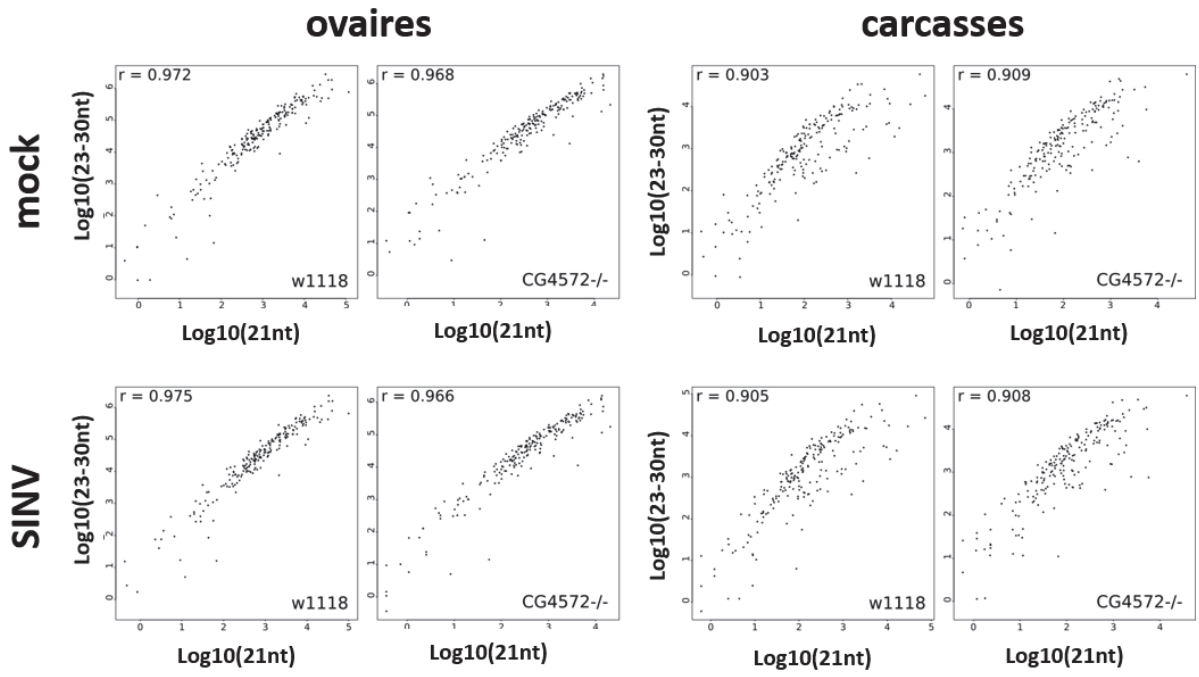


Figure S26. Corrélation positive entre les petits ARN interférents de 23-30 nucléotides et ceux de 21 nucléotides dans toutes les conditions. Les coefficients de corrélation de *Pearson* sont indiqués en haut à gauche de chaque panel (r).

RESULTATS EXPERIMENTAUX (II)

II. Réactivation d'éléments transposables lors de l'infection par les virus Sigma et DCV chez *Drosophila simulans*

(Article en préparation)

1. Introduction

Le virus sigma est un virus enveloppé, contenant un génome à ARN simple brin de polarité négative. Il appartient à l'ordre des *Mononegavirales* et à la famille des *Rhabdoviridae*. Il a été découvert chez *D. melanogaster* d'où son nom *D. melanogaster Sigma Virus* (DMelSV). Contrairement à *D. simulans* qui n'est jamais infectée naturellement par DMelSV (Webster et al., 2015), d'autres espèces de drosophiles sont fréquemment infectées par le virus sigma, telles que *D. affinis* et *D. athabasca*, *D. obscura*, *D. tristis*, *D. immigrans* et *D. ananassae* (Longdon et al., 2012). Ainsi, nous avons infecté expérimentalement cette espèce de drosophile, permettant d'initier les voies d'ARNi, incluant les voies siARN et piARN dirigées contre DMelSV lors de l'infection expérimentale. Une caractéristique très intéressante de DMelSV est sa transmission verticale par l'intermédiaire du sperme et des œufs (Brun and Plus, 1980; Fleuriet, 1988; L'heritier, 1958; Longdon et al., 2010) alors que la transmission horizontale est plutôt rare ou inexistante. Par conséquent, ce modèle viral permet l'analyse de l'impact de l'infection sur les quantités des transcrits d'ET dans les carcasses des drosophiles mais aussi au niveau des ovaires infectés par le virus. D'autant plus que le contrôle de l'expression des ET par la voie des piARN est principalement mis en place au niveau des ovaires (Iwasaki et al., 2015).

Des analyses bio-informatiques ont également été menées à partir de données de séquençage à haut débit du transcriptome de drosophiles infectées par un autre virus spécifique de drosophile, *Drosophila C Virus* (DCV) (Merkling et al., 2015). Ces données ont été générées par un autre laboratoire concernant une autre question biologique (Merkling et al., 2015). DCV appartient à la famille des *Dicistroviridae*. Il n'est pas enveloppé et comporte un génome ARN simple brin, de polarité positive. Contrairement au virus sigma, DCV infecte naturellement *D. simulans* et il code une protéine, nommée 1A, qui est un suppresseur de la voie d'ARNi, empêchant le clivage des ARN double brin par Dcr-2 et impactant la production de siARN (van Rij et al., 2006). Par conséquent, DCV est très pathogène pour la drosophile.

L'ensemble de ces analyses a été mis en perspective pour comprendre l'effet d'une infection virale par un virus naturel de drosophile sur la transcription des ET, dans le but d'améliorer les connaissances des interactions hôtes-pathogènes et de la diversité génétique des insectes.

Trois questions principales sont posées dans cette seconde partie :

- L'activité des ET est-elle modulée au cours de l'infection par DMelSV et DCV ?
- La modulation des ET est-elle associée à un changement du taux de production des petits ARN interférents lors de l'infection par DMelSV ?
- Quels sont les points de divergence avec l'étude de la modulation de l'activité des ET lors de l'infection des Makindu par SINV ?

2. Etude de la réplication du virus Sigma (DMelSV) chez *D. simulans*

Nous avons infecté la même lignée de *D. simulans* que dans la première partie des résultats. Il s'agit des Makindu infectées expérimentalement par injection intrathoracique d'un volume de 207 nl de DMelSV. Ce dernier n'induit pas d'effets cytopathiques en culture cellulaire *in vitro*, ce qui implique qu'il ne peut pas être quantifié par des méthodes classiques de titration virale. Par conséquent, dans le but de détecter les drosophiles infectées par DMelSV, nous avons utilisé un test au dioxyde de carbone (CO_2), révélant une paralysie puis la mort des drosophiles infectées (L'Heritier, 1948).

Tout d'abord, le virus DMelSV a été titré en réalisant le test CO_2 à partir de *D. melanogaster* infectées expérimentalement (**Figure S27.**). Nous avons déterminé le nombre d'unités infectieuses du stock de DMelSV en réalisant une série de dilutions de celui-ci, que nous avons ensuite injecté aux drosophiles non infectées (L'heritier, 1958; Longdon et al., 2010). Au jour 14 après infection, les drosophiles mortes ont été associées à une détection de réplication virale du DMelSV. En se basant sur la méthode de Reed et Muench, le titre viral du stock de DMelSV a été estimé à 1,91 log10(DID50/drosophile), soit la dose infectieuse permettant la mort de 50% des drosophiles (DID50, pour *50% Drosophila infectious dose*) ; (Reed and Muench, 1938).

En se basant sur ces premières analyses, nous avons infecté la lignée Makindu en utilisant le stock non dilué (pur) de DMelSV, permettant l'infection de la majorité des *D. melanogaster*. Chez Makindu, nous avons observé une augmentation progressive de la mortalité à partir du jour 7 après infection et jusqu'au jour 11 après infection (**Figure 66.**). Ces résultats révèlent que DMelSV se réplique expérimentalement chez *D. simulans* puisque toutes les drosophiles ont succombé à l'infection suite au test au CO_2, au jour 11 après l'infection. (**Figure 66.**).

De plus, de la même façon que pour la première partie des résultats, nous avons détecté l'infection chronique de NV chez *D. simulans* Makindu infectées par DMelSV et non infectées par DMelSV (mock) à partir de nos analyses de RNA-seq au niveau des carcasses (**Figure 67.**).

On note que l'infection chronique par NV semble être à un niveau basal de réplication assez bas puisque très peu de transcrits sont détectés. Par ailleurs, les transcrits des virus DAV et DCV n'ont pas été détectés (**Figure 67.**). Les drosophiles Makindu infectées expérimentalement par DMelSV révèlent un nombre de transcrits de DMelSV très supérieur dans les carcasses par rapport aux ovaires.

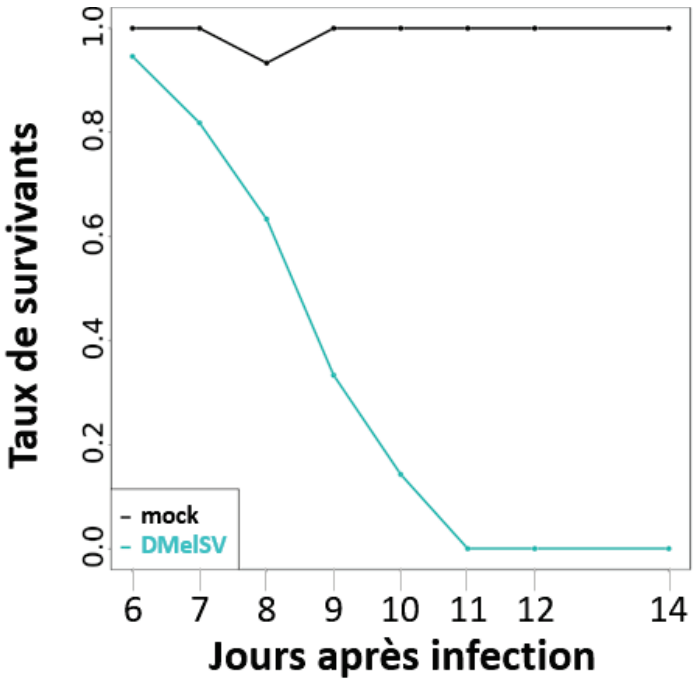


Figure 66. Réplication de DMelSV chez Makindu. Les drosophiles ont été infectées avec un volume de 207 nl de DMelSV par rapport à une condition non infectée (mock) où du milieu de préparation de DMelSV seul a été injecté. Les drosophiles vivantes ont été quantifiées tout au long de la cinétique de réplication, en réalisant le test au CO_2 (J6 à J14 jours après infection). La courbe en bleu représente les drosophiles infectées par DMelSV et la courbe noire représente les drosophiles non infectées. n = 1 expérience (incluant 5 à 35 drosophiles).

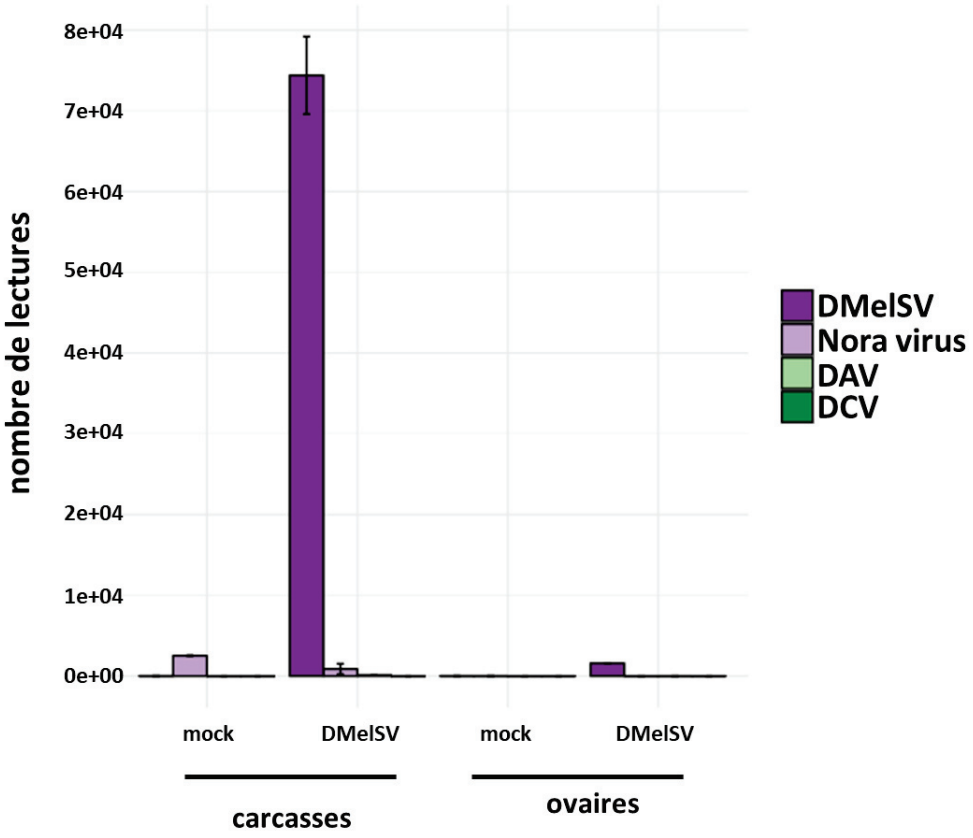


Figure 67. Détection des transcrits viraux chez Makindu. Détermination du nombre de lectures (normalisation par DESeq2) correspondant aux transcrits viraux par RNA-seq pour les infections chroniques suivantes : NV, DAV, DCV, dans les tissus (carcasses et ovaires) issus de drosophiles infectées expérimentalement par DMelSV ou non infectées (mock). n = 2 expériences indépendantes (30 tissus par échantillon).

Dans cette partie, nous avons validé la réplication virale de DMelSV dans les drosophiles Makindu et retrouvé l'infection chronique des Makindu par NV. Par la suite, nous allons étudier la production des petits ARN interférents viraux lors de l'infection par DMelSV dans les carcasses et les ovaires des Makindu.

3. Production des petits ARN interférents viraux chez *D. simulans*

Comme attendu, la réponse de type ARNi mise en place lors de l'infection par DMelSV est principalement médiée par les siARN (**Figure 68.**). Ces siARN ciblant DMelSV sont peu abondants et majoritairement détectés dans les carcasses par rapport aux ovaires. Les proportions de brins sens et antisens sont équivalentes dans les carcasses et les ovaires (**Figure 68 A.**). De plus, on retrouve une distribution homogène des brins sens et antisens des siARN, le long du génome de DMelSV (**Figure 68 B.**). Une infime proportion des transcrits sens et antisens des petits ARN interférents de 23 à 30 nucléotides de longueur ciblant DMelSV a

été détectée dans les carcasses et les ovaires, correspondant au bruit de fond (**Figure 68 A.**). Par conséquent, ces derniers ne présentent pas la signature ping-pong suggérant qu'ils ne sont pas impliqués dans la voie secondaire des piARN.

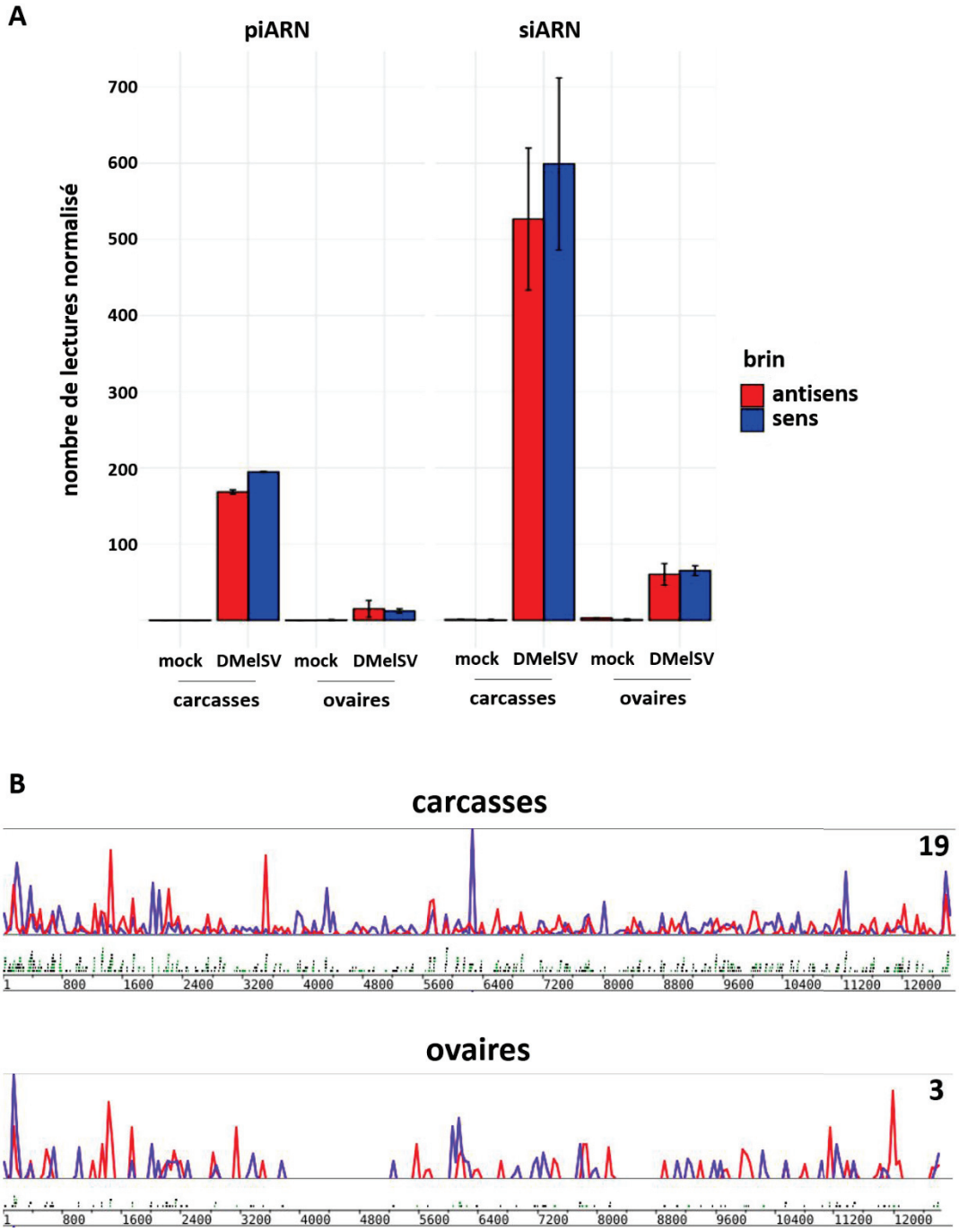


Figure 68. Analyse des siARN et des piARN ciblant DMelSV chez *D. simulans* Makindu. (A) Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des drosophiles Makindu. L'ensemble des comptes des lectures a été normalisé par le nombre de lectures de miARN. **(B)** Distribution des petits ARN interférents de 21 nucléotides ciblant SINV, le long du génome viral, incluant les brins sens (bleu) et antisens (rouge). La valeur numérique en haut à droite correspond au nombre de lectures pour le pic le plus haut. n = 2 expériences indépendantes (50 tissus par échantillon).

Comme nous l'avons décrit précédemment, l'analyse du RNA-seq nous a permis de détecter les transcrits de NV (**Figure 67.**). Cependant, de façon surprenante, l'analyse du séquençage des petits ARN interférents a révélé (en plus des siARN ciblant DMelSV et NV) la production importante de siARN dirigés contre DAV (**Figure S28.**), alors qu'aucun transcrit de DAV n'a été détecté en RNA-seq (**Figure 67.**). Ces observations suggèrent que DAV est efficacement contrôlé par ces siARN viraux, permettant sa persistance à un niveau très faible dans les drosophiles, et le rendant ainsi indétectable. De façon surprenante, dans les carcasses infectées par DMelSV, le nombre total des transcrits sens et antisens des siARN ciblant DAV est supérieur à ceux ciblant les virus NV (moyenne du log2 *fold change* = 3,36) et DMelSV (moyenne du log2 *fold change* = 4,99) et est détecté presque uniquement dans les carcasses infectées par DMelSV par rapport aux carcasses non infectées par DMelSV (moyenne du log2 *fold change* = 9,43) ; (**Figure S28 A.**). Pourtant, l'analyse de l'expression de DAV n'a révélé aucun transcrit viral spécifique dans les drosophiles infectées de façon persistante par DMelSV Hap23, utilisé pour produire le stock viral de DMelSV (**Figure S29.**). A l'inverse, la production de siARN ciblant NV est relativement proche dans les carcasses infectées par DMelSV en comparaison avec les carcasses non infectées par DMelSV, mais reste plus élevée que la production de siARN ciblant DMelSV (**Figure S28 A.** ; moyenne du log2 *fold change* = 1,63). Pour chacun de ces virus, les siARN viraux présentent des proportions équivalentes de brins sens et antisens dans les carcasses et les ovaires (**Figure S28 B.**). Par ailleurs, l'analyse des brins des petits ARN interférents de 23 à 30 nucléotides de long ciblant DAV n'a permis de détecter quasiment que le brin sens, et seulement dans les carcasses lors de l'infection par DMelSV (**Figure S28 B.**). Les brins sens des piARN ciblant NV sont également prédominants par rapport aux brins antisens mais sont détectés à la fois dans les carcasses non infectées par DMelSV et les carcasses infectées par DMelSV (**Figure S28 B.**). Finalement, seuls les petits ARN interférents de 23 à 30 nucléotides de long ciblant DMelSV présentent une proportion équivalente de brins sens et antisens, mais ces derniers sont détectés en très faible quantité par rapport aux brins sens détectés pour les virus DAV et NV (**Figure S28 A-B.**).

4. Dérégulation des voies des petits ARN interférents et de l'expression des éléments transposables lors de l'infection par DMelSV chez *D. simulans*

L'analyse des données de séquençage des petits ARN interférents (Small RNA-seq) et du transcriptome (RNA-seq), effectuée à partir des carcasses et des ovaires de *D. simulans* Makindu, nous a permis d'identifier un impact de l'infection par DMelSV sur le contrôle génomique (piARN et siARN) et ainsi sur le niveau de transcrits des ET. Ces analyses ont été effectuées à partir d'échantillons extraits à 8 jours après infection, au moment où le virus DMelSV se réplique et induit potentiellement des réponses antivirales. L'ensemble des

expérimentations a été fait selon la même procédure que pour la première partie des résultats.

Carcasses :

Ici, l'analyse de 237 familles et sous-familles d'ET annotées révèle une augmentation globale significative du nombre de lectures de transcrits d'ET, dans les carcasses infectées par DMelSV, par rapport aux échantillons non infectés (**Figure 69 A.** ; moyenne log2 *fold change* = 0,31 ; *p-value* = 1,0e-17). Ceci est associé à une diminution significative des petits ARN interférents de 23-30 nucléotides (piARN) (**Figure 69 B.** ; moyenne log2 *fold change* = -0,76 ; *p-value* = 2,1e-35) et de 21 nucléotides (siARN) (**Figure 69 C.** ; moyenne log2 *fold change* = -0,30 ; *p-value* = 5,2e- 14) dirigés contre les ET. Ces résultats suggèrent une réactivation globale des ET par une diminution des voies d'ARNi contrôlant les ET. Cependant, l'association entre l'augmentation de l'activité des ET et la diminution des petits ARN interférents dirigés contre les ET n'est pas significative (**Figure S30 A-B.**), excluant également une dépendance entre la production des siARN et des piARN ciblant les ET (**Figure S30 C.**). Néanmoins, chacune des familles d'ET dont l'activité est augmentée significativement présente une diminution de la production de piARN ciblant les ET (points rouges de la **Figure 30 A.**). Ces résultats suggèrent que la modulation des quantités de transcrits d'ET n'est pas seulement due aux voies des petits ARN interférents mais peut impliquer une autre voie de régulation lors de l'infection par DMelSV. De plus, la classe des rétrovirus endogènes (ERV), représentant la classe d'ET la plus abondante, est significativement plus impactée que les autres classes d'ET (transposons ADN et rétrotransposons à LTR) dans les carcasses (**Figure 70.** ; *Pearson's Chi-squared test* ; *p-value* = 3.2e-2).

Ovaires :

A l'inverse, dans les ovaires, une infime diminution des quantités de transcrits des ET a été identifiée (**Figure 69 D.** ; moyenne log2 *fold change* = -0,18 ; *p-value* = 0,005). De plus, on observe une légère augmentation des petits ARN interférents de 23-30 nucléotides (**Figure 69 E.** ; moyenne log2 *fold change* = 0,31 ; *p-value* = 4,0e-37) et des petits ARN interférents de 21 nucléotides (**Figure 69 F.** ; moyenne log2 *fold change* = 0,24 ; *p-value* = 4,4e-24). Dans cet organe, on confirme que les ET sont principalement contrôlés par la voie des piARN. Malgré la détection des transcrits de DMelSV dans les ovaires des Makindu, il semblerait que ce tissu soit protégé de la modulation des transcrits d'ET, permettant la protection des cellules germinales.

Par ailleurs, les transcrits de gènes impliqués dans la voie des piARN (*piwi*, *aub*, *ago3* et *zuc*) (**Figure S31 A.**) et dans la voie des siARN (*dcr2*, *ago2*, *r2d2* et *loqs*) (**Figure S31 B.**) ne sont pas significativement modulés lors de l'infection, dans les carcasses et dans les ovaires (*Wilcoxon paired test* ; *p-value* > 0,05). Ces résultats indiquent que la modulation de la production des petits ARN interférents n'implique pas une augmentation de la transcription des effecteurs de ces voies lors de l'infection par DMelSV.

Ainsi, lors de l'infection expérimentale par DMelSV, nous avons observé une augmentation des quantités de transcrits des ET et une diminution des petits ARN interférents, incluant les piARN et les siARN ciblant les ET, dans les carcasses. Les transcrits de DMelSV ont été faiblement détectés dans les ovaires des Makindu et quasiment aucune modulation de l'activité des ET n'a été détectée lors de l'infection par DMelSV dans cet organe. Cependant, une augmentation de la production des piARN et des siARN ciblant les ET a été révélée dans les ovaires après infection des Makindu par DMelSV. L'ensemble de ces résultats indique que la modulation des ET ne semble pas être dépendante des quantités de petits ARN interférents dans les carcasses et les ovaires.

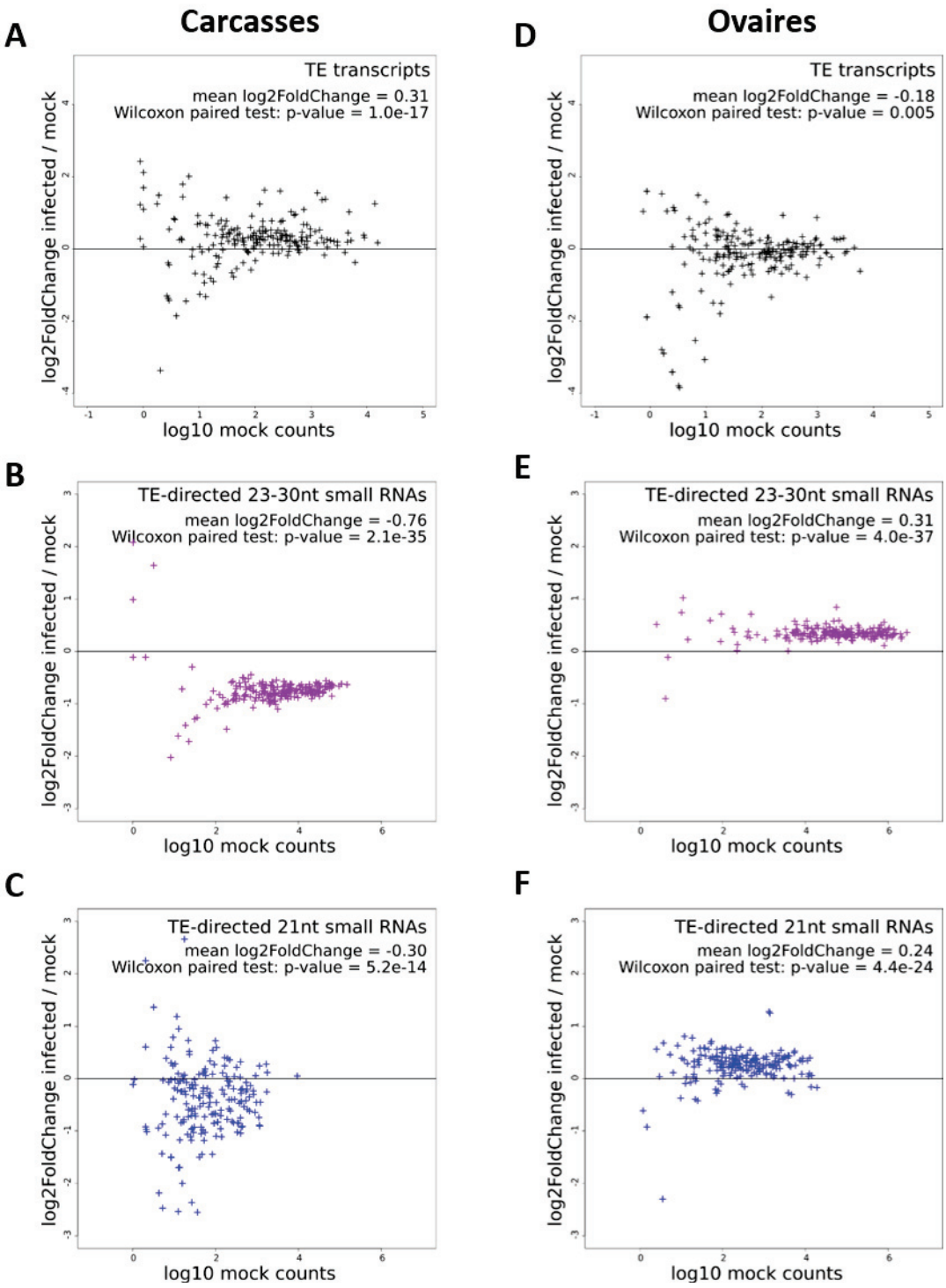


Figure 69. Modulation de l'abondance des transcrits d'ET et des productions de petits ARN interférents ciblant les ET chez Makindu lors de l'infection par DMelSV, dans les carcasses (A-C) et dans les ovaires (D-F). Chaque point représente le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par DMelSV et la condition non infectée (mock), pour les transcrits d'ET **(A, D)**, pour les piARN **(B, E)** et pour les siARN **(C, F)**. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). n = 2 expériences indépendantes.

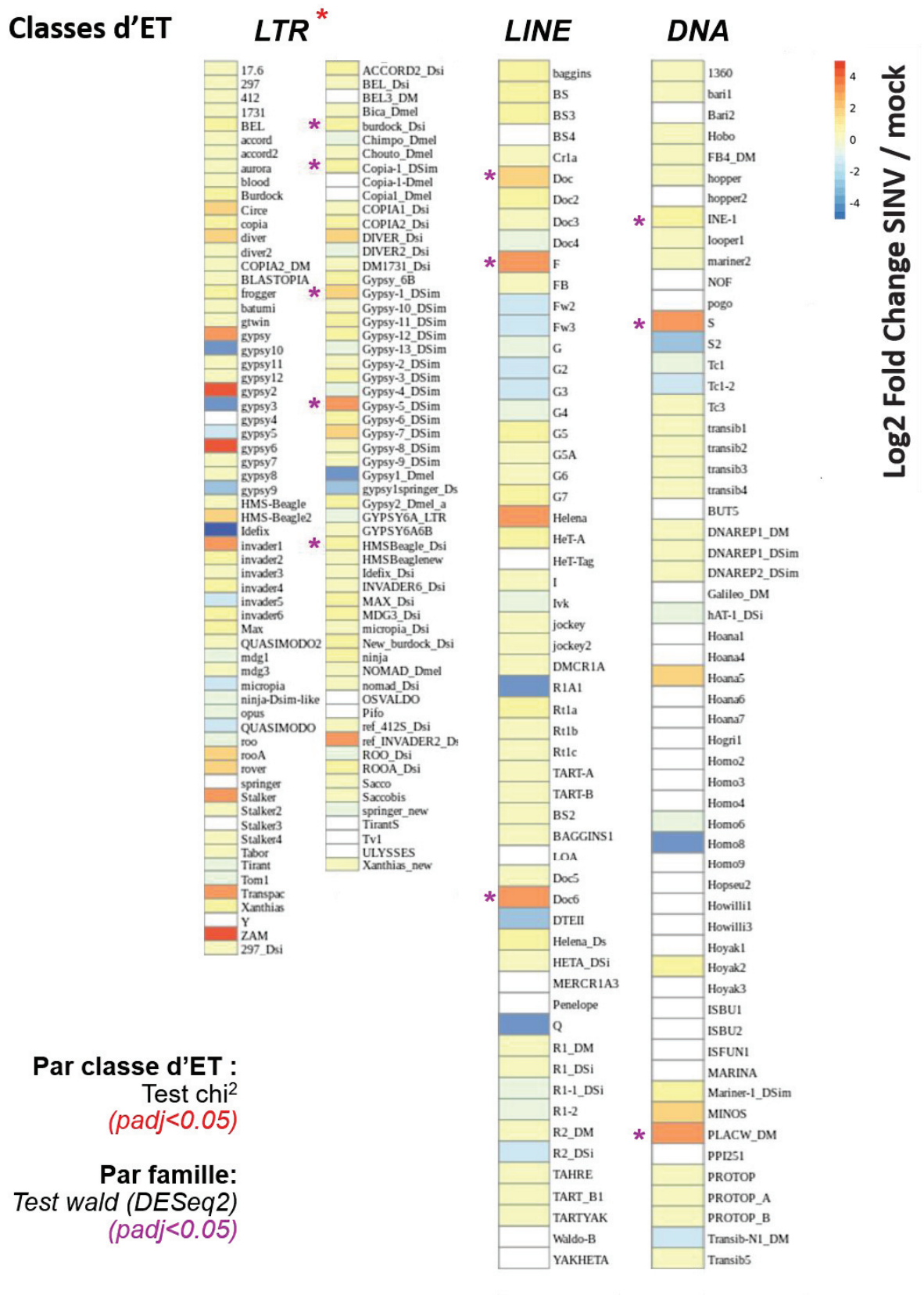


Figure 70. Les ERV sont significativement plus modulés que les autres classes d'ET chez Makindu, lors de l'infection par DMelSV, dans les carcasses. Les *heatmaps* représentent le log2 du ratio des comptes par famille d'ET annotée, entre la condition infectée par DMelSV et la condition non infectée (mock). Une augmentation des

quantités de transcrits d'ET est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. Les familles d'ET en blanc ne sont pas exprimées. La significativité de la différence du nombre de famille d'ET entre classes, modulée lors de l'infection, est établie en comparant le nombre de familles d'ET significativement modulées entre les trois classes d'ET ; rétrotransposons à LTR, LINE et transposons à ADN, et en utilisant le test de khi-deux d'indépendance (test chi2*, Pearson's Chi-squared test = test chisq*). La significativité de la modulation par famille d'ET est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock par DESeq2 (*test wald*). **padj < 0,05* (*p-value* ajustée par le nombre de famille d'ET par classe = padj). n = 2 expériences indépendantes (30 tissus par échantillon de RNA-seq).

Dans cette partie, de façon similaire à l'étude de l'infection des Makindu par SINV (**Figure 44.**), on retrouve une réponse de type ARNi, principalement médiée par les siARN lors de l'infection des Makindu par DMelSV. De plus, la proportion des brins sens et antisens détectée correspondant aux siARN ciblant DMelSV est semblable à celle des siARN ciblant SINV dans les carcasses et les ovaires. Cependant, nous avons observé des points de divergence concernant la production de petits ARN interférents viraux de 23 à 30 nucléotides entre DMelSV et SINV. En effet, alors que dans les Makindu infectées par DMelSV on observe une proportion équivalente de brins sens et antisens, dans les Makindu infectées par SINV, nous avons principalement détecté le brin sens, incluant un très faible nombre de lectures. Pour autant, l'ensemble de ces petits ARN interférents viraux (DMelSV et SINV) de 23 à 30 nucléotides, ne présentent pas la signature ping-pong caractéristique des piARN secondaires. Par ailleurs, puisque les drosophiles Makindu sont chroniquement infectées par NV, nous avons détecté les transcrits de NV dans les carcasses des Makindu infectées ou non par DMelSV (**Figure 67.**) et dans celles infectées ou non par SINV (**Figure 39.**). Cependant, tandis que nous avons observé un nombre similaire de lectures correspondant aux transcrits de NV dans les carcasses non infectées et infectées par SINV (**Figure 39.**), très peu de transcrits de NV ont été détectés dans les carcasses infectées par DMelSV par rapport aux carcasses non infectées (**Figure 67.**). Ces carcasses non infectées présentent une proportion équivalente de transcrits de NV par rapport aux carcasses des Makindu étudiées dans l'étude de SINV. Par ailleurs, la production de siARN ciblant NV est très élevée dans les carcasses non infectées et infectées par SINV (**Figure 48.**) par rapport à celle identifiée dans les carcasses des Makindu non infectées ou infectées par DMelSV, présentant également une grande quantité de siARN ciblant DAV uniquement lors de l'infetion par DMelSV (**Figure S28.**). Concernant les petits ARN de 23 à 30 nucléotides ciblant NV, le brin sens est prédominant dans toutes les conditions, et il semble y en avoir moins dans les carcasses des Makindu non infectées ou infectées par DMelSV (**Figure S28.**) par rapport aux carcasses des expériences concernant SINV (**Figure 48**.). De plus, comme pour les siARN ciblant DAV, les petits ARN interférents de 23 à 30 nucléotides de long ciblant DAV ne sont détectés que dans les carcasses infectées par DMelSV et ne révèlent quasiment que le brin sens (**Figure S28.**).

Ces analyses nous ont permis d'avoir une connaissance précise des infections virales existantes dans le modèle Makindu et sont importantes à prendre en compte dans l'étude de la modulation de l'activité des ET et des petits ARN interférents lors de l'infection virale.

Carcasses :

En effet, la modulation de l'activité des ET n'est pas la même lors de l'infection par DMelSV par rapport à SINV (**Figure 71.**). Dans les carcasses, nous avons observé un profil opposé de modulation de l'activité des ET et de la production des petits ARN interférents entre les infections par les virus DMelSV et SINV (**Figure 71 A-B.**). En effet, dans les carcasses Makindu chroniquement infectées par NV, on observe une augmentation de l'activité des ET et une diminution des quantités de petits ARN interférents dirigés contre les ET suite à l'infection par DMelSV en comparaison avec la condition non infectée (**Figure 71 A.**). Dans cette même comparaison, nous avons observé une production importante de siARN ciblant DAV. A l'inverse, lors de l'infection des Makindu par SINV, ces drosophiles sont également chroniquement infectées par NV mais révèlent une diminution de l'activité des ET et une augmentation des quantités de petits ARN interférents ciblant les ET dans les carcasses suite à l'infection virale en comparaison avec la condition non infectée (**Figure 71 B.**). De plus, les modulations des quantités de ces petits ARN interférents dans les carcasses n'impliquent pas de modulation du niveau de transcrits des gènes codant les principaux effecteurs de ces voies d'ARNi (**Figure 71 A-B.**).

Ovaires :

Dans les ovaires, il n'y a quasiment pas de modulation de l'activité des ET lors de l'infection par DMelSV et par SINV (**Figure 71 C-D.**). Cependant, dans les ovaires, la stabilité au niveau des quantités de transcrits d'ET n'implique pas les mêmes modulations de la production des petits ARN interférents entre les deux virus (**Figure 71 C-D.**). Lors de l'infection par DMelSV, nous avons détecté une légère augmentation des piARN et des siARN ciblant les ET (**Figure 71 C.**) tandis qu'une forte modulation en miroir des siARN et des piARN ciblant les ET a été observée dans les ovaires lors de l'infection par SINV (**Figure 71 D.**). De plus, les modulations de quantités de petits ARN interférents dans les ovaires n'impliquent pas de modulation du niveau de transcrits des gènes codant les principaux effecteurs de ces voies d'ARN interférence (**Figure 71 C-D.**). Ces résultats révèlent que chez *D. simulans* Makindu, l'infection expérimentale par un virus de drosophile n'a pas le même impact sur l'activité des ET et sur les quantités de petits ARN interférents ciblant les ET par rapport à un arbovirus, tel que SINV.

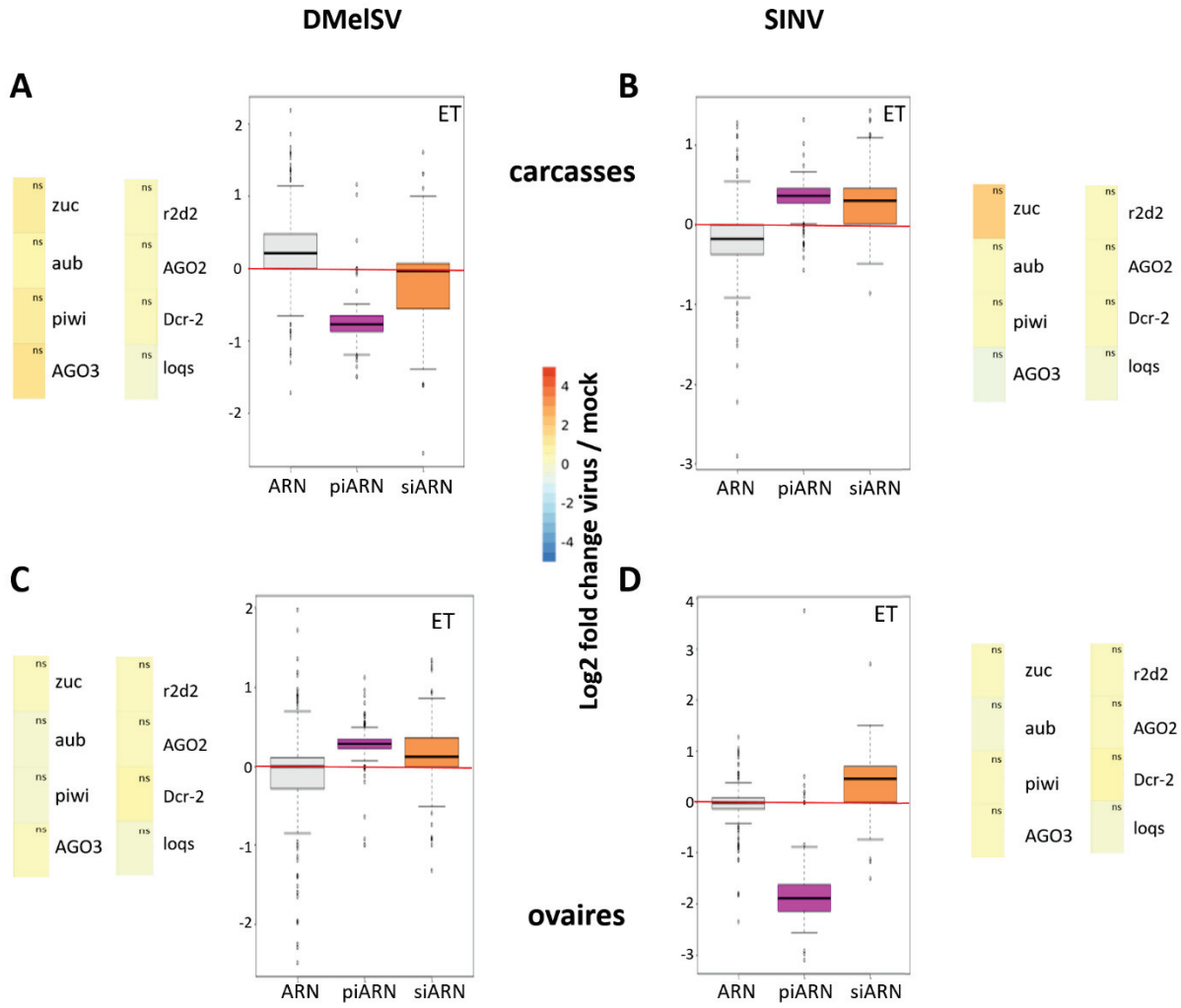


Figure 71. Comparaison des modulations de l'activité des ET et des petits ARN interférents ciblant les ET entre les infections par DMelSV et SINV. Dans les carcasses **(A-B)** et dans les ovaires **(C-D)**, pour les virus DMelSV (A-C) et SINV (B-D), nous avons comparé la condition infectée avec la condition non infectée (mock). Les *boxplots* représentent les modulations en log2 *fold change* des quantités de transcrits des ET et des petits ARN interférents (piARN et siARN). Les *heatmaps* représentent le log2 *fold change* des comptes des transcrits des gènes des voies piARN (*zuc*, *aub*, *piwi*, *ago3*) ; siARN (*r2d2*, *ago2*, *dcr2*, *loqs*). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. Les modulations des taux de transcrits des gènes sont déterminées par DESeq2 et ne sont pas significatives (ns, non significatif ; *p-value* > 0,05). n = 2 expériences indépendantes.

5. Dérégulation des gènes impliqués dans la voie Toll lors de l'infection par DMelSV chez *D. simulans*

De plus, l'analyse de « Gene Ontology », permettant de déterminer s'il existe un enrichissement en certaines fonctions parmi les gènes différentiellement exprimés par

rapport à des gènes pris aléatoirement, nous a permis de détecter une forte modulation des gènes impliqués dans la régulation de la voie Toll, lors de l'infection par DMelSV au niveau des carcasses des Makindu (**Figure 72.**). Nous avons notamment détecté dans les carcasses une induction significative du gène *PGRP-SB1* (**Figure 73.** ; log2 *fold change* = 0,69 ; *p- value* = 3,48e-4) lors de l'infection par DMelSV, comme décrite précédemment en drosophile entière par Tsai et al., (2008). Cependant, contrairement à Tsai et al., (2008), la modulation observée de l'expression du gène *PGRP-SD* n'est pas significative (log2 *fold change* = 0,60 ; *p-value* = 5,81e-2), mais une induction significative de *PGRP-SA* (log2 *fold change* = 1,91 ; *p-value* = 7,90e-20) a été révélée lors de l'infection par DMelSV, dans les carcasses (**Figure 73.**). Les PGRP (*peptidoglycan recognition proteins*) sont les premiers récepteurs qui reconnaissent des composants de surface au niveau des membranes des bactéries (Lemaitre and Hoffmann, 2007). L'expression des gènes *PGRP-SD* et *PGRP-SA* est nécessaire à l'activation de la voie Toll (Bischoff et al., 2004; Tanji and Ip, 2005), tandis que le gène PGRP-SB1 est régulé par la voie Imd (Zaidman-Rémy et al., 2011) mais présente également une activité bactéricide (Mellroth and Steiner, 2006). De plus, les gènes *toll* (log2 *fold change* = 0,57 ; *p-value* = 2,18e-3) et *spaetzle* (log2 *fold change* = 0,76 ; *p-value* = 2,19e-3), impliqués dans la voie Toll, sont significativement surexprimés lors de l'infection par DMelSV dans les carcasses (**Figure 73.**). Enfin, dans ces mêmes tissus, en accord avec Swevers et al., (2018) et Tsai et al., (2008), l'induction significative du gène *drosomycine* (log2 *fold change* = 1,50 ; *p-value* = 1,26e-17), codant l'AMP Drosomycine, et exprimé sous le contôle de la voie Toll a également été observée lors de l'infection par DMelSV (**Figure 73.**). Ces résultats montrent une stimulation de l'expression de plusieurs gènes impliqués directement dans la voie Toll mais aussi des gènes codant des effecteurs en amont ou en aval de la voie Toll.

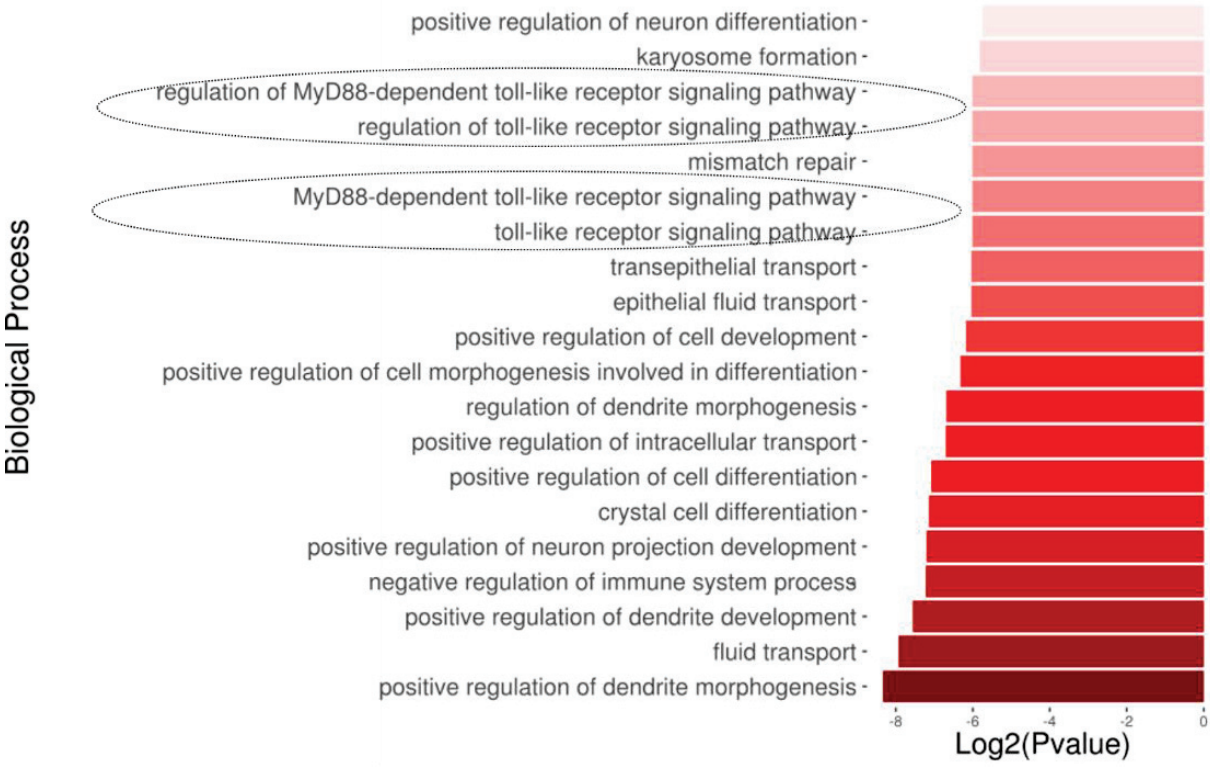


Figure 72. Parmi les gènes différentiellement exprimés lors de l'infection par DMelSV, on trouve un enrichissement en gènes impliqués dans les processus de régulation de la voie Toll. L'analyse de Gene Ontology (GO) a été réalisée pour déterminer s'il existe un enrichissement en certaines fonctions parmi les gènes différentiellement exprimés lors de l'infection par DMelSV par rapport à la condition non infectée. Les résulats sont exprimés en log2 de la *p-value* pour l'ensemble des gènes classés dans un processus biologique donné.

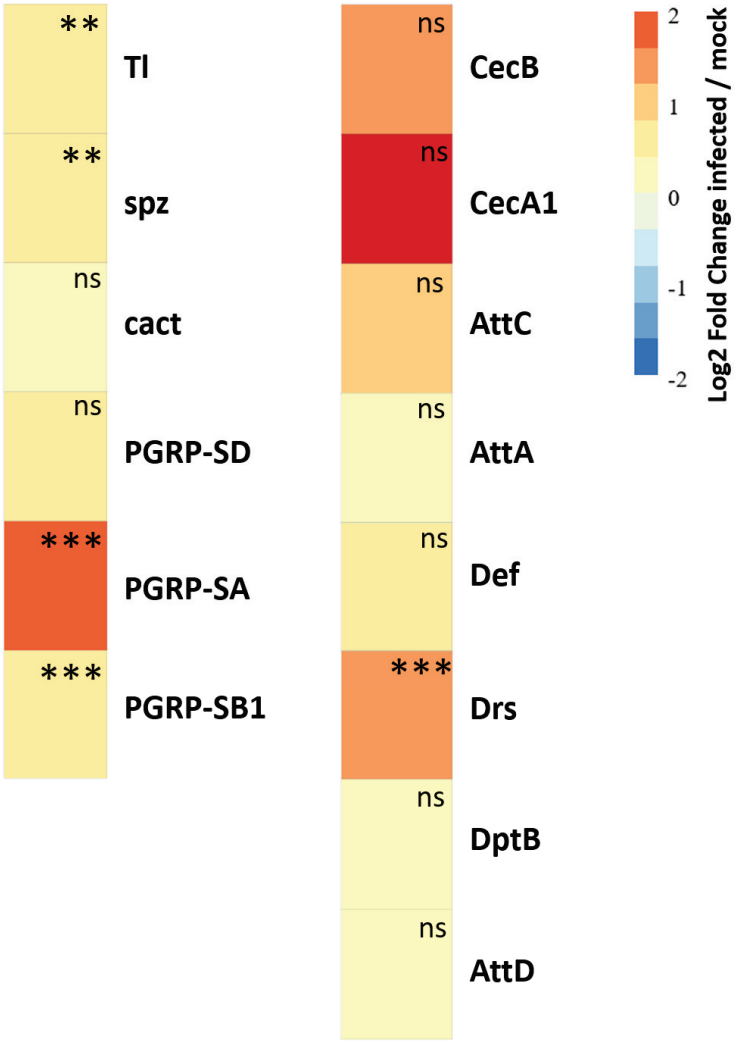


Figure 73. Surexpression de gènes impliqués dans la voie Toll et de certains AMP chez Makindu, lors de l'infection par DMelSV, dans les carcasses. Les *heatmaps* représentent le log2 *fold change* des comptes des gènes, entre la condition infectée par DMelSV et la condition non infectée (mock). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock par DESeq2 (*test wald*). **p-value < 0,05 ; P** < 0.01 ; P** < 0.001*. n = 2 expériences indépendantes (30 tissus par échantillon de RNA-seq). Légende : Tl (*toll*), spz (*spätzle*), cact (*cactus*), PGRP-SD (*Peptidoglycan-recognition protein SD*), PGRP-SA (*Peptidoglycan recognition protein SA*), PGRP-SB1 (*Peptidoglycan recognition protein SB1*), CecB (*Cecropin B*), CecA1 (*Cecropin A1*), AttC (*Attacin-C*), AttA (*Attacin-A*), AttD (*Attacin-D*), Def (*Defensi*n), Drs (*Drosomycin*), DptB (*Diptericin B*).

6. Dérégulation des quantités de transcrits des éléments transposables lors de l'infection par DCV chez *D. melanogaster*

A. Réplication de DCV et tolérance virale

Des données de séquençage de transcriptome (RNA-seq) d'échantillons de drosophiles entières (*D. melanogaster*) et de corps gras ont été analysées de la même façon que précédemment, permettant notamment l'étude des transcrits des ET et des virus. Ils se réfèrent à l'article de Merkling et al., (2015). Ce jeu de données ne présente pas de réplicat, excluant l'analyse statistique par DESeq2 et il n'y a pas de donnée de Small RNA-seq. Par ailleurs, les mêmes drosophiles (w^{1118}) ont été mutées pour l'une des trois H3K9 méthyltransférases, appelée *G9a*, et les transcriptomes de ces mutants entiers et des corps gras correspondants ont été séquencés (Merkling et al., 2015). Les marques H3K9me2 sont essentielles pour la mise en place de l'hétérochromatine et la régulation des gènes (Hublitz et al., 2009). La méthyltransférase *G9a* induit des diméthylations H3K9 chez *Drosophila*, mais est associée à des régions euchromatiques (Kramer et al., 2011; Stabell et al., 2006). L'ensemble de ces drosophiles, présentant un fond génétique identique, avait été affranchi de toute infection virale chronique et de la bactérie endosymbiotique *Wolbachia* (Merkling et al., 2015).

Je vais tout d'abord vous présenter le jeu de données de Merkling et collaborateurs, impliquant une caractérisation de la réplication de DCV par ces auteurs (Merkling et al., 2015). Les drosophiles ont été infectées expérimentalement par injection intrathoracique du virus DCV (10000 TCID50) (Merkling et al., 2015). Tout d'abord, Merkling et collaborateurs ont montré une sensibilité des drosophiles sauvages infectées par DCV, à partir du jour 4 après infection et aboutissant à la mortalité de l'ensemble des drosophiles au jour 13 après infection (**Figure S32 A.**). De plus, les drosophiles mutées pour le gène *G9a* présentent une sensibilité accrue à l'infection par DCV, puisqu'elles succombent toutes à l'infection au jour 6. Les drosophiles ($G9a^{+/+}$ et $G9a^{-/-}$) non infectées (mock) ne présentent aucune mortalité au cours du temps. Par ailleurs, la perte de la tolérance virale des drosophiles $G9a^{-/-}$ n'est pas due à une augmentation du titre viral de DCV, puisqu'il n'y a pas de différence significative des charges virales entre les drosophiles sauvages et celles mutées pour *G9a*, lors de l'infection par DCV, en mouches entières (**Figure S32 B.**) et au niveau du corps gras (**Figure S32 C.**). On remarque une importante réplication de DCV dans le corps gras pour l'ensemble des drosophiles (Merkling et al., 2015). En effet, le corps gras est un organe impliqué dans le métabolisme et l'immunité (Lemaitre and Hoffmann, 2007) et il est une cible très importante de DCV (Deddouche et al., 2008). Enfin, Merkling et collaborateurs ont montré que la perte de la tolérance virale des mutants *G9a*, lors de l'infection par DCV, était due à une forte augmentation de l'expression de gènes dépendants de la voie Jak/STAT, tels que *vir-1*, *TotA*, et *TotM* (**Figure S32 D-E**).

B. Modulation des quantités de transcrits des ET

L'analyse du transcriptome par RNA-seq a été réalisée à 24 heures après infection par DCV pour s'affranchir de la grande mortalité des drosophiles, observée notamment pour les mutants *G9a* (Merkling et al., 2015). A partir de ces données de RNA-seq, nous avons identifié une augmentation significative globale des quantités de transcrits des ET lors de l'infection par DCV, dans les drosophiles w^{1118} entières (**Figure 74 A.** ; moyenne log2 *fold change* = 0,28 ; *p-value* = 9,1e-16) et dans le corps gras (**Figure 74 B.** ; moyenne log2 *fold change* = 0,73 ; p- value = *2,3e-29*). De plus, l'impact sur la transcription des ET semble davantage marqué dans le corps gras, où DCV se réplique abondamment (Deddouche et al., 2008). Ces résultats révèlent une réactivation globale des ET lors de l'infection par DCV. De la même façon que précédemment, il serait intéressant d'analyser les répertoires de petits ARN interférents ciblant les ET, dans le but de déterminer si cette modulation dépend d'une diminution de la voie de contrôle génomique des ET.

De plus, les quantités de transcrits des ET sont également augmentées chez les mutants *G9a*, dans les drosophiles entières (**Figure 74 C.** ; moyenne log2 *fold change* = 0,57 ; *p-value* = 3,2e-7), et dans le corps gras (**Figure 74 D.** ; moyenne log2 *fold change* = 1,07 ; *p-value* = 4,9e- 17), lors de l'infection par DCV, en comparaison avec la condition non infectée (mock). Ces résultats suggèrent que la méthyltransférase H3K9, nommée G9a, permettant la formation de l'hétérochromatine, n'est pas indispensable à la modulation des transcrits d'ET, lors de l'infection par DCV, puisque l'augmentation de l'activité des ET est également retrouvée chez les drosophiles infectées par DCV mais dépourvues de méthyltransférase *G9a* fonctionnelle (**Figure 74 C-D.**). Néanmoins, il est intéressant de noter que la modulation observée semble encore plus marquée chez les mutants *G9a*, par rapport aux drosophiles sauvages w^{1118}, à partir des échantillons de drosophiles entières et de corps gras (**Figure 74.**). De plus, la comparaison entre les comptes des mutants *G9a* et les comptes des drosophiles sauvages w^{1118} révèle une diminution des quantités de transcrits d'ET chez les drosophiles entières (**Figure 75 A.** ; moyenne log2 *fold change* = -0,28 ; *p-value* = 6,4e-9) et dans le corps gras (**Figure 75 B.** ; moyenne log2 *fold change* = -0,33 ; *p-value* = 5,3e-6) en condition non infectée et également lors de l'infection par DCV à partir des drosophiles entières (**Figure 75 C.** ; moyenne log2 *fold change* = -0,33 ; *p-value* = 2,3e-4) et du corps gras (**Figure 75 D.** ; moyenne log2 *fold change* = -0,14 ; *p-value* = 1,7e-3). Etonnament, nos résultats indiquent une diminution de l'activité des ET à partir des mutants *G9a* (drosophiles entières et corps gras) infectés et non infectés par DCV (**Figure 75.**). En effet, une augmentation de l'activité des ET était davantage attendue puisqu'il a été montré que les marques épigénétiques répressives (H3K3me2) sont répandues au niveau de régions d'ADN proches des ET et sont impliquées dans la répression épigénétique des ET (Lee and Karpen, 2017).

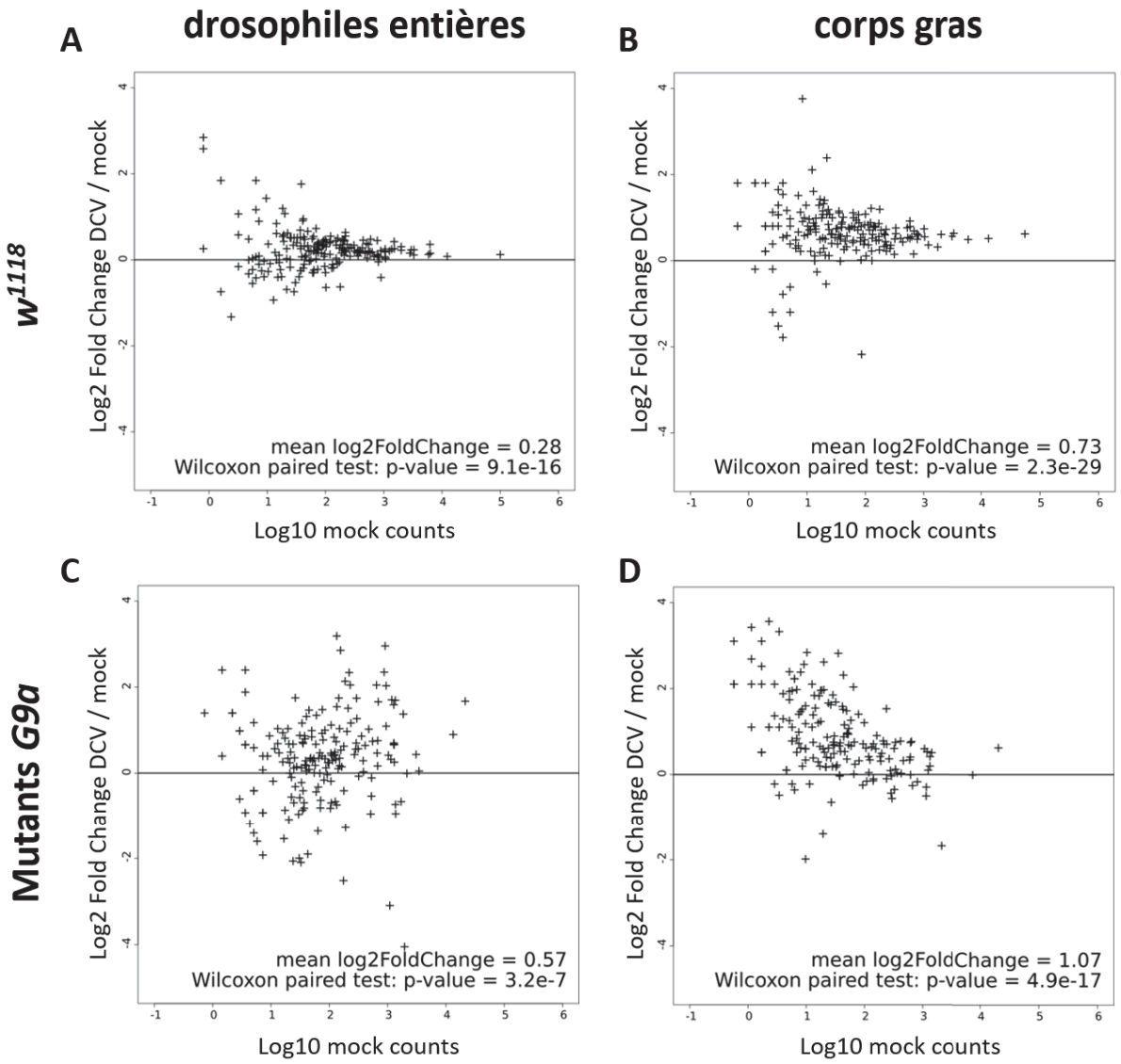


Figure 74. Modulation de l'abondance des transcrits des ET chez *D. melanogaster*, lors de l'infection par DCV dans les drosophiles entières (A ; C) et dans le corps gras (B ; D). Chaque point représente le log2 du ratio des comptes des transcrits par famille d'ET annotée, entre la condition infectée par DCV et la condition non infectée (mock), pour les drosophiles w^{1118} (A-B) et pour les mutants *G9a* (C-D). La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock et en utilisant les tests appariés de Wilcoxon (*Wilcoxon paired tests*). n = 1 expérience (incluant 30 drosophiles entières ou 100 corps gras).

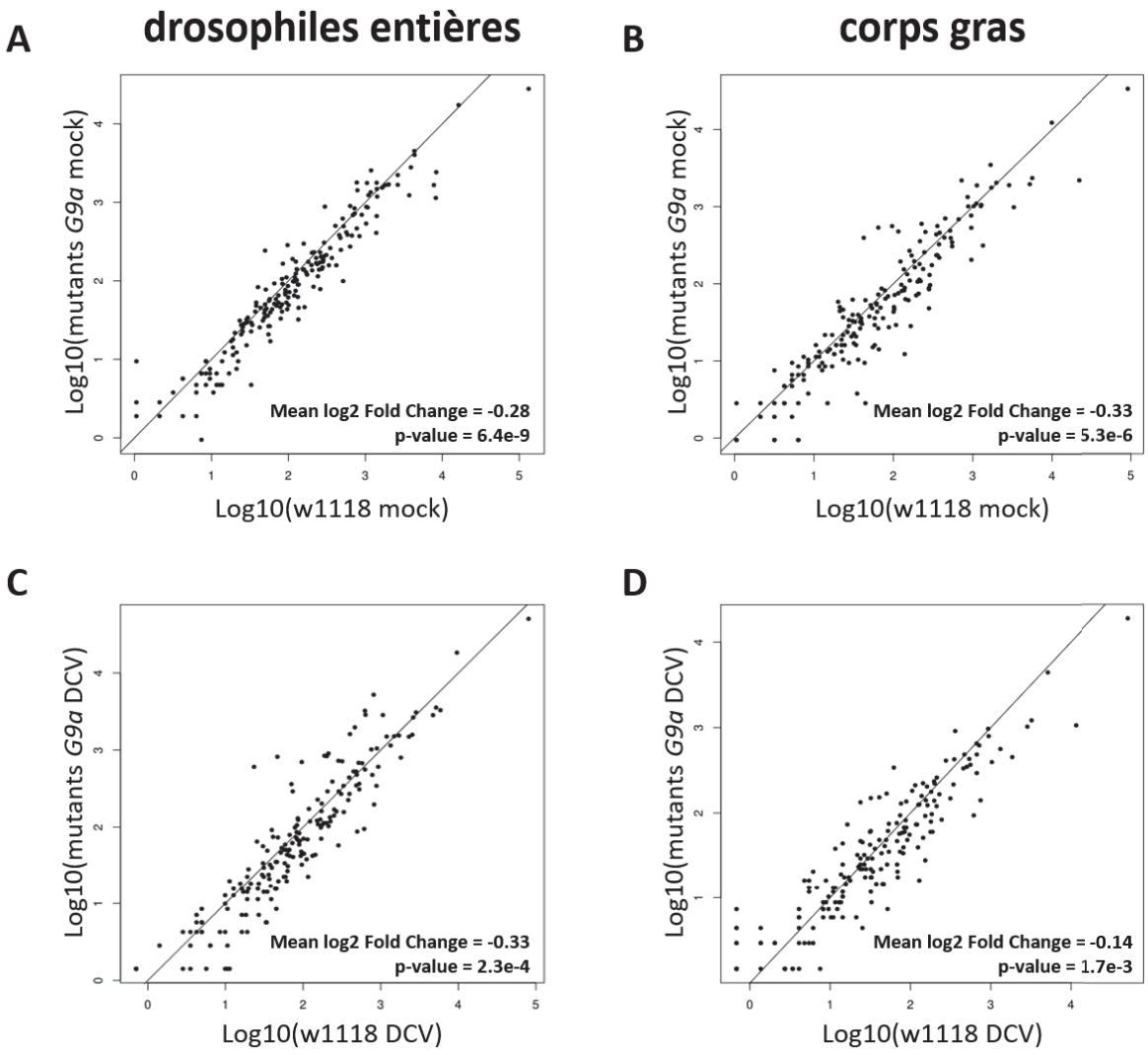


Figure 75. Comparaison des quantités de transcrits d'ET entre les *w*^{1118} et les mutants *G9a*. Chaque point représente une famille d'ET. La droite est y = x. Les comptes des transcrits d'ET sont comparés entre *w*^{1118} et mutants *G9a* en condition mock **(A-B)** et en condition infectée par DCV **(C-D)**, à partir des drosophiles entières (A ; C) et dans le corps gras (B ; D). Les comptes des mutants *G9a* sont significativement inférieurs à ceux des *w*^{1118} en condition non infectée, dans les drosophiles entières (moyenne du log2 *fold change* = -0,28 ; *p-value* = 6,4e-9) et dans le corps gras (moyenne du log2 *fold change* = -0,33 ; *p-value* = 5,3e-6) et lors de l'infection par DCV, dans les drosophiles entières (moyenne du log2 *fold change* = -0,33 ; *p-value* = 2,3e-4) et dans le corps gars (moyenne du log2 *fold change* = -0,14 *; p- value* = 1,7e-3). Test statistique utilisé : Wilcoxon paired test ; n = 1 expérience (incluant 30 drosophiles entières ou 100 corps gras).

Discussion (II)

Les analyses de deux virus infectant naturellement la drosophile (DMelSV et DCV) révèlent une réactivation globale des ET lors de l'infection virale (**Figures 69. et 74.**). De plus, l'effet est davantage marqué au niveau des organes cibles de réplication du virus, comme décrit pour le corps gras lors de l'infection par DCV chez *D. melanogaster* où la modulation des transcrits des ET est plus intense. De la même façon, lors de l'infection par DMelSV chez *D. simulans*, la modulation des ET est principalement observée au niveau des carcasses, par rapport aux ovaires, où la charge virale semble inférieure. Par ailleurs, comme présenté précédemment, ces résultats révèlent un profil de modulation opposé par rapport à celui observé lors de l'infection par SINV dans les carcassses chez *D. simulans* (**Figure 71.**). Ces résultats montrent un impact différent de l'infection par un virus spécifique de drosophile et par un arbovirus sur l'activité des ET et sur la production des petits ARN interférents ciblant les ET.

D'autre part, il semblerait également que les virus contenant un génome ARN de polarité négative, tel que DMelSV, produisent moins d'ARN double brin que les autres virus à ARN (Obbard et al., 2009), et aucune information n'a été décrite concernant le rôle antiviral de la voie des siARN lors de l'infection par DMelSV, ou encore l'existence d'un suppresseur d'ARNi codé par DMelSV. Concernant DCV, la voie des siARN est modérée puisque la protéine 1A de DCV est un suppresseur de la voie d'ARN interférence, empêchant le clivage des ARN double brin par Dcr-2 et ne permettant pas la production maximale de siARN (van Rij et al., 2006). De la même façon que lors de l'infection par DCV (**Figure 74.**), l'infection par DMelSV engendre une augmentation de l'activité des ET, et cette dernière implique aussi une diminution des quantités des siARN et des piARN ciblant les ET dans les carcasses (**Figure 69.**). Ainsi, on peut supposer que lors de l'infection virale, la machinerie de l'ARNi n'est pas stimulée de façon optimale, ce qui engendre une faible production de siARN dirigés contre les ET (**Figure 76.**). Ces résultats sont également cohérents avec ceux obtenus chez *D. melanogaster Dcr2*$^{[L811fsX]}$, présentant une mutation au niveau du gène codant la protéine Dcr-2 et où on retrouve une augmentation des quantités de transcrits d'ET et une diminution de la production des piARN ciblant les ET lors de l'infection virale par SINV par rapport à la condition mock (**Figure 63.**). Nous proposons dans ces conditions, incluant l'infection par DMelSV et DCV, une saturation de la voie des piARN par les transcrits viraux lors de l'infection virale, induisant une diminution des piARN ciblant les ET dans les carcasses (**Figure 76.**). Il serait intéressant de confirmer notre modèle en quantifiant les petits ARN interférents (*Small RNA-seq*) et particulièrement les piARN ciblant les ET lors de l'infection par DCV. De plus, comme mentionné précédemment aucun suppresseur d'ARNi n'a été décrit pour DMelSV, mais cela reste à appronfondir au regard de nos résultats. Ainsi, on peut proposer que lors de l'infection virale, lorsque la voie des siARN est inexistante ou non optimale (mutants *Dcr2*$^{[L811fsX]}$ ou suppresseur d'ARNi), la

voie des piARN est saturée par des ARN viraux, représentant une menace plus importante que les ET en somatique et induisant une diminution des piARN ciblant les ET.

Dans les ovaires, il n'y a quasiment pas de modulation de l'activité des ET lors de l'infection par DMelSV chez *D. simulans* (**Figure 76.**). Ces résultats sont similaires à ceux observés dans les ovaires lors de l'infection par SINV chez *D. simulans* (**Figure 71.**). Néanmoins, les quantités de petits ARN interférents ne sont pas modulées de la même façon entre ces deux virus, suggérant à nouveau que la pression de sélection du maintien de l'activité des ET est plus forte dans les ovaires que dans les carcasses mais n'implique pas les mêmes mécanismes entre DMelSV et SINV. Cependant, dans ces deux conditions on détecte une augmentation des quantités de siARN ciblant les ET (**Figure 76.**). Ces résultats nécessitent d'être approfondis, par exemple en étudiant l'impact d'autres voies de réponse immunitaire sur les modulations des petits ARN interférents induit lors de l'infection virale.

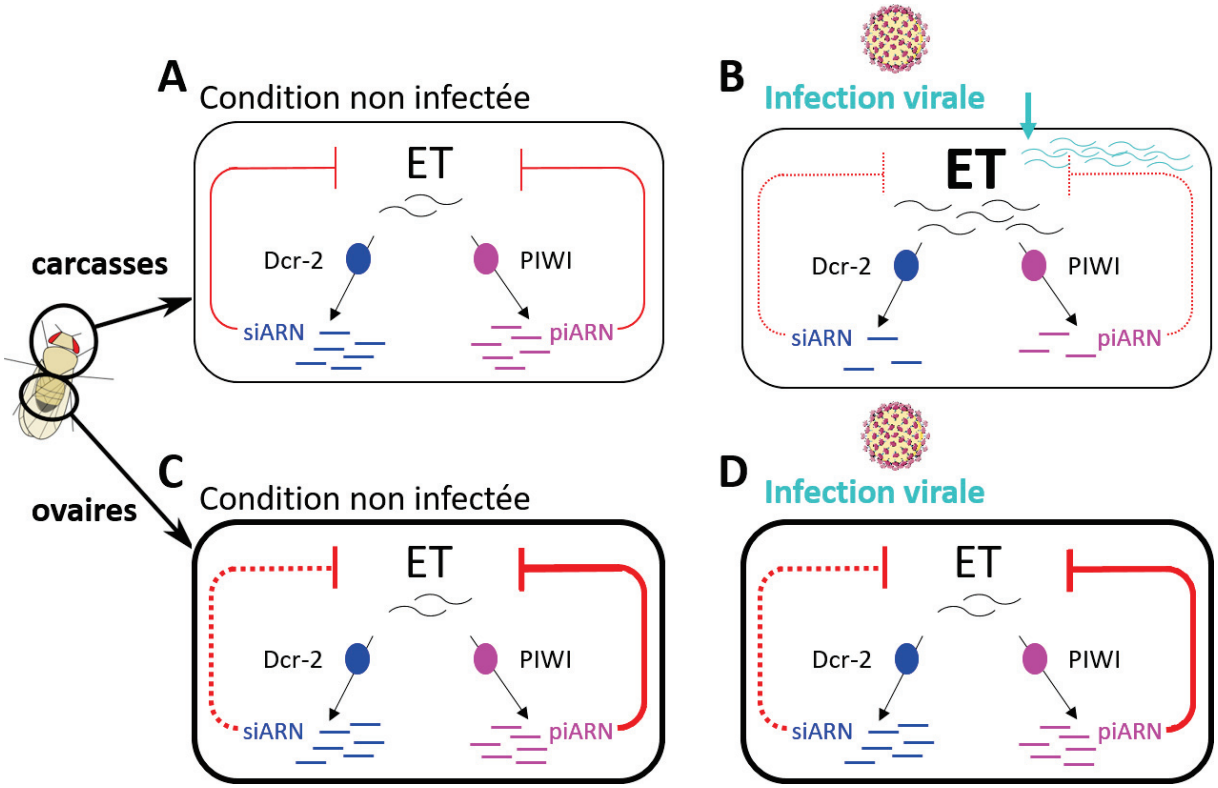


Figure 76. Modèle de l'impact de l'infection par DMelSV et DCV sur les ET. Les mécanismes de modulation de la transcription des ET et de la production des petits ARN interférents dirigés contre les ET et les virus ont été étudiés de façon indépendante dans les carcasses (A-B) et les ovaires (C-D) chez la drosophile. Nous avons analysé ces modulations lors de l'infection virale (B-D) en comparaison avec la condition non infectée (A-C). **(A)** En condition non infectée, les transcrits d'ET sont contrôlés par les piARN et les siARN, dont les principaux effecteurs sont respectivement Dicer-2 et PIWI. **(B)** L'infection virale implique la production d'ARN double brin qui ne sont pas clivés de façon optimale par Dicer-2 (DCV code un suppresseur d'ARN interférence qui cible Dcr-2 alors que pour DMelSV cette étape reste à élucider), induisant une diminution des siARN ciblant les ET et une saturation de la machinerie piARN par les transcrits viraux, et résultant en une diminution des piARN ciblant les ET et enfin une augmentation des quantités de transcrits des ET. **(C-D)** Dans les deux conditions (non infectée et infectée

par DMelSV), les ET sont contrôlés principalement par les piARN, comme l'indique les traits rouges pleins. Dans cet organe, il n'y a quasiment pas d'impact sur la transcription des ET lors de l'infection des drosophiles. Les ovaires semblent protégés des effets de l'infection, se mettant en place dans le reste du corps de la drosophile, comme illustré par les traits noirs épais.

De plus, dans le cas de DMelSV, les quantités de petits ARN interférents ciblant les ET (piARN et siARN) sont également modulées lors de l'infection virale (**Figure 69.**) mais ne sont pas significativement associées aux quantités de transcrits d'ET (**Figure S30.**). Ces résultats suggèrent que la modulation des quantités de transcrits d'ET n'est pas seulement due aux voies des petits ARN interférents mais peut impliquer une autre voie de régulation lors de l'infection par DMelSV. Une voie de régulation candidate pourrait être la voie Toll. En effet, nous avons observé une dérégulation importante de l'expression des gènes impliqués dans la voie Toll, dans les carcasses lors de l'infection par DMelSV, chez *D. simulans* Makindu. Par ailleurs, le rôle de la voie Toll dans la réponse à l'infection par DMelSV est controversé. En effet, une étude a permis l'identification du gène de résistance principal de DMelSV, nommé *ref(2)P*, qui est un composant de la voie de signalisation Toll (Avila et al., 2002; Carré-Mlouka et al., 2007). Dans une autre étude, l'expression de différents gènes impliqués dans la régulation de la voie Toll ou en aval de cette voie, était significativement stimulée ; cette stimulation était visible mais non significative pour les composants clés de cette voie, tel que le gène *toll* (Tsai et al., 2008). Nos résultats ont également permis de détecter une induction significative des gènes codant les récepteurs impliqués dans l'activation de la voie Toll (PGRP), mais également des gènes principaux de cette voie tels que les gènes *toll*, *spätzle* et *cactus* (Arnot et al., 2010; Meng et al., 1999)(**Figure 73.**). Cependant, Carpenter et collaborateurs n'ont pas détecté de stimulation de la voie Toll lors de l'infection par DMelSV (Carpenter et al., 2009). Ces divergences pourraient être expliquées par le mode d'infection (infection expérimentale par injection ou infection naturelle) des drosophiles, le génotype des drosophiles étudiées et par les infections persistantes impliquées (Carpenter et al., 2009). De plus, il a été montré que la voie Toll est importante pour la tolérance virale des drosophiles *D. melanogaster* infectées oralement par DCV (Ferreira et al., 2014), incluant une augmentation de l'expression des gènes régulés par la voie Toll, tels que le gène codant la cytokine Spätzle (Kemp et al., 2013). Ces résultats sont également cohérents avec ceux obtenus chez *D. melanogaster* $Dcr2^{[L811fsX]}$, où on retrouve une stimulation de certains gènes de la voie Toll lors de l'infection par SINV dans les carcasses (données non montrées ; *toll* : log2 *fold change* = 2,96 ; *p-value* = 3,87e-33 ; *PGRP-SA* : log2 *fold change* = 1,83 ; *p-value* = 1,07e-3). Par conséquent, lors de l'infection par DMelSV ou DCV, il serait également intéressant de déterminer l'impact de la voie Toll sur les quantités de transcrits d'ET, en utilisant des mutants de drosophiles, tels que pour les gènes *toll* ou *späztle*.

Par la suite, l'objectif principal serait d'étudier l'impact de ces modulations d'activité des ET sur les taux de transposition sous-jacents, lors d'infections virales. En effet, la modulation de la transposition des ET, induisant une variabilité génétique, est impliquée dans les processus de diversité génétique des organismes qui sont d'un intérêt fondamental pour comprendre les interactions hôte-pathogène. Il serait également intéressant d'élargir ces études à d'autres virus naturels de drosophile, responsables d'infections chroniques. Enfin, ces résultats révèlent un impact des infections virales sur l'activité des ET qui semble dépendre du virus, et qui peut être en partie expliqué par le fait que les virus n'induisent pas les mêmes voies de réponse immunitaire. De plus, on peut penser que l'impact de l'infection virale (par exemple SINV) sur les quantités de petits ARN interférents dirigés contre les ET dépend aussi partiellement de l'espèce de drosophile, comme illustré par les différences de production des petits ARN interférents ciblant les ET dans les ovaires entre *D. simulans* Makindu et *D. melanogaster* w^{1118}.

DONNEES SUPPLEMENTAIRES (II)

Données supplémentaires (II)

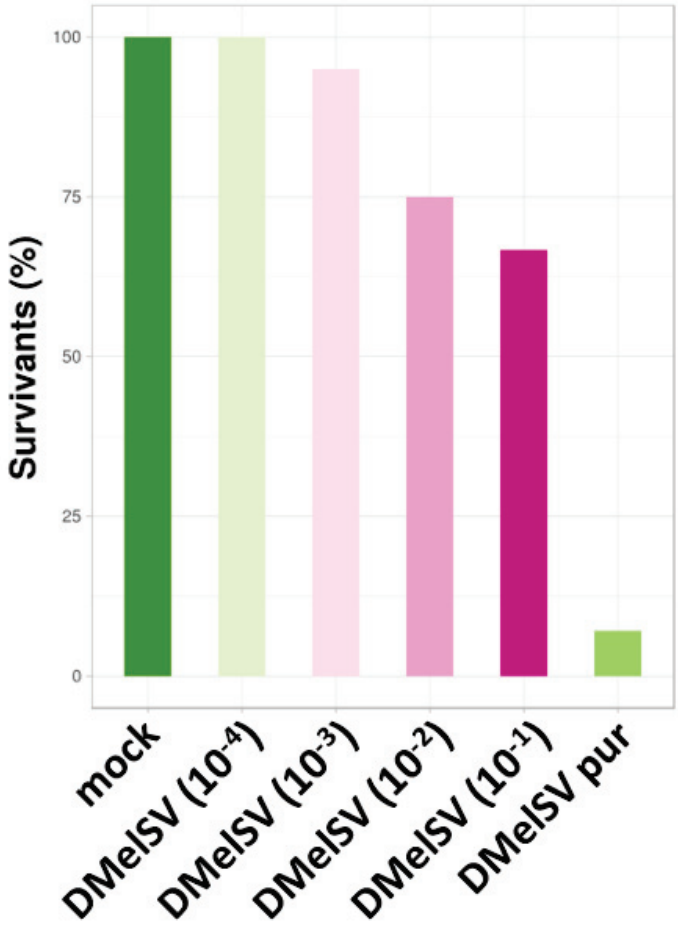


Figure S27. Titration *in vivo* de DMelSV chez *D.* melanogaster. Les drosophiles ont été infectées avec un volume d'injection de 207 nl à partir du stock de DMelSV dilué au 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} ou pur (non dilué) en comparaison avec la condition non infectée (mock). Les drosophiles vivantes ont été quantifiées au jour 14 après l'infection. n = 1 expérience (15 à 20 drosophiles par condition).

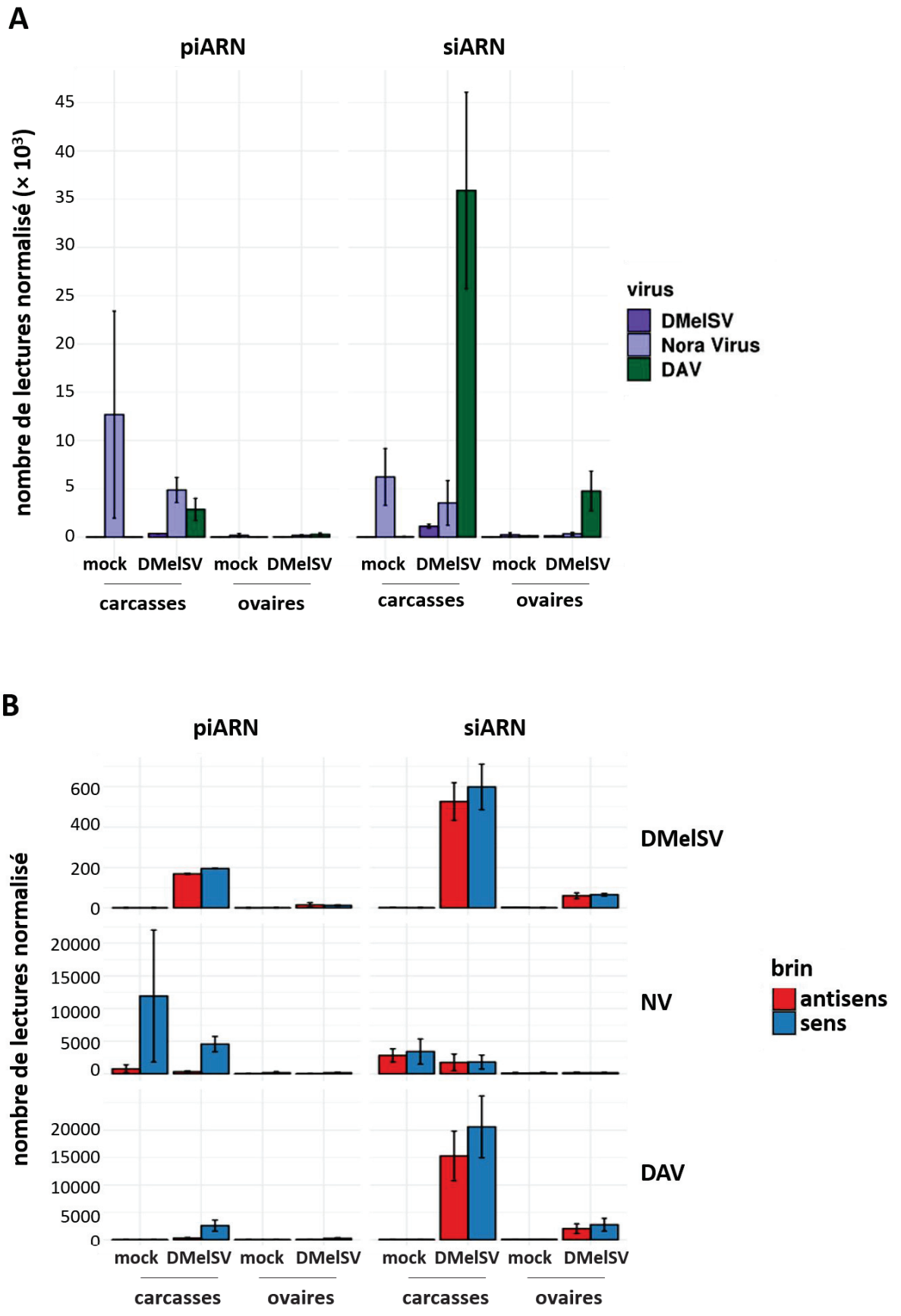


Figure S28. Analyse des siARN et des piARN viraux chez *D. simulans* Makindu. (A) Nombre total de lectures des transcrits des petits ARN interférents viraux (ciblant DMelSV, NV et DAV) de 21 nucléotides (siARN) et de 23-30

nucléotides (piARN) dans les carcasses et les ovaires des Makindu. **(B)** Nombre de lectures des brins sens (bleu) et antisens (rouge) des petits ARN interférents viraux (ciblant DMelSV, NV et DAV) de 21 nucléotides (siARN) et de 23-30 nucléotides (piARN) dans les carcasses et les ovaires des Makindu. L'ensemble des comptes des lectures a été normalisé par le nombre de lectures des miARN. n = 2 expériences indépendantes (50 tissus par échantillon).

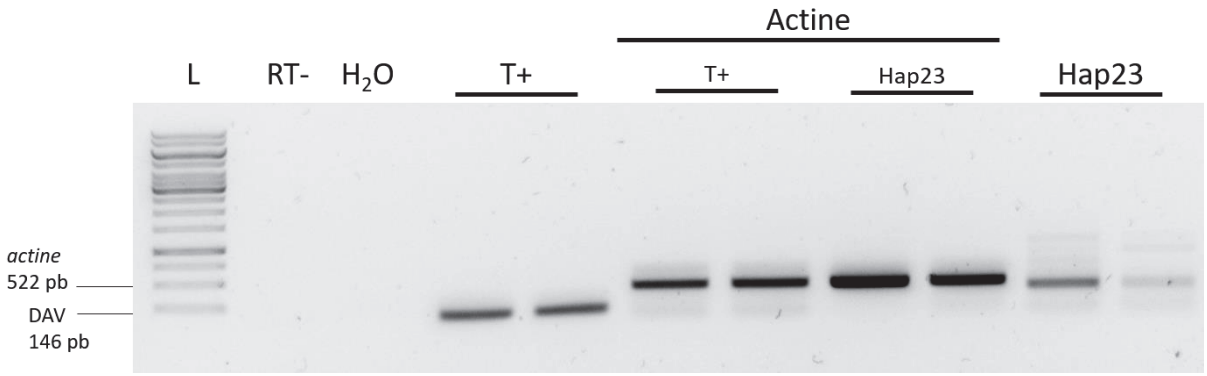


Figure S29. Absence de détection des transcrits de DAV dans les drosophiles chroniquement infectées par DMelSV. Détection de DAV par RT-PCR dans les drosophiles chroniquement infectées par DMelSV (Hap23), en utilisant des amorces spécifiques de DAV. Témoin d'ARN par détection du gène de ménage *actine*, pour les deux témoins positifs de détection de DAV (cellules embryonaires de drosophiles infectées par DAV) et pour les drosophiles infectées par DMelSV Hap23. Les témoins négatifs correspondent à la condition RT- (témoin négatif de transcription à partir des drosophiles infectées par Hap23 et au témoin de PCR (H_2O). Résultats représentatifs de 2 expériences ; 5 drosophiles par échantillon d'ARN.

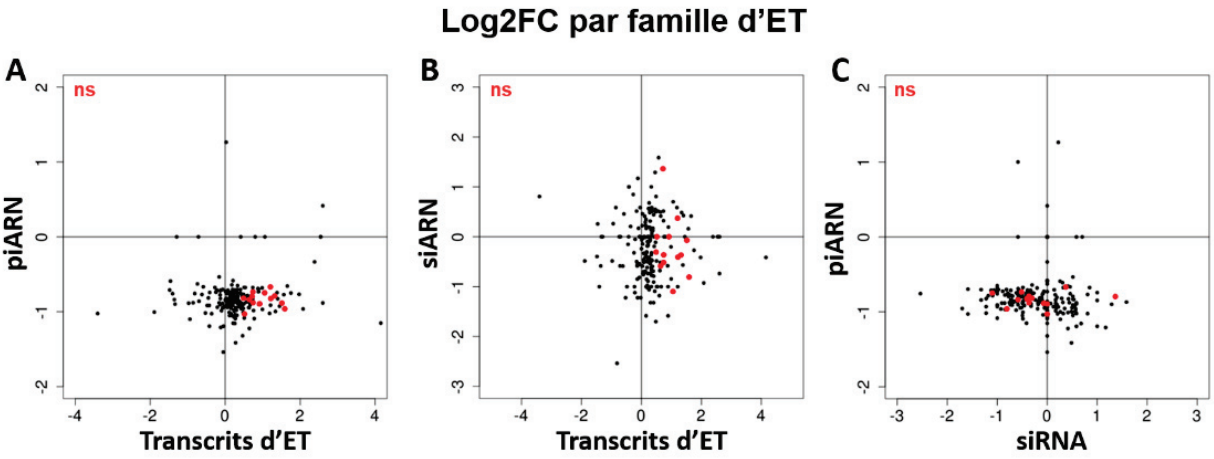


Figure S30. L'augmentation de l'abondance des transcrits d'ET n'est pas significativement dépendante de la diminution des productions de piARN et siARN ciblant les ET, lors de l'infection par DMelSV, dans les carcasses. Chaque point représente le log2 du ratio des comptes par famille d'ET entre la condition infectée par SINV et la condition non infectée (mock). **(A)** Dépendance entre les transcrits d'ET et les piARN. **(B)** Dépendance entre les transcrits d'ET et les siARN. **(C)** Dépendance entre les piARN et les siARN. Le test exact de Fisher a été utilisé pour les analyses statistiques (*Fisher exact test ; ns : non significatif ; *P* < 0,05 ; ***P* < 0,001 ; ****P* < 0,001). n = 2 expériences indépendantes.

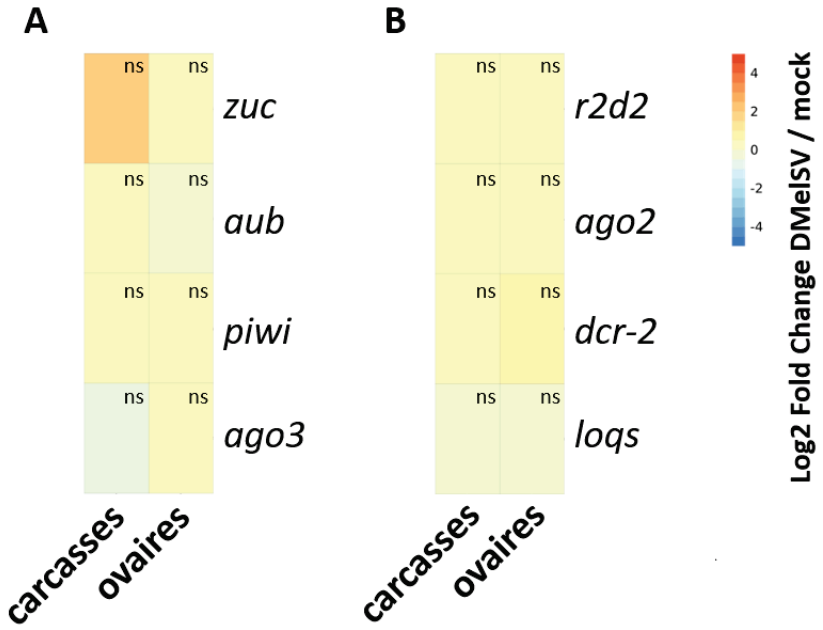


Figure S31. Absence de modulation significative de la transcription des gènes impliqués dans les voies piARN (A) et siARN (B) chez Makindu, lors de l'infection par DMelSV. Les *heatmaps* représentent le log2 *fold change* des comptes des gènes, entre la condition infectée par DMelSV et la condition non infectée (mock). Une augmentation des quantités de transcrits de ces gènes est visualisée dans la gamme de couleur allant du jaune au rouge alors qu'une diminution est représentée dans la gamme de couleur allant du vert clair au bleu. La significativité de la modulation est établie en comparant les comptes de la condition infectée par rapport aux comptes de la condition mock par DESeq2 (*test wald*). *ns : non significatif ;* (**p-value* < 0,05) . Légende : *zuc* (*zucchini*), *aub* (*aubergine*), *piwi* (*P-element induced wimpy testis*), *ago3* (*argonaute 3*), *r2d2* (*r2d2*), *ago2* (*argonaute 2*), *dcr-2* (*dicer-2*), *loqs* (*loquacious*). n = 2 expériences indépendantes (30 tissus par échantillon de RNA-seq).

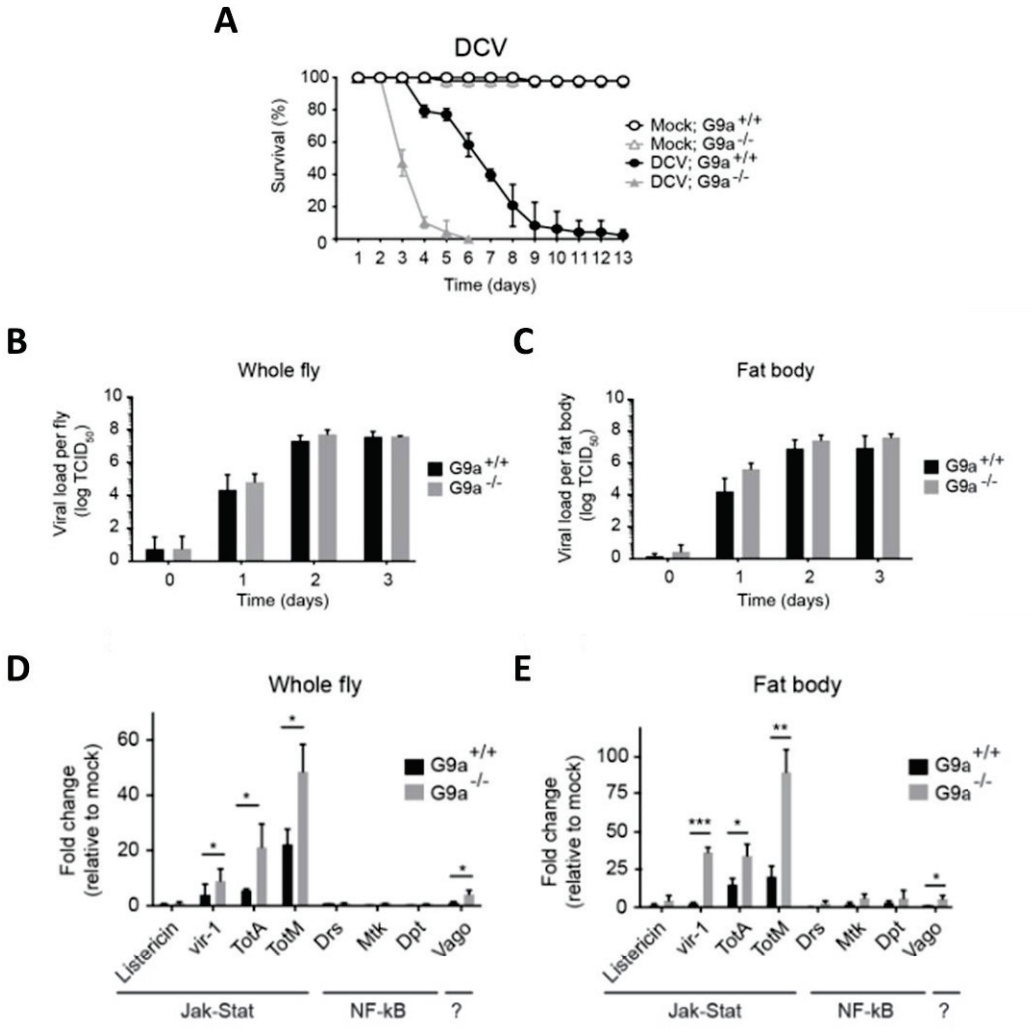


Figure S32. Le régulateur épigénétique G9a est nécessaire à la tolérance des drosophiles lors de l'infection par DCV (Merkling et al., 2015). Comparaison entre la condition infectée par DCV et la condition non infectée (mock) et comparaison entre les drosophiles sauvages ($G9a^{+/+}$) et les mutants *G9a* ($G9a^{-/-}$). **(A)** Courbe de survie des drosophiles sauvages et des mutants *G9a* lors de l'infection par DCV, en comparaison avec la condition non infectée (mock). **(B-C)** Titrations virales en TCID50/ml des drosophiles entières (C) et des corps gras (D). **(D-E)** Expression des gènes induits à 24 heures après infection par DCV, par RT-PCR quantitative, à partir des drosophiles entières (D) et des corps gras (E). n = 3 expériences indépendantes ; **P* < 0,05*; ** P* < 0,01*; *** P* < 0,001 *(Student's t-test).* Données de (Merkling et al., 2015)**.**

DISCUSSION GENERALE

DISCUSSION GENERALE

Nous avons démontré pour la première fois les effets des infections virales sur l'activité transcriptionnelle des ET. Par la suite, il sera primordial de déterminer si les taux de transposition des ET sont impactés par ces variations de quantité de transcrits d'ET lors de l'infection virale. Ces résultats sont d'une importance fondamentale en génomique évolutive et génétique des insectes, incluant le fait que les infections virales peuvent moduler les taux de mutations à travers la modulation de l'activité des ET, et ainsi la vitesse d'évolution des génomes et de diversification génétique des organismes.

Par ailleurs, les modulations de l'activité des ET semblent être spécifiques du virus, et en particulier des mécanismes de réponse immunitaire mis en place lors de l'infection. En effet, dans le cas de SINV (génome ARN simple brin +), la diffusion systémique de siARN viraux par l'intermédiaire de la voie de l'import de l'ARN double brin (Saleh et al., 2009, 2006) joue un rôle dans les modulations de l'activité des ET observées dans les carcasses des *D. melanogaster* w^{1118}, incluant une diminution des quantités de transcrits d'ET associée à une augmentation de la production de piARN et siARN ciblant les ET. Ces résultats ont été retrouvés au niveau des carcasses lors de l'infection par SINV de *D. simulans* Makindu, suggérant un effet non spécifique de l'espèce de drosophile. A l'inverse, dans le cas des virus DMelSV (génome ARN simple brin -) et DCV (génome ARN simple brin +), on a observé une augmentation des quantités de transcrits d'ET qui n'est cependant pas associée de façon significative à la diminution de la production des petits ARN interférents dans le cas de DMelSV (absence de données de Small RNA-seq pour les analyses concernant DCV), suggérant l'implication d'une autre voie de régulation dans la modulation de l'activité des ET lors de l'infection virale. Ainsi, en utilisant deux virus naturels de drosophile, l'effet de la réplication virale sur l'activité des ET est similaire. Cependant, l'infection par un arbovirus (SINV), n'infectant pas naturellement la drosophile, engendre des effets différents sur les quantités de transcrits des ET. Cette divergence pourrait s'expliquer par l'importante induction de la machinerie d'ARN interférence lors de l'infection par SINV et par la diffusion systémique des siARN dans l'organisme des drosophiles infectées (Saleh et al., 2009). En effet, lors des infections par DMelSV ou DCV, la stimulation des voies d'ARN interférence semble atténuée ; mais d'autres voies de l'immunité innée, telles que la voie Toll, semblent être davantage dérégulées (Carré-Mlouka et al., 2007; Ferreira et al., 2014; Kemp et al., 2013). La voie Toll est importante pour la survie des drosophiles *D. melanogaster* infectées oralement par DCV (Ferreira et al., 2014), incluant une augmentation de l'expression des gènes régulés par la voie Toll, tels que le gène codant la cytokine Spätzle (Kemp et al., 2013). Concernant DMelSV, en plus de la surexpression de gènes impliqués directement ou indirectement dans la voie Toll (Carpenter et al., 2009; Tsai et al., 2008), une étude a permis l'identification du gène de résistance principal de DMelSV, nommé *ref(2)P*, un composant de la voie de signalisation Toll (Avila et al., 2002; Carré-Mlouka et al., 2007). A l'inverse, la réplication de SINV n'est pas

impactée dans les mutants de *D. melanogaster* de la voie Toll (*dif*$^{-/-}$), excluant un contrôle de SINV par la voie Toll chez la drosophile (Avadhanula et al., 2009).

De plus, la comparaison des productions de siARN viraux entre SINV et DMelSV révèle une quantité inférieure de siARN lors de l'infection par DMelSV. Contrairement à l'infection par les virus DMelSV et DCV (Ferreira et al., 2014; Tsai et al., 2008), aucun gène impliqué dans la voie Toll n'est significativement modulé lors de l'infection par SINV. En outre, ces virus présentent des variations au niveau de leur réponse virale, de leur tropisme et de leur pathogenèse chez la drosophile (Gomariz-Zilber et al., 1995; Kemp et al., 2013; Tsai et al., 2008). Aucun suppresseur d'ARNi n'a été rapporté pour SINV et DMelSV (Berry et al., 2009; Li et al., 2002; van Rij et al., 2006) alors que la protéine 1A de DCV est un suppresseur d'ARNi (Fareh et al., 2018; Saleh et al., 2006). Lors de l'infection par les virus SINV et DCV de drosophiles mutées au niveau du gène codant la protéine Dcr-2, la réplication virale de ces virus est exacerbée, conduisant à la mort des drosophiles (Galiana-Arnoux et al., 2006). Aucune donnée n'a été rapportée à propos des interactions du DMelSV avec la machinerie d'ARNi, mis à part l'absence de production de piARN viraux chez la drosophile (Petit et al., 2016). Cette absence de piARN viraux a également été rapportée lors des infections par SINV et DCV (Petit et al., 2016). Ces résultats sont en accord avec nos résultats pour les virus SINV et DMelSV. Il serait intéressant par la suite d'analyser l'impact du suppresseur d'ARNi de DCV sur l'activité des ET, notamment en infectant des drosophiles avec le virus sauvage ou le virus muté pour la séquence codante de la protéine 1A de DCV (van Rij et al., 2006). De plus, l'analyse des transcrits d'ET et des petits ARN interférents chez des drosophiles mutées pour un gène de la voie Toll, tels que les gènes *toll* et *dif* (Avadhanula et al., 2009), permettrait de préciser l'hypothèse de son implication dans la modulation des quantités de transcrits d'ET observée lors de l'infection virale par DMelSV et DCV. Par ailleurs, il est intéressant de discuter le fait d'observer une modulation plus importante de l'activité des ET dans le corps gras par rapport aux drosophiles entières lors de l'infection par DCV. En effet, le corps gras représente une cible spécifique de réplication de DCV (Merkling et al., 2015). Par conséquent, il serait particulièrement pertinent d'analyser la voie Toll dans ce tissu et d'utiliser des mutants de drosophiles invalidant cette voie pour déterminer son rôle dans la modulation de l'activité des ET lors de l'infection virale.

Par ailleurs, au niveau des ovaires, alors qu'à l'inverse de SINV, DMelSV semble s'y répliquer, on observe tout de même peu d'impact sur la transcription des ET et la production des petits ARN interférents ciblant les ET. De plus, il n'y a pas de transmission verticale de DCV rapportée dans la littérature. Aucune modulation significative des gènes impliqués dans la voie Toll n'a été observée dans nos analyses en ovaires (données non montrées). Dans cet organe, la voie de l'import de l'ARN double brin et la voie des siARN ne semblent pas être impliquées dans la faible modulation de l'activité des ET observée chez *D. melanogaster*. Il serait intéressant de déterminer le rôle d'autres voies de réponse immunitaire, telles que les voies Toll, Jak/STAT et Imd dans la modulation des quantités de petits ARN interférents ciblant les ET, résultant en une quasi absence de modulation de l'activité des ET (Ullastres et al., 2019). En effet, en

fonction des virus ces voies sont plus ou moins impliquées dans le contrôle de l'infection virale. Comme mentionné ci-dessus, la voie Toll semble être un candidat potentiel pour l'étude de la modulation de l'activité des ET et des petits ARN interférents chez *Drosophila* lors de l'infection par DMelSV et par DCV (Kemp et al., 2013; Tsai et al., 2008). L'activation de la voie Jak/STAT par DCV a également été décrite chez *Drosophila*, incluant une réplication virale et une sensibilité plus élevées chez les drosophiles mutées pour le gène *hopscotch* (Agaisse and Perrimon, 2004; Dostert et al., 2005). Concernant SINV, l'invalidation de la voie Imd et de la voie Jak/STAT conduit à une augmentation de la réplication virale chez les drosophiles mutées pour les gènes codant respectivement la protéine Relish (Imd) et STAT (Jak/STAT) (Avadhanula et al., 2009; Sansone et al., 2015). Ainsi, l'augmentation de la réplication virale dans les drosophiles mutées pour les gènes codant des effecteurs impliqués dans ces voies de réponse antivirale peut éventuellement interférer avec les quantités de petits ARN interférents, et par conséquent sur le maintien de l'activité des ET. En effet, les gonades présentent une plus forte pression de sélection que dans la lignée somatique, permettant d'éliminer les mutations dans la lignée germinale et pouvant favoriser une limitation des modulations de l'activité des ET.

L'ensemble de ces résultats révèle un impact des infections virales sur l'activité des ET qui semble virus-spécifique et notamment dépendante des voies de réponse immunitaire antivirales induites lors de l'infection virale chez *Drosophila*.

Dans mes travaux de thèse, l'ensemble des drosophiles étudiées a été infecté expérimentalement par injection intrathoracique, incluant nos modèles viraux SINV, DMelSV et les infections par DCV, réalisées avec le même mode d'infection par Merkling et collaborateurs (Merkling et al., 2015). Ce mode d'administration virale peut être discuté du fait qu'il engendre une réplication virale systémique qui contourne les barrières de protection naturelle de l'hôte (Mondotte and Saleh, 2018). Pour autant, l'infection de drosophiles par SINV est communément réalisée par injection intrathrocique (Avadhanula et al., 2009; Z. Huang et al., 2013; Lamiable et al., 2016b; Petit et al., 2016; Saleh et al., 2009; Tassetto et al., 2017) mais n'exclut pas d'autres modes d'infection telle que l'ingestion orale (Sansone et al., 2015). Cependant, contrairement à ce qui est décrit dans la littérature, l'infection par DMelSV a également été réalisée par injection intrathoracique et non en utilisant une lignée de *D. melanogaster* infectée de façon persistante (Carré-Mlouka et al., 2007; Petit et al., 2016; Tsai et al., 2008). En effet, pour les études concernant DMelSV nous avons choisi d'étudier *D. simulans* Makindu, qui n'est pas naturellement infectée par DMelSV. Par ailleurs, l'infection de drosophiles par DCV inclut différents modes d'infection, notamment l'ingestion orale (Ferreira et al., 2014; Sansone et al., 2015) et l'injection intrathoracique (Chtarbanova et al., 2014; Lamiable et al., 2016b; Merkling et al., 2015; Petit et al., 2016). DCV se réplique dans l'ensemble du corps de la drosophile lorsqu'il est injecté ou administré oralement, incluant une réplication dans l'intestin, et également dans les hémocytes (Ferreira et al., 2014). Cependant, il a notamment été montré l'importance de plusieurs composants de la voie Toll

dans la tolérance virale des drosophiles lors de l'infection orale (naturelle) par DCV, alors que cela n'a pas été rapporté lors d'infection systémique (infection par piqûre) (Ferreira et al., 2014). Cette résistance virale implique le facteur de transcription de type NFκB, Dorsal. La translocation de Dorsal du cytoplasme jusqu'au noyau dans le corps gras a été montrée lors des deux types d'infection, révélant une activation de la voie Toll indépendamment du mode d'infection, mais une activité antivirale de cette voie uniquement lors de l'infection orale (Ferreira et al., 2014; Mondotte and Saleh, 2018). Ces résultats démontrent l'importance du mode d'infection des drosophiles dans l'étude des voies de l'immunité innée lors d'une infection virale. Par conséquent, il serait intéressant par la suite d'analyser l'impact de l'activité et du taux de transposition des ET lors d'infections naturelles de DMelSV, incluant les drosophiles chroniquement infectées et en utilisant des modes d'infection plus naturels pour DCV et SINV, telle que l'ingestion orale. Cependant, ces derniers modes d'infection ne permettent pas un contrôle et une répétabilité d'infection des drosophiles nécessaires aux études réalisées dans mes travaux de thèse, incluant l'impact d'une infection virale sur l'activité des ET et l'analyse globale des voies d'ARN interférence mises en jeu lors de ces modulations.

De plus, un autre facteur important à prendre en compte lors de l'étude d'interactions hôte-virus est le temps d'infection des drosophiles. Nous avons choisi de réaliser nos expériences et analyses à 6 jours après infection pour SINV et à 8 jours après infection pour DMelSV. Cela a été déterminé en fonction de la cinétique de réplication virale (test CO_2 pour DMelSV) et également pensé dans l'optique d'une éventuelle réplication virale dans les ovaires. Dans la littérature, les expériences impliquent généralement 2 à 6 jours d'infection pour SINV (Lamiable et al., 2016b; Petit et al., 2016; Saleh et al., 2009), tandis que pour l'étude de l'infection par DMelSV des drosophiles chroniquement infectées sont utilisées. Il serait intéressant de déterminer l'impact de l'infection virale sur l'activité des ET à différents temps après infection pour préciser la dynamique de ces modulations et confirmer nos résulats. En effet, on peut s'attendre à ce que les modulations observées soient corrélées notamment avec la cinétique de production des siARN viraux lors de l'infection par SINV. Cette production de petits ARN interférents ou d'autres processus biologiques peuvent également être impactés par la dose infectieuse utilisée pour infecter les drosophiles. Par exemple, lors de l'infection par DCV, le contrôle de l'infection implique le mécanisme de phagocytose à des doses élevées de DCV, alors qu'il n'est pas nécessaire à des doses plus faibles (Lamiable et al., 2016a; Nainu et al., 2015). Par ailleurs, du fait de l'utilisation de *D. melanogaster* infectées de façon persistante dans la littérature, nous avons déterminé le volume d'infection de DMelSV en fonction de la cinétique du test CO_2, indiquant le nombre de drosophiles infectées. Les drosophiles Makindu ont été infectées avec deux doses de SINV (14900 et 2300 PFU par drosophile), révélant une différence importante de réplication virale de SINV entre ces deux doses. La dose à 2300 PFU, induisant une faible réplication virale, n'a pas été retenue pour la suite des expérimentations, mais il serait intéressant d'analyser la production des piARN et

des siARN ciblant SINV et les ET à partir des drosophiles infectées avec des faibles doses infectieuses chez *D. melanogaster*. On peut également envisager que SINV soit capable d'infecter les ovaires à des doses infectieuses inférieures, impliquant une induction plus lente et plus faible des siARN viraux ou d'autres voies de réponse immunitaire.

Mes travaux de thèse révèlent un impact des infections virales sur l'activité des ET et sur les quantités des petits ARN interférents ciblant les ET chez la drosophile. Lors de l'étude d'un arbovirus (SINV), nous avons identifié une interférence entre la voie des piARN et la voie des siARN, impliquée dans la modulation de l'activité des ET. En effet, nos résultats suggèrent l'existence de substrats d'ARN communs entre les siARN et les piARN ciblant les ET, pouvant en partie expliquer les modulations de l'activité des ET observées en somatique lors de l'infection virale. Ces axes de recherche pourraient être élargis aux Arthropodes, et en particulier aux moustiques véhiculant de nombreux arbovirus et représentant une importante menace pour la santé humaine et animale. Néanmoins, les voies d'ARNi présentent certains points de divergence entre la drosophile et le moustique. Concernant la voie des piARN, une expansion des protéines PIWI a été révélée chez les moustiques *Aedes*, *Culex* et *Anopheles* (Campbell et al., 2008; Lewis et al., 2016; Wang et al., 2018). De plus, la production de piARN viraux présentant des propriétés antivirales a été montrée pour de nombreux arbovirus tels que DENV (Wang et al., 2018), RVFV (Dietrich et al., 2017), ZIKV (Saldaña et al., 2017), CHIKV (Morazzani et al., 2012) chez leur hôte arthropode. La présence de séquences virales appelées EVE au sein du génome du moustique a été décrite à partir de rétrovirus mais aussi de virus à ARN, principalement des familles Flavivirus et Rhabdovirus (Cook et al., 2006; Crochu et al., 2004; Fort et al., 2012; Palatini et al., 2017; Suzuki et al., 2017; Whitfield et al., 2017), impliquant leur intégration au niveau de cluster de piARN et permettant la production de piARN viraux dérivés de ces EVE (Palatini et al., 2017). A l'inverse, la production de piARN viraux est controversée chez la drosophile (Petit et al., 2016; Wu et al., 2010) et ces derniers ne semblent pas présenter d'activité antivirale (Petit et al., 2016). Concernant la voie des siARN, le rôle fondamental de la machinerie d'ARNi au niveau de la défense anti-virale est décrit chez la drosophile et le moustique (Galiana-Arnoux et al., 2006; Keene et al., 2004). Cependant, par exemple, le gène *ago2* présente une seule copie chez *D. melanogaster* tandis qu'il est dupliqué chez *Culex pipiens* (Campbell et al., 2008). La production d'endo-siRNA a été révélée chez *Drosophila* (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008) et chez les moustiques (Adelman et al., 2012; Arensburger et al., 2011; Biryukova and Ye, 2015). Néanmoins, ces endo-siARN sont impliqués dans le contrôle somatique des ET chez *Drosophila* tandis que cette caractéristique reste à élucider chez le moustique (Biryukova and Ye, 2015). Par ailleurs, l'activité et le nombre de copies des ET sont très variables entre et au sein des espèces d'insectes, notamment au niveau des génomes des moustiques et des drosophiles (Boulesteix and Biémont, 2005; Petersen et al., 2019). Ainsi, il serait intéressant par la suite d'expérimenter les effets de la variation de l'activité des ET ou du contenu en ET sur la réplication virale. Cet aspect est particulièrement intéressant en ce qui concerne la réplication

d'arbovirus au sein de leurs vecteurs arthropodes, incluant la possibilité d'impact de l'activité des ET sur la compétence vectorielle et ainsi sur la transmission d'arbovirus. Cette perspective de recherche permettrait la comparaison de la réplication virale et de la production de petits ARN interférents en fonction du niveau d'activité des ET. De façon intéressante, il a été précédemment décrit que l'activité de transcriptase inverse de rétrotransposons à LTR est impliquée dans la production de siARN viraux secondaires, participant au renforcement du contrôle de l'infection et permettant ainsi la persistance virale chez la drosophile (Goic et al., 2013) et la tolérance virale lors de l'infection, notamment par CHIKV ou DENV chez les moustiques du genre *Aedes* (Goic et al., 2016). Ainsi, plusieurs éléments suggèrent dès lors l'implication des ET dans l'immunité antivirale chez le moustique (Nag and Kramer, 2017). Cette perspective de recherche pourrait aboutir à l'identification de facteurs modulant la réponse des insectes suite à l'infection virale.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

Abu Bakar F, Ng LFP. 2018. Nonstructural Proteins of Alphavirus—Potential Targets for Drug Development. *Viruses* **10**. doi:10.3390/v10020071

Adelman ZN, Anderson MAE, Liu M, Zhang L, Myles KM. 2012. Sindbis virus induces the production of a novel class of endogenous siRNAs in Aedes aegypti mosquitoes. *Insect Mol Biol* **21**:357–368. doi:10.1111/j.1365-2583.2012.01141.x

Adouchief S, Smura T, Sane J, Vapalahti O, Kurkela S. 2016. Sindbis virus as a human pathogen-epidemiology, clinical picture and pathogenesis. *Rev Med Virol* **26**:221–241. doi:10.1002/rmv.1876

Adrion JR, Song MJ, Schrider DR, Hahn MW, Schaack S. 2017. Genome-Wide Estimates of Transposable Element Insertion and Deletion Rates in Drosophila Melanogaster. *Genome Biol Evol* **9**:1329–1340. doi:10.1093/gbe/evx050

Agaisse H, Perrimon N. 2004. The roles of JAK/STAT signaling in Drosophila immune responses. *Immunol Rev* **198**:72–82.

Agarwal A, Parida M, Dash PK. 2017. Impact of transmission cycles and vector competence on global expansion and emergence of arboviruses. *Reviews in Medical Virology* **27**:e1941. doi:10.1002/rmv.1941

Aiewsakun P, Katzourakis A. 2015. Endogenous viruses: Connecting recent and ancient viral evolution. *Virology*, 60th Anniversary Issue **479–480**:26–37. doi:10.1016/j.virol.2015.02.011

Akbari OS, Antoshechkin I, Amrhein H, Williams B, Diloreto R, Sandler J, Hay BA. 2013. The developmental transcriptome of the mosquito Aedes aegypti, an invasive species and major arbovirus vector. *G3 (Bethesda)* **3**:1493–1509. doi:10.1534/g3.113.006742

Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, Braman V, Chambeyron S, Vieira C. 2013. Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. *EMBO Rep* **14**:458–464. doi:10.1038/embor.2013.38

Akkouche A, Rebollo R, Burlet N, Esnault C, Martinez S, Viginier B, Terzian C, Vieira C, Fablet M. 2012. tirant, a Newly Discovered Active Endogenous Retrovirus in Drosophila simulans. *J Virol* **86**:3675–3681. doi:10.1128/JVI.07146-11

Ambrose RL, Lander GC, Maaty WS, Bothner B, Johnson JE, Johnson KN. 2009. Drosophila A virus is an unusual RNA virus with a T=3 icosahedral core and permuted RNA-dependent RNA polymerase. *J Gen Virol* **90**:2191–2200. doi:10.1099/vir.0.012104-0

Aminetzach YT, Macpherson JM, Petrov DA. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in Drosophila. *Science* **309**:764–767. doi:10.1126/science.1112699

Antoniewski C. 2014. Computing siRNA and piRNA overlap signatures. *Methods Mol Biol* **1173**:135–146. doi:10.1007/978-1-4939-0931-5_12

Arensburger P, Hice RH, Wright JA, Craig NL, Atkinson PW. 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. doi:10.1186/1471-2164-12-606

Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B, Bartholomay L, Bidwell S, Caler E, Camara F, Campbell CL, Campbell KS, Casola C, Castro MT, Chandramouliswaran I, Chapman SB, Christley S, Costas J, Eisenstadt E, Feschotte C, Fraser-Liggett C, Guigo R, Haas B, Hammond M, Hansson BS, Hemingway J, Hill SR, Howarth C, Ignell R, Kennedy RC, Kodira CD, Lobo NF, Mao C, Mayhew G, Michel K, Mori A, Liu N, Naveira H, Nene V, Nguyen N, Pearson MD, Pritham EJ, Puiu D, Qi Y, Ranson H, Ribeiro JMC, Roberston HM, Severson DW, Shumway M, Stanke M, Strausberg RL, Sun C, Sutton G, Tu ZJ, Tubio JMC, Unger MF, Vanlandingham DL, Vilella AJ, White O, White JR, Wondji CS, Wortman J, Zdobnov EM, Birren B, Christensen BM, Collins FH, Cornel A, Dimopoulos G, Hannick LI, Higgs S, Lanzaro GC, Lawson D, Lee NH, Muskavitch MAT, Raikhel AS, Atkinson PW. 2010. Sequencing of Culex

quinquefasciatus establishes a platform for mosquito comparative genomics. *Science* **330**:86–88. doi:10.1126/science.1191864
Arnheiter H, Davis NL, Wertz G, Schubert M, Lazzarini RA. 1985. Role of the nucleocapsid protein in regulating vesicular stomatitis virus RNA synthesis. *Cell* **41**:259–267.
Arnot CJ, Gay NJ, Gangloff M. 2010. Molecular mechanism that induces activation of Spätzle, the ligand for the Drosophila Toll receptor. *J Biol Chem* **285**:19502–19509. doi:10.1074/jbc.M109.098186
Asgari S. 2013. MicroRNA functions in insects. *Insect Biochemistry and Molecular Biology* **43**:388–397. doi:10.1016/j.ibmb.2012.10.005
Avadhanula V, Weasner BP, Hardy GG, Kumar JP, Hardy RW. 2009. A Novel System for the Launch of Alphavirus RNA Synthesis Reveals a Role for the Imd Pathway in Arthropod Antiviral Response. *PLOS Pathogens* **5**:e1000582. doi:10.1371/journal.ppat.1000582
Avila A, Silverman N, Diaz-Meco MT, Moscat J. 2002. The Drosophila Atypical Protein Kinase C-Ref(2)P Complex Constitutes a Conserved Module for Signaling in the Toll Pathway. *Molecular and Cellular Biology* **22**:8787–8795. doi:10.1128/MCB.22.24.8787-8795.2002
Ayres JS, Schneider DS. 2012. Tolerance of infections. *Annu Rev Immunol* **30**:271–294. doi:10.1146/annurev-immunol-020711-075030
Ballinger MJ, Bruenn JA, Taylor DJ. 2012. Phylogeny, integration and expression of sigma virus-like genes in Drosophila. *Mol Phylogenet Evol* **65**:251–258. doi:10.1016/j.ympev.2012.06.008
Barckmann B, El-Barouk M, Pélisson A, Mugat B, Li B, Franckhauser C, Fiston Lavier A-S, Mirouze M, Fablet M, Chambeyron S. 2018. The somatic piRNA pathway controls germline transposition over generations. *Nucleic Acids Res* **46**:9524–9536. doi:10.1093/nar/gky761
Barrón MG, Fiston-Lavier A-S, Petrov DA, González J. 2014. Population genomics of transposable elements in Drosophila. *Annu Rev Genet* **48**:561–581. doi:10.1146/annurev-genet-120213-092359
Beek M van den, Silva B da, Pouch J, Chaouche M el amine A, Carré C, Antoniewski C. 2018. Dual-layer transposon repression in heads of Drosophila melanogaster. *bioRxiv* 267039. doi:10.1101/267039
Beerntsen BT, James AA, Christensen BM. 2000. Genetics of Mosquito Vector Competence. *Microbiol Mol Biol Rev* **64**:115–137.
Belshaw R, Pereira V, Katzourakis A, Talbot G, Paces J, Burt A, Tristem M. 2004. Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci USA* **101**:4894–4899. doi:10.1073/pnas.0307800101
Bennett KE, Flick D, Fleming KH, Jochim R, Beaty BJ, Black WC. 2005. Quantitative Trait Loci That Control Dengue-2 Virus Dissemination in the Mosquito Aedes aegypti. *Genetics* **170**:185–194. doi:10.1534/genetics.104.035634
Bergman CM, Bensasson D. 2007. Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in Drosophila melanogaster. *Proc Natl Acad Sci USA* **104**:11340–11345. doi:10.1073/pnas.0702552104
Berry B, Deddouche S, Kirschner D, Imler J-L, Antoniewski C. 2009. Viral Suppressors of RNA Silencing Hinder Exogenous and Endogenous Small RNA Pathways in Drosophila. *PLoS ONE* **4**. doi:10.1371/journal.pone.0005866
Bhattacharya T, Newton ILG, Hardy RW. 2017. Wolbachia elevates host methyltransferase expression to block an RNA virus early during infection. *PLoS Pathog* **13**:e1006427. doi:10.1371/journal.ppat.1006427
Biémont C. 2010. A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* **186**:1085–1093. doi:10.1534/genetics.110.124180
Biémont C. 2008. Genome size evolution: Within-species variation in genome size. *Heredity* **101**:297–298. doi:10.1038/hdy.2008.80
Biémont C, Vieira C. 2006. Genetics: junk DNA as an evolutionary force. *Nature* **443**:521–524. doi:10.1038/443521a

Biessmann H, Carter SB, Mason JM. 1990. Chromosome ends in Drosophila without telomeric DNA sequences. *Proc Natl Acad Sci U S A* **87**:1758–1761.
Biryukova I, Ye T. 2015. Endogenous siRNAs and piRNAs derived from transposable elements and genes in the malaria vector mosquito Anopheles gambiae. *BMC Genomics* **16**. doi:10.1186/s12864-015-1436-1
Bischoff V, Vignal C, Boneca IG, Michel T, Hoffmann JA, Royet J. 2004. Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. *Nature Immunology* **5**:1175. doi:10.1038/ni1123
Bonatti S, Migliaccio G, Blobel G, Walter P. 1984. Role of signal recognition particle in the membrane assembly of Sindbis viral glycoproteins. *Eur J Biochem* **140**:499–502. doi:10.1111/j.1432-1033.1984.tb08130.x
Bosco G, Campbell P, Leiva-Neto JT, Markow TA. 2007. Analysis of Drosophila species genome size and satellite DNA content reveals significant differences among strains as well as between species. *Genetics* **177**:1277–1290. doi:10.1534/genetics.107.075069
Boulesteix M, Biémont C. 2005. Transposable elements in mosquitoes. *Cytogenet Genome Res* **110**:500–509. doi:10.1159/000084983
Boulesteix M, Weiss M, Biémont C. 2006. Differences in genome size between closely related species: the Drosophila melanogaster species subgroup. *Mol Biol Evol* **23**:162–167. doi:10.1093/molbev/msj012
Bowen NJ, McDonald JF. 2001. Drosophila euchromatic LTR retrotransposons are much younger than the host species in which they reside. *Genome Res* **11**:1527–1540. doi:10.1101/gr.164201
Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**:1089–1103. doi:10.1016/j.cell.2007.01.043
Bronkhorst AW, van Cleef KWR, Vodovar N, Ince IA, Blanc H, Vlak JM, Saleh M-C, van Rij RP. 2012. The DNA virus Invertebrate iridescent virus 6 is a target of the Drosophila RNAi machinery. *Proc Natl Acad Sci USA* **109**:E3604-3613. doi:10.1073/pnas.1207213109
Bronkhorst AW, van Rij RP. 2014. The long and short of antiviral defense: small RNA-based immunity in insects. *Current Opinion in Virology* **7**:19–28. doi:10.1016/j.coviro.2014.03.010
Broughton JP, Lovci MT, Huang JL, Yeo GW, Pasquinelli AE. 2016. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Mol Cell* **64**:320–333. doi:10.1016/j.molcel.2016.09.004
Brower-Toland B, Findley SD, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin SCR, Lin H. 2007. Drosophila PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev* **21**:2300–2311. doi:10.1101/gad.1564307
Brown RS, Wan JJ, Kielian M. 2018. The Alphavirus Exit Pathway: What We Know and What We Wish We Knew. *Viruses* **10**. doi:10.3390/v10020089
Brown S, Hu N, Hombría JC. 2001. Identification of the first invertebrate interleukin JAK/STAT receptor, the Drosophila gene domeless. *Curr Biol* **11**:1700–1705.
Brun G, Plus N. 1980. The viruses of Drosophila. *Academic Press* Pp. 625-702 in A. ASHBURNER and T. R. E WRIGHT, eds.
Campbell CL, Black WC, Hess AM, Foy BD. 2008. Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. *BMC Genomics* **9**:425. doi:10.1186/1471-2164-9-425
Capy P, Gibert P. 2004. Drosophila melanogaster, Drosophila simulans: so similar yet so different. *Genetica* **120**:5–16.
Capy P, Langin T, Higuet D, Maurer P, Bazin C. 1997. Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? *Genetica* **100**:63–72.
Carpenter J, Hutter S, Baines JF, Roller J, Saminadin-Peter SS, Parsch J, Jiggins FM. 2009. The Transcriptional Response of Drosophila melanogaster to Infection with the Sigma Virus (Rhabdoviridae). *PLoS One* **4**. doi:10.1371/journal.pone.0006838

Carpenter JA, Obbard DJ, Maside X, Jiggins FM. 2007. The recent spread of a vertically transmitted virus through populations of Drosophila melanogaster. *Mol Ecol* **16**:3947–3954. doi:10.1111/j.1365-294X.2007.03460.x

Carré-Mlouka A, Gaumer S, Gay P, Petitjean AM, Coulondre C, Dru P, Bras F, Dezélée S, Contamine D. 2007. Control of sigma virus multiplication by the ref(2)P gene of Drosophila melanogaster: an in vivo study of the PB1 domain of Ref(2)P. *Genetics* **176**:409–419. doi:10.1534/genetics.106.063826

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. *Mol Ecol* **22**:1503–1517. doi:10.1111/mec.12170

Chao J-L, Tsai Y-C, Chiu S-J, Sun YH. 2004. Localized Notch signal acts through eyg and upd to promote global growth in Drosophila eye. *Development* **131**:3839–3847. doi:10.1242/dev.01258

Charroux B, Royet J. 2010. Drosophila immune response: From systemic antimicrobial peptide production in fat body cells to local defense in the intestinal tract. *Fly* **4**:40–47. doi:10.4161/fly.4.1.10810

Chen X-G, Jiang Xuanting, Gu J, Xu M, Wu Y, Deng Y, Zhang C, Bonizzoni M, Dermauw W, Vontas J, Armbruster P, Huang X, Yang Y, Zhang H, He W, Peng H, Liu Y, Wu K, Chen J, Lirakis M, Topalis P, Leeuwen TV, Hall AB, Jiang Xiaofang, Thorpe C, Mueller RL, Sun C, Waterhouse RM, Yan G, Tu ZJ, Fang X, James AA. 2015. Genome sequence of the Asian Tiger mosquito, Aedes albopictus, reveals insights into its biology, genetics, and evolution. *PNAS* **112**:E5907–E5915. doi:10.1073/pnas.1516410112

Chengula AA, Mdegela RH, Kasanga CJ. 2013. Socio-economic impact of Rift Valley fever to pastoralists and agro pastoralists in Arusha, Manyara and Morogoro regions in Tanzania. *SpringerPlus* **2**. doi:10.1186/2193-1801-2-549

Chotkowski HL, Ciota AT, Jia Y, Puig-Basagoiti F, Kramer LD, Shi P-Y, Glaser RL. 2008. West Nile virus infection of Drosophila melanogaster induces a protective RNAi response. *Virology* **377**:197–206.

Chtarbanova S, Lamiable O, Lee K-Z, Galiana D, Troxler L, Meignin C, Hetru C, Hoffmann JA, Daeffler L, Imler J-L. 2014. Drosophila C Virus Systemic Infection Leads to Intestinal Obstruction. *J Virol* **88**:14057–14069. doi:10.1128/JVI.02320-14

Chung W-J, Okamura K, Martin R, Lai EC. 2008. Endogenous RNA Interference Provides a Somatic Defense against Drosophila Transposons. *Current Biology* **18**:795–802. doi:10.1016/j.cub.2008.05.006

Chuong EB, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics* **18**:71–86. doi:10.1038/nrg.2016.139

Claycomb JM. 2014. Ancient Endo-siRNA Pathways Reveal New Tricks. *Current Biology* **24**:R703–R715. doi:10.1016/j.cub.2014.06.009

Cook S, Bennett SN, Holmes EC, De Chesse R, Moureau G, de Lamballerie X. 2006. Isolation of a new strain of the flavivirus cell fusing agent virus in a natural mosquito population from Puerto Rico. *J Gen Virol* **87**:735–748. doi:10.1099/vir.0.81475-0

Crochu S, Cook S, Attoui H, Charrel RN, De Chesse R, Belhouchet M, Lemasson J-J, de Micco P, de Lamballerie X. 2004. Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. mosquitoes. *Journal of General Virology* **85**:1971–1980. doi:10.1099/vir.0.79850-0

Czech B, Hannon GJ. 2016. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem Sci* **41**:324–337. doi:10.1016/j.tibs.2015.12.008

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, Hannon GJ, Brennecke J. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**:798–802. doi:10.1038/nature07007

Czech B, Preall JB, McGinn J, Hannon GJ. 2013. A Transcriptome-wide RNAi Screen in the Drosophila Ovary Reveals Factors of the Germline piRNA Pathway. *Molecular Cell* **50**:749–761. doi:10.1016/j.molcel.2013.04.007

Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D, Batterham P, Feyereisen R, Wilson TG, ffrench-Constant RH. 2002. A single p450 allele associated with insecticide resistance in Drosophila. *Science* **297**:2253–2256. doi:10.1126/science.1074170

Dasgupta R, Free HM, Zietlow SL, Paskewitz SM, Aksoy S, Shi L, Fuchs J, Hu C, Christensen BM. 2007. Replication of flock house virus in three genera of medically important insects. *J Med Entomol* **44**:102–110. doi:10.1603/0022-2585(2007)44[102:rofhvi]2.0.co;2

Davis MM, Engström Y. 2012. Immune response in the barrier epithelia: lessons from the fruit fly Drosophila melanogaster. *J Innate Immun* **4**:273–283. doi:10.1159/000332947

Dävring L, Sunner M. 1973. Female meiosis and embryonic mitosis in Drosophila melanogaster. I. Meiosis and fertilization. *Hereditas* **73**:51–64.

de Curtis I, Simons K. 1988. Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells. *Proc Natl Acad Sci USA* **85**:8052–8056. doi:10.1073/pnas.85.21.8052

de Groot RJ, Hardy WR, Shirako Y, Strauss JH. 1990. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *EMBO J* **9**:2631–2638.

de la Chaux N, Wagner A. 2011. BEL/Pao retrotransposons in metazoan genomes. *BMC Evol Biol* **11**:154. doi:10.1186/1471-2148-11-154

Deddouche S, Matt N, Budd A, Mueller S, Kemp C, Galiana-Arnoux D, Dostert C, Antoniewski C, Hoffmann JA, Imler J-L. 2008. The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activity in drosophila. *Nat Immunol* **9**:1425–1432. doi:10.1038/ni.1664

Delorme-Axford E, Klionsky DJ. 2019. Inflammatory-dependent Sting activation induces antiviral autophagy to limit zika virus in the Drosophila brain. *Autophagy* **15**:1–3. doi:10.1080/15548627.2018.1539585

Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**:231–235. doi:10.1038/nature03049

Desset S, Conte C, Dimitri P, Calco V, Dastugue B, Vaury C. 1999. Mobilization of two retroelements, ZAM and Idefix, in a novel unstable line of Drosophila melanogaster. *Mol Biol Evol* **16**:54–66. doi:10.1093/oxfordjournals.molbev.a026038

Desset S, Meignin C, Dastugue B, Vaury C. 2003. COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. *Genetics* **164**:501–509.

Dewannieux M, Heidmann T. 2013. Endogenous retroviruses: acquisition, amplification and taming of genome invaders. *Curr Opin Virol* **3**:646–656. doi:10.1016/j.coviro.2013.08.005

Dhileepan K, Azuolas JK, Gibson CA. 1996. Evidence of vertical transmission of Ross River and Sindbis viruses (Togaviridae: Alphavirus) by mosquitoes (Diptera: Culicidae) in southeastern Australia. *J Med Entomol* **33**:180–182.

Dietrich I, Jansen S, Fall G, Lorenzen S, Rudolf M, Huber K, Heitmann A, Schicht S, Ndiaye EH, Watson M, Castelli I, Brennan B, Elliott RM, Diallo M, Sall AA, Failloux A-B, Schnettler E, Kohl A, Becker SC. 2017. RNA Interference Restricts Rift Valley Fever Virus in Multiple Insect Systems. *mSphere* **2**. doi:10.1128/mSphere.00090-17

Dietzgen RG, Kondo H, Goodin MM, Kurath G, 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. doi:10.1016/j.virusres.2016.10.010

Dionne M. 2014. Immune-metabolic interaction in Drosophila. *Fly (Austin)* **8**:75–79. doi:10.4161/fly.28113

Dönertas D, Sienski G, Brennecke J. 2013. Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. *Genes Dev* **27**:1693–1705. doi:10.1101/gad.221150.113

Dostert C, Jouanguy E, Irving P, Troxler L, Galiana-Arnoux D, Hetru C, Hoffmann JA, Imler J-L. 2005. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of drosophila. *Nat Immunol* **6**:946–953. doi:10.1038/ni1237

Dowsett AP, Young MW. 1982. Differing levels of dispersed repetitive DNA among closely related species of Drosophila. *Proc Natl Acad Sci USA* **79**:4570–4574.

Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, Pollard DA, Sackton TB, Larracuente AM, Singh ND, Abad JP, Abt DN, Adryan B, Aguade M, Akashi H, Anderson WW, Aquadro CF, Ardell DH, Arguello R, Artieri CG, Barbash DA, Barker D, Barsanti P, Batterham P, Batzoglou S, Begun D, Bhutkar A, Blanco E, Bosak SA, Bradley RK, Brand AD, Brent MR, Brooks AN, Brown RH, Butlin RK, Caggese C, Calvi BR, Bernardo de Carvalho A, Caspi A, Castrezana S, Celniker SE, Chang JL, Chapple C, Chatterji S, Chinwalla A, Civetta A, Clifton SW, Comeron JM, Costello JC, Coyne JA, Daub J, David RG, Delcher AL, Delehaunty K, Do CB, Ebling H, Edwards K, Eickbush T, Evans JD, Filipski A, Findeiss S, Freyhult E, Fulton L, Fulton R, Garcia ACL, Gardiner A, Garfield DA, Garvin BE, Gibson G, Gilbert D, Gnerre S, Godfrey J, Good R, Gotea V, Gravely B, Greenberg AJ, Griffiths-Jones S, Gross S, Guigo R, Gustafson EA, Haerty W, Hahn MW, Halligan DL, Halpern AL, Halter GM, Han MV, Heger A, Hillier L, Hinrichs AS, Holmes I, Hoskins RA, Hubisz MJ, Hultmark D, Huntley MA, Jaffe DB, Jagadeeshan S, Jeck WR, Johnson J, Jones CD, Jordan WC, Karpen GH, Kataoka E, Keightley PD, Kheradpour P, Kirkness EF, Koerich LB, Kristiansen K, Kudrna D, Kulathinal RJ, Kumar S, Kwok R, Lander E, Langley CH, Lapoint R, Lazzaro BP, Lee S-J, Levesque L, Li R, Lin C-F, Lin MF, Lindblad-Toh K, Llopart A, Long M, Low L, Lozovsky E, Lu J, Luo M, Machado CA, Makalowski W, Marzo M, Matsuda M, Matzkin L, McAllister B, McBride CS, McKernan B, McKernan K, Mendez-Lago M, Minx P, Mollenhauer MU, Montooth K, Mount SM, Mu X, Myers E, Negre B, Newfeld S, Nielsen R, Noor MAF, O'Grady P, Pachter L, Papaceit M, Parisi MJ, Parisi M, Parts L, Pedersen JS, Pesole G, Phillippy AM, Ponting CP, Pop M, Porcelli D, Powell JR, Prohaska S, Pruitt K, Puig M, Quesneville H, Ram KR, Rand D, Rasmussen MD, Reed LK, Reenan R, Reily A, Remington KA, Rieger TT, Ritchie MG, Robin C, Rogers Y-H, Rohde C, Rozas J, Rubenfield MJ, Ruiz A, Russo S, Salzberg SL, Sanchez-Gracia A, Saranga DJ, Sato H, Schaeffer SW, Schatz MC, Schlenke T, Schwartz R, Segarra C, Singh RS, Sirot L, Sirota M, Sisneros NB, Smith CD, Smith TF, Spieth J, Stage DE, Stark A, Stephan W, Strausberg RL, Strempel S, Sturgill D, Sutton G, Sutton GG, Tao W, Teichmann S, Tobari YN, Tomimura Y, Tsolas JM, Valente VLS, Venter E, Venter JC, Vicario S, Vieira FG, Vilella AJ, Villasante A, Walenz B, Wang J, Wasserman M, Watts T, Wilson D, Wilson RK, Wing RA, Wolfner MF, Wong A, Wong GK-S, Wu C-I, Wu G, Yamamoto D, Yang H-P, Yang S-P, Yorke JA, Yoshida K, Zdobnov E, Zhang P, Zhang Y, Zimin AV, Baldwin J, Abdouelleil A, Abdulkadir J, Abebe A, Abera B, Abreu J, Acer SC, Aftuck L, Alexander A, An P, Anderson E, Anderson S, Arachi H, Azer M, Bachantsang P, Barry A, Bayul T, Berlin A, Bessette D, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Bourzgui I, Brown A, Cahill P, Channer S, Cheshatsang Y, Chuda L, Citroen M, Collymore A, Cooke P, Costello M, D'Aco K, Daza R, De Haan G, DeGray S, DeMaso C, Dhargay N, Dooley K, Dooley E, Doricent M, Dorje P, Dorjee K, Dupes A, Elong R, Falk J, Farina A, Faro S, Ferguson D, Fisher S, Foley CD, Franke A, Friedrich D, Gadbois L, Gearin G, Gearin CR, Giannoukos G, Goode T, Graham J, Grandbois E, Grewal S, Gyaltsen K, Hafez N, Hagos B, Hall J, Henson C, Hollinger A, Honan T, Huard MD, Hughes L, Hurhula B, Husby ME, Kamat A, Kanga B, Kashin S, Khazanovich D, Kisner P, Lance K, Lara M, Lee W, Lennon N, Letendre F, LeVine R, Lipovsky A, Liu X, Liu J, Liu S, Lokyitsang T, Lokyitsang Y, Lubonja R, Lui A, MacDonald P, Magnisalis V, Maru K, Matthews C, McCusker W, McDonough S, Mehta T, Meldrim J, Meneus L, Mihai O, Mihalev A, Mihova T, Mittelman R, Mlenga V, Montmayeur A, Mulrain L, Navidi A, Naylor J, Negash T, Nguyen T, Nguyen N, Nicol R, Norbu C, Norbu N, Novod N, O'Neill B, Osman S, Markiewicz E, Oyono OL, Patti C, Phunkhang P, Pierre F, Priest M, Raghuraman S, Rege F, Reyes R, Rise C, Rogov P, Ross K, Ryan E, Settipalli S, Shea T, Sherpa N, Shi L, Shih D, Sparrow T, Spaulding J, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Strader C, Tesfaye S, Thomson T, Thoulutsang Y, Thoulutsang D, Topham K, Topping I, Tsamla T, Vassiliev H, Vo A, Wangchuk T, Wangdi T, Weiand M, Wilkinson J, Wilson A, Yadav S, Young G, Yu Q, Zembek L, Zhong D, Zimmer A, Zwirko Z, Jaffe DB, Alvarez P, Brockman W, Butler J, Chin C, Gnerre S, Grabherr M, Kleber M,

Mauceli E, MacCallum I. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**:203–218. doi:10.1038/nature06341

Duffy S. 2018. Why are RNA virus mutation rates so damn high? *PLOS Biology* **16**:e3000003. doi:10.1371/journal.pbio.3000003

Eleftherianos I, Won S, Chtarbanova S, Squiban B, Ocorr K, Bodmer R, Beutler B, Hoffmann JA, Imler J-L. 2011. ATP-sensitive potassium channel (KATP)–dependent regulation of cardiotropic viral infections. *Proc Natl Acad Sci U S A* **108**:12024–12029. doi:10.1073/pnas.1108926108

Fablet M. 2014. Host control of insect endogenous retroviruses: small RNA silencing and immune response. *Viruses* **6**:4447–4464. doi:10.3390/v6114447

Fablet M, Jacquet A, Rebollo R, Haudry A, Rey C, Salces-Ortiz J, Bajad P, Burlet N, Jantsch MF, Guerreiro MPG, Vieira C. 2019. Dynamic Interactions Between the Genome and an Endogenous Retrovirus: Tirant in Drosophila simulans Wild-Type Strains. *G3 (Bethesda)* **9**:855–865. doi:10.1534/g3.118.200789

Fablet M, McDonald JF, Biémont C, Vieira C. 2006. Ongoing loss of the tirant transposable element in natural populations of Drosophila simulans. *Gene* **375**:54–62. doi:10.1016/j.gene.2006.02.008

Fareh M, van Lopik J, Katechis I, Bronkhorst AW, Haagsma AC, van Rij RP, Joo C. 2018. Viral suppressors of RNAi employ a rapid screening mode to discriminate viral RNA from cellular small RNA. *Nucleic Acids Res* **46**:3187–3197. doi:10.1093/nar/gkx1316

Fearns R, Peeples ME, Collins PL. 1997. Increased expression of the N protein of respiratory syncytial virus stimulates minigenome replication but does not alter the balance between the synthesis of mRNA and antigenome. *Virology* **236**:188–201. doi:10.1006/viro.1997.8734

Ferreira ÁG, Naylor H, Esteves SS, Pais IS, Martins NE, Teixeira L. 2014. The Toll-Dorsal Pathway Is Required for Resistance to Viral Oral Infection in Drosophila. *PLOS Pathogens* **10**:e1004507. doi:10.1371/journal.ppat.1004507

Feschotte C, Gilbert C. 2012. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat Rev Genet* **13**:283–296. doi:10.1038/nrg3199

Feschotte C, Jiang N, Wessler SR. 2002. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* **3**:329–341. doi:10.1038/nrg793

Filone CM, Hanna SL, Caino MC, Bambina S, Doms RW, Cherry S. 2010. Rift valley fever virus infection of human cells and insect hosts is promoted by protein kinase C epsilon. *PLoS ONE* **5**:e15483. doi:10.1371/journal.pone.0015483

Fleuriet. 1988. Maintenance of a gereditary virus: the sigma virus in populations of its host, Drosophila melanogaster.

Fleuriet A. 1981. Comparison of various physiological traits in flies (Drosophila melanogaster) of wild origin, infected or uninfected by the hereditary Rhabdovirus sigma. *Arch Virol* **69**:261–272.

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

Francy DB, Jaenson TG, Lundström JO, Schildt EB, Espmark A, Henriksson B, Niklasson B. 1989. Ecologic studies of mosquitoes and birds as hosts of Ockelbo virus in Sweden and isolation of Inkoo and Batai viruses from mosquitoes. *Am J Trop Med Hyg* **41**:355–363.

Franz AWE, Kantor AM, Passarelli AL, Clem RJ. 2015. Tissue Barriers to Arbovirus Infection in Mosquitoes. *Viruses* **7**:3741–3767. doi:10.3390/v7072795

Fu G, Brkić J, Hayder H, Peng C. 2013. MicroRNAs in Human Placental Development and Pregnancy Complications. *Int J Mol Sci* **14**:5519–5544. doi:10.3390/ijms14035519

Fukunaga R, Han BW, Hung J-H, Xu J, Weng Z, Zamore PD. 2012. Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals. *Cell* **151**:533–546. doi:10.1016/j.cell.2012.09.027

Fuller SD. 1987. The T=4 envelope of Sindbis virus is organized by interactions with a complementary T=3 capsid. *Cell* **48**:923–934.

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler J-L. 2006. Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. *Nature Immunology* **7**:590–597. doi:10.1038/ni1335

Gammon DB, Mello CC. 2015. RNA interference-mediated antiviral defense in insects. *Curr Opin Insect Sci* **8**:111–120. doi:10.1016/j.cois.2015.01.006

Ganko EW, Greene CS, Lewis JA, Bhattacharjee V, McDonald JF. 2006. LTR retrotransposon-gene associations in Drosophila melanogaster. *J Mol Evol* **62**:111–120. doi:10.1007/s00239-004-0312-4

Gay P. 1978. [Drosophila genes which intervene in multiplication of sigma virus (author's transl)]. *Mol Gen Genet* **159**:269–283.

Gerdes P, Richardson SR, Mager DL, Faulkner GJ. 2016. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. *Genome Biol* **17**:100. doi:10.1186/s13059-016-0965-5

Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler ELW, Zapp ML, Weng Z, Zamore PD. 2008. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320**:1077–1081. doi:10.1126/science.1157396

Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**:94–108. doi:10.1038/nrg2504

Girardi E, Lefèvre M, Chane-Woon-Ming B, Paro S, Claydon B, Imler J-L, Meignin C, Pfeffer S. 2015. Cross-species comparative analysis of Dicer proteins during Sindbis virus infection. *Scientific Reports* **5**:10693. doi:10.1038/srep10693

Girardi E, Miesen P, Pennings B, Frangeul L, Saleh M-C, van Rij RP. 2017. Histone-derived piRNA biogenesis depends on the ping-pong partners Piwi5 and Ago3 in Aedes aegypti. *Nucleic Acids Res* **45**:4881–4892. doi:10.1093/nar/gkw1368

Glaser RL, Meola MA. 2010. The native Wolbachia endosymbionts of Drosophila melanogaster and Culex quinquefasciatus increase host resistance to West Nile virus infection. *PLoS ONE* **5**:e11977. doi:10.1371/journal.pone.0011977

Gleason RJ, Anand A, Kai T, Chen X. 2018. Protecting and Diversifying the Germline. *Genetics* **208**:435–471. doi:10.1534/genetics.117.300208

Goic B, Stapleford KA, Frangeul L, Doucet AJ, Gausson V, Blanc H, Schemmel-Jofre N, Cristofari G, Lambrechts L, Vignuzzi M, Saleh M-C. 2016. Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat Commun* **7**. doi:10.1038/ncomms12410

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

Gomariz-Zilber E, Poras M, Thomas-Orillard M. 1995. Drosophila C virus: experimental study of infectious yields and underlying pathology in Drosophila melanogaster laboratory populations. *J Invertebr Pathol* **65**:243–247. doi:10.1006/jipa.1995.1037

González J, Karasov TL, Messer PW, Petrov DA. 2010. Genome-wide patterns of adaptation to temperate environments associated with transposable elements in Drosophila. *PLoS Genet* **6**:e1000905. doi:10.1371/journal.pgen.1000905

González J, Lenkov K, Lipatov M, Macpherson JM, Petrov DA. 2008. High rate of recent transposable element-induced adaptation in Drosophila melanogaster. *PLoS Biol* **6**:e251. doi:10.1371/journal.pbio.0060251

Goriaux C, Desset S, Renaud Y, Vaury C, Brasset E. 2014. Transcriptional properties and splicing of the flamenco piRNA cluster. *EMBO Rep* **15**:411–418. doi:10.1002/embr.201337898

Graham SC, Assenberg R, Delmas O, Verma A, Gholami A, Talbi C, Owens RJ, Stuart DI, Grimes JM, Bourhy H. 2008. Rhabdovirus Matrix Protein Structures Reveal a Novel Mode of Self-Association. *PLoS Pathog* **4**. doi:10.1371/journal.ppat.1000251

Grandbastien M-A, Audeon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa A-PP, Le QH, Melayah D, Petit M, Poncet C, Tam SM, Van Sluys M-A, Mhiri C. 2005. Stress activation and genomic

impact of Tnt1 retrotransposons in Solanaceae. *Cytogenet Genome Res* **110**:229–241. doi:10.1159/000084957
Green J, Griffiths G, Louvard D, Quinn P, Warren G. 1981. Passage of viral membrane proteins through the Golgi complex. *J Mol Biol* **152**:663–698.
Greiser-Wilke I, Moenning V, Kaaden OR, Figueiredo LT. 1989. Most alphaviruses share a conserved epitopic region on their nucleocapsid protein. *J Gen Virol* **70 (Pt 3)**:743–748. doi:10.1099/0022-1317-70-3-743
Grentzinger T, Chambeyron S. 2014. Fast and accurate method to purify small noncoding RNAs from Drosophila ovaries. *Methods Mol Biol* **1093**:171–182. doi:10.1007/978-1-62703-694-8_14
Gubler DJ. 2001. Human arbovirus infections worldwide. *Ann N Y Acad Sci* **951**:13–24.
Guerreiro MPG. 2012. What makes transposable elements move in the Drosophila genome? *Heredity (Edinb)* **108**:461–468. doi:10.1038/hdy.2011.89
Guida V, Cernilogar FM, Filograna A, De Gregorio R, Ishizu H, Siomi MC, Schotta G, Bellenchi GC, Andrenacci D. 2016. Production of Small Noncoding RNAs from the flamenco Locus Is Regulated by the gypsy Retrotransposon of Drosophila melanogaster. *Genetics* **204**:631–644. doi:10.1534/genetics.116.187922
Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. *Science* **315**:1587–1590. doi:10.1126/science.1140494
Ha M, Kim VN. 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**:509–524. doi:10.1038/nrm3838
Habayeb MS, Cantera R, Casanova G, Ekström J-O, Albright S, Hultmark D. 2009. The Drosophila Nora virus is an enteric virus, transmitted via feces. *Journal of Invertebrate Pathology* **101**:29–33. doi:10.1016/j.jip.2009.02.003
Habayeb MS, Ekengren SK, Hultmark D. 2006. Nora virus, a persistent virus in Drosophila, defines a new picorna-like virus family. *Journal of General Virology* **87**:3045–3051. doi:10.1099/vir.0.81997-0
Hahn YS, Strauss EG, Strauss JH. 1989. Mapping of RNA- temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *J Virol* **63**:3142–3150.
Hancks DC, Kazazian HH. 2016. Roles for retrotransposon insertions in human disease. *Mobile DNA* **7**:9. doi:10.1186/s13100-016-0065-9
Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, Brennecke J. 2013. The Genetic Makeup of the Drosophila piRNA Pathway. *Molecular Cell* **50**:762–777. doi:10.1016/j.molcel.2013.04.031
Harada K, Yukuhiro K, Mukai T. 1990. Transposition rates of movable genetic elements in Drosophila melanogaster. *Proc Natl Acad Sci USA* **87**:3248–3252. doi:10.1073/pnas.87.8.3248
Hardy WR, Hahn YS, de Groot RJ, Strauss EG, Strauss JH. 1990. Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus. *Virology* **177**:199–208.
Hardy WR, Strauss JH. 1989. Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J Virol* **63**:4653–4664.
Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R, Gene Ontology Consortium. 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* **32**:D258-261. doi:10.1093/nar/gkh036

Harrison DA, McCoon PE, Binari R, Gilman M, Perrimon N. 1998. Drosophila unpaired encodes a secreted protein that activates the JAK signaling pathway. *Genes Dev* **12**:3252–3263.
Harsh S, Ozakman Y, Kitchen SM, Paquin-Proulx D, Nixon DF, Eleftherianos I. 2018. Dicer-2 Regulates Resistance and Maintains Homeostasis against Zika Virus Infection in Drosophila. *J Immunol* **201**:3058–3072. doi:10.4049/jimmunol.1800597
Harty RN, Brown ME, McGettigan JP, Wang G, Jayakar HR, Huibregtse JM, Whitt MA, Schnell MJ. 2001. Rhabdoviruses and the cellular ubiquitin-proteasome system: a budding interaction. *J Virol* **75**:10623–10629. doi:10.1128/JVI.75.22.10623-10629.2001
Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. 2008. Wolbachia and virus protection in insects. *Science* **322**:702. doi:10.1126/science.1162418
Hefti E, Bishop DH, Dubin DT, Stollar V. 1975. 5' nucleotide sequence of sindbis viral RNA. *J Virol* **17**:149–159.
Hertig M, Wolbach SB. 1924. Studies on Rickettsia-Like Micro-Organisms in Insects. *J Med Res* **44**:329-374.7.
Hess AM, Prasad AN, Ptitsyn A, Ebel GD, Olson KE, Barbacioru C, Monighetti C, Campbell CL. 2011. Small RNA profiling of Dengue virus-mosquito interactions implicates the PIWI RNA pathway in anti-viral defense. *BMC Microbiol* **11**:45. doi:10.1186/1471-2180-11-45
Hesson JC, Verner-Carlsson J, Larsson A, Ahmed R, Lundkvist Å, Lundström JO. 2015. Culex torrentium Mosquito Role as Major Enzootic Vector Defined by Rate of Sindbis Virus Infection, Sweden, 2009. *Emerging Infect Dis* **21**:875–878. doi:10.3201/eid2105.141577
Holden PR, Brookfield JF, Jones P. 1993. Cloning and characterization of an ftsZ homologue from a bacterial symbiont of Drosophila melanogaster. *Mol Gen Genet* **240**:213–220.
Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JMC, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, Cai S, Center A, Chaturverdi K, Christophides GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanigan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic I, Levitsky A, Liang Y, Lin J-J, Lobo NF, Lopez JR, Malek JA, McIntosh TC, Meister S, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, Zhao S, Zhu SC, Zhimulev I, Coluzzi M, della Torre A, Roth CW, Louis C, Kalush F, Mural RJ, Myers EW, Adams MD, Smith HO, Broder S, Gardner MJ, Fraser CM, Birney E, Bork P, Brey PT, Venter JC, Weissenbach J, Kafatos FC, Collins FH, Hoffman SL. 2002. The genome sequence of the malaria mosquito Anopheles gambiae. *Science* **298**:129–149. doi:10.1126/science.1076181
Hombría JC-G, Brown S, Häder S, Zeidler MP. 2005. Characterisation of Upd2, a Drosophila JAK/STAT pathway ligand. *Dev Biol* **288**:420–433. doi:10.1016/j.ydbio.2005.09.040
Huang CRL, Burns KH, Boeke JD. 2012. Active Transposition in Genomes. *Annu Rev Genet* **46**:651–675. doi:10.1146/annurev-genet-110711-155616
Huang H, Li Y, Szulwach KE, Zhang G, Jin P, Chen D. 2014. AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification. *J Cell Biol* **206**:217–230. doi:10.1083/jcb.201401002
Huang X, Tóth KF, Aravin AA. 2017. piRNA Biogenesis in Drosophila melanogaster. *Trends in Genetics* **33**:882–894. doi:10.1016/j.tig.2017.09.002
Huang XA, Yin H, Sweeney S, Raha D, Snyder M, Lin H. 2013. A Major Epigenetic Programming Mechanism Guided by piRNAs. *Developmental Cell* **24**:502–516. doi:10.1016/j.devcel.2013.01.023

Huang Z, Kingsolver MB, Avadhanula V, Hardy RW. 2013. An Antiviral Role for Antimicrobial Peptides during the Arthropod Response to Alphavirus Replication. *J Virol* **87**:4272–4280. doi:10.1128/JVI.03360-12

Hublitz P, Albert M, Peters AHFM. 2009. Mechanisms of transcriptional repression by histone lysine methylation. *Int J Dev Biol* **53**:335–354. doi:10.1387/ijdb.082717ph

Hughes TT, Allen AL, Bardin JE, Christian MN, Daimon K, Dozier KD, Hansen CL, Holcomb LM, Ahlander J. 2012. Drosophila as a genetic model for studying pathogenic human viruses. *Virology* **423**:1–5. doi:10.1016/j.virol.2011.11.016

Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L, Hannon GJ. 2012. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **491**:279–283. doi:10.1038/nature11502

Iwasaki YW, Murano K, Ishizu H, Shibuya A, Iyoda Y, Siomi MC, Siomi H, Saito K. 2016. Piwi Modulates Chromatin Accessibility by Regulating Multiple Factors Including Histone H1 to Repress Transposons. *Molecular Cell* **63**:408–419. doi:10.1016/j.molcel.2016.06.008

Iwasaki YW, Siomi MC, Siomi H. 2015. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annual Review of Biochemistry* **84**:405–433. doi:10.1146/annurev-biochem-060614-034258

Jakob NJ, Kleespies RG, Tidona CA, Müller K, Gelderblom HR, Darai G. 2002. Comparative analysis of the genome and host range characteristics of two insect iridoviruses: Chilo iridescent virus and a cricket iridovirus isolate. *J Gen Virol* **83**:463–470. doi:10.1099/0022-1317-83-2-463

Jason Pitts R. 2014. A blood-free protein meal supporting oogenesis in the Asian tiger mosquito, Aedes albopictus (Skuse). *Journal of Insect Physiology* **64**:1–6. doi:10.1016/j.jinsphys.2014.02.012

Jones BC, Wood JG, Chang C, Tam AD, Franklin MJ, Siegel ER, Helfand SL. 2016. A somatic piRNA pathway in the Drosophila fat body ensures metabolic homeostasis and normal lifespan. *Nature Communications* **7**:13856. doi:10.1038/ncomms13856

Jose J, Snyder JE, Kuhn RJ. 2009. A structural and functional perspective of alphavirus replication and assembly. *Future Microbiol* **4**:837–856. doi:10.2217/fmb.09.59

Jouan-Dufournel I, Cosset FL, Contamine D, Verdier G, Biémont C. 1996. Transposable elements behavior following viral genomic stress in Drosophila melanogaster inbred line. *J Mol Evol* **43**:19–27.

Jousset FX, Bergoin M, Revet B. 1977. Characterization of the Drosophila C virus. *J Gen Virol* **34**:269–283. doi:10.1099/0022-1317-34-2-269

Kakumani PK, Ponia SS, S RK, Sood V, Chinnappan M, Banerjea AC, Medigeshi GR, Malhotra P, Mukherjee SK, Bhatnagar RK. 2013. Role of RNA interference (RNAi) in dengue virus replication and identification of NS4B as an RNAi suppressor. *J Virol* **87**:8870–8883. doi:10.1128/JVI.02774-12

Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker SE. 2002. The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol* **3**:RESEARCH0084.

Kapitonov VV, Jurka J. 2001. Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* **98**:8714–8719. doi:10.1073/pnas.151269298

Kapun M, Nolte V, Flatt T, Schlötterer C. 2010. Host Range and Specificity of the Drosophila C Virus. *PLOS ONE* **5**:e12421. doi:10.1371/journal.pone.0012421

Karlikow M, Goic B, Mongelli V, Salles A, Schmitt C, Bonne I, Zurzolo C, Saleh M-C. 2016. Drosophila cells use nanotube-like structures to transfer dsRNA and RNAi machinery between cells. *Sci Rep* **6**:27085. doi:10.1038/srep27085

Karlikow M, Goic B, Saleh M-C. 2014. RNAi and antiviral defense in Drosophila: Setting up a systemic immune response. *Developmental & Comparative Immunology*, Drosophila Immunity **42**:85–92. doi:10.1016/j.dci.2013.05.004

Katzourakis A, Gifford RJ. 2010. Endogenous Viral Elements in Animal Genomes. *PLoS Genet* **6**. doi:10.1371/journal.pgen.1001191

Kawai T, Akira S. 2009. The roles of TLRs, RLRs and NLRs in pathogen recognition. *Int Immunol* **21**:317–337. doi:10.1093/intimm/dxp017

Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H. 2008. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453**:793–797. doi:10.1038/nature06938

Kawaoka S, Izumi N, Katsuma S, Tomari Y. 2011. 3' end formation of PIWI-interacting RNAs in vitro. *Mol Cell* **43**:1015–1022. doi:10.1016/j.molcel.2011.07.029

Keene KM, Foy BD, Sanchez-Vargas I, Beaty BJ, Blair CD, Olson KE. 2004. RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of Anopheles gambiae. *PNAS* **101**:17240–17245. doi:10.1073/pnas.0406983101

Kemp C, Mueller S, Goto A, Barbier V, Paro S, Bonnay F, Dostert C, Troxler L, Hetru C, Meignin C, Pfeffer S, Hoffmann JA, Imler J-L. 2013. Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in Drosophila. *J Immunol* **190**:650–658. doi:10.4049/jimmunol.1102486

Kidwell MG. 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**:49–63.

Kielian M, Chanel-Vos C, Liao M. 2010. Alphavirus Entry and Membrane Fusion. *Viruses* **2**:796–825. doi:10.3390/v2040796

Kim A, Terzian C, Santamaria P, Pélisson A, Purd'homme N, Bucheton A. 1994. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of Drosophila melanogaster. *Proc Natl Acad Sci U S A* **91**:1285–1289.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**:R36. doi:10.1186/gb-2013-14-4-r36

Kim KH, Rümenapf T, Strauss EG, Strauss JH. 2004. Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology* **323**:153–163. doi:10.1016/j.virol.2004.03.009

Kingsolver MB, Huang Z, Hardy RW. 2013. Insect Antiviral Innate Immunity: Pathways, Effectors, and Connections. *Journal of Molecular Biology*, Antiviral Innate Immunity (Part I) **425**:4921–4936. doi:10.1016/j.jmb.2013.10.006

Klenov MS, Lavrov SA, Stolyarenko AD, Ryazansky SS, Aravin AA, Tuschl T, Gvozdev VA. 2007. Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Res* **35**:5430–5438. doi:10.1093/nar/gkm576

Koito A, Ikeda T. 2013. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol* **4**:28. doi:10.3389/fmicb.2013.00028

Kramer JM, Kochinke K, Oortveld MAW, Marks H, Kramer D, de Jong EK, Asztalos Z, Westwood JT, Stunnenberg HG, Sokolowski MB, Keleman K, Zhou H, van Bokhoven H, Schenck A. 2011. Epigenetic regulation of learning and memory by Drosophila EHMT/G9a. *PLoS Biol* **9**:e1000569. doi:10.1371/journal.pbio.1000569

Kurkela S, Helve T, Vaheri A, Vapalahti O. 2008a. Arthritis and arthralgia three years after Sindbis virus infection: clinical follow-up of a cohort of 49 patients. *Scand J Infect Dis* **40**:167–173. doi:10.1080/00365540701586996

Kurkela S, Rätti O, Huhtamo E, Uzcátegui NY, Nuorti JP, Laakkonen J, Manni T, Helle P, Vaheri A, Vapalahti O. 2008b. Sindbis virus infection in resident birds, migratory birds, and humans, Finland. *Emerging Infect Dis* **14**:41–47. doi:10.3201/eid1401.070510

Laine M, Luukkainen R, Toivanen A. 2004. Sindbis viruses and other alphaviruses as cause of human arthritic disease. *J Intern Med* **256**:457–471. doi:10.1111/j.1365-2796.2004.01413.x

Lamiable O, Arnold J, de Faria IJ da S, Olmo RP, Bergami F, Meignin C, Hoffmann JA, Marques JT, Imler J-L. 2016a. Analysis of the Contribution of Hemocytes and Autophagy to Drosophila Antiviral Immunity. *J Virol* **90**:5415–5426. doi:10.1128/JVI.00238-16

Lamiable O, Kellenberger C, Kemp C, Troxler L, Pelte N, Boutros M, Marques JT, Daeffler L, Hoffmann JA, Roussel A, Imler J-L. 2016b. Cytokine Diedel and a viral homologue suppress the IMD pathway in Drosophila. *PNAS* **113**:698–703. doi:10.1073/pnas.1516122113

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J, International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921. doi:10.1038/35057062

LaPointe AT, Moreno-Contreras J, Sokoloski KJ. 2018. Increasing the Capping Efficiency of the Sindbis Virus nsP1 Protein Negatively Affects Viral Infection. *bioRxiv*. doi:10.1101/453068

Lark T, Keck F, Narayanan A. 2018. Interactions of Alphavirus nsP3 Protein with Host Proteins. *Front Microbiol* **8**. doi:10.3389/fmicb.2017.02652

Lau NC, Robine N, Martin R, Chung W-J, Niki Y, Berezikov E, Lai EC. 2009. Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. *Genome Res* **19**:1776–1785. doi:10.1101/gr.094896.109

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. 2013. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* **27**:390–399. doi:10.1101/gad.209841.112

Lee S, Owen KE, Choi HK, Lee H, Lu G, Wengler G, Brown DT, Rossmann MG, Kuhn RJ. 1996. Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly. *Structure* **4**:531–541.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**:415–419. doi:10.1038/nature01957

Lee YCG, Karpen GH. 2017. Pervasive epigenetic effects of Drosophila euchromatic transposable elements impact their evolution. *eLife* **6**.
Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW. 2004. Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways. *Cell* **117**:69–81. doi:10.1016/S0092-8674(04)00261-2
Léger P, Lara E, Jagla B, Sismeiro O, Mansuroglu Z, Coppée JY, Bonnefoy E, Bouloy M. 2013. Dicer-2- and Piwi-Mediated RNA Interference in Rift Valley Fever Virus-Infected Mosquito Cells. *Journal of Virology* **87**:1631–1648. doi:10.1128/JVI.02795-12
Lemaitre B, Hoffmann J. 2007. The host defense of Drosophila melanogaster. *Annu Rev Immunol* **25**:697–743. doi:10.1146/annurev.immunol.25.022106.141615
Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell* **86**:973–983.
Lemm JA, Rümenapf T, Strauss EG, Strauss JH, Rice CM. 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J* **13**:2925–2934.
Lentz TL, Burrage TG, Smith AL, Tignor GH. 1983. The acetylcholine receptor as a cellular receptor for rabies virus. *Yale J Biol Med* **56**:315–322.
Lerat E, Burlet N, Biémont C, Vieira C. 2011. Comparative analysis of transposable elements in the melanogaster subgroup sequenced genomes. *Gene* **473**:100–109. doi:10.1016/j.gene.2010.11.009
Lerat E, Fablet M, Modolo L, Lopez-Maestre H, Vieira C. 2017. TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. *Nucleic Acids Res* **45**:e17. doi:10.1093/nar/gkw953
Leulier F, Rodriguez A, Khush RS, Abrams JM, Lemaitre B. 2000. The Drosophila caspase Dredd is required to resist gram-negative bacterial infection. *EMBO Rep* **1**:353–358. doi:10.1093/embo-reports/kvd073
Leulier F, Vidal S, Saigo K, Ueda R, Lemaitre B. 2002. Inducible expression of double-stranded RNA reveals a role for dFADD in the regulation of the antibacterial response in Drosophila adults. *Curr Biol* **12**:996–1000.
Levis R, Schlesinger S, Huang HV. 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J Virol* **64**:1726–1733.
Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen FM. 1993. Transposons in place of telomeric repeats at a Drosophila telomere. *Cell* **75**:1083–1093.
Lewis SH, Quarles KA, Yang Y, Tanguy M, Frézal L, Smith SA, Sharma PP, Cordaux R, Gilbert C, Giraud I, Collins DH, Zamore PD, Miska EA, Sarkies P, Jiggins FM. 2018. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat Ecol Evol* **2**:174–181. doi:10.1038/s41559-017-0403-4
Lewis SH, Salmela H, Obbard DJ. 2016. Duplication and Diversification of Dipteran Argonaute Genes, and the Evolutionary Divergence of Piwi and Aubergine. *Genome Biol Evol* **8**:507–518. doi:10.1093/gbe/evw018
L'heritier P. 1958. The hereditary virus of Drosophila. *Adv Virus Res* **5**:195–245.
L'Heritier P. 1957. The hereditary virus of Drosophila. *Adv Virus Res 5, 195–245*.
L'Heritier P. 1948. Sensitivity to CO2 in Drosophila; a review. *Heredity* **2**:325–348.
Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, Kittler ELW, Zapp ML, Klattenhoff C, Schulz N, Theurkauf WE, Weng Z, Zamore PD. 2009. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* **137**:509–521. doi:10.1016/j.cell.2009.04.027
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**:2078–2079. doi:10.1093/bioinformatics/btp352

Li H, Li WX, Ding SW. 2002. Induction and suppression of RNA silencing by an animal virus. *Science* **296**:1319–1321. doi:10.1126/science.1070948

Li W, Prazak L, Chatterjee N, Grüninger S, Krug L, Theodorou D, Dubnau J. 2013. Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat Neurosci* **16**:529–531. doi:10.1038/nn.3368

Lim EXY, Lee WS, Madzokere ET, Herrero LJ. 2018. Mosquitoes as Suitable Vectors for Alphaviruses. *Viruses* **10**. doi:10.3390/v10020084

Lindgreen S. 2012. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res Notes* **5**:337. doi:10.1186/1756-0500-5-337

Liu X, Jiang F, Kalidas S, Smith D, Liu Q. 2006. Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. *RNA* **12**:1514–1520. doi:10.1261/rna.101606

Lobigs M, Zhao HX, Garoff H. 1990. Function of Semliki Forest virus E3 peptide in virus assembly: replacement of E3 with an artificial signal peptide abolishes spike heterodimerization and surface expression of E1. *J Virol* **64**:4346–4355.

Longdon B, Day JP, Schulz N, Leftwich PT, de Jong MA, Breuker CJ, Gibbs M, Obbard DJ, Wilfert L, Smith SCL, McGonigle JE, Houslay TM, Wright LI, Livraghi L, Evans LC, Friend LA, Chapman T, Vontas J, Kambouraki N, Jiggins FM. 2017. Vertically transmitted rhabdoviruses are found across three insect families and have dynamic interactions with their hosts. *Proc Biol Sci* **284**. doi:10.1098/rspb.2016.2381

Longdon B, Obbard DJ, Jiggins FM. 2010. Sigma viruses from three species of Drosophila form a major new clade in the rhabdovirus phylogeny. *Proc Biol Sci* **277**:35–44. doi:10.1098/rspb.2009.1472

Longdon B, Wilfert L, Jiggins FM, Dietzgen R, Kuzmin I. 2012. The sigma viruses of Drosophila. *Rhabdoviruses: Molecular Taxonomy, Evolution, Genomics, Ecology, Cytopathology and Control* 117–132.

Lopez W, Page AM, Carlson DJ, Ericson BL, Cserhati MF, Guda C, Carlson KA. 2018. Analysis of immune-related genes during Nora virus infection of Drosophila melanogaster using next generation sequencing. *AIMS Microbiol* **4**:123–139. doi:10.3934/microbiol.2018.1.123

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550. doi:10.1186/s13059-014-0550-8

Lundström JO, Lindström KM, Olsen B, Dufva R, Krakower DS. 2001. Prevalence of sindbis virus neutralizing antibodies among Swedish passerines indicates that thrushes are the main amplifying hosts. *J Med Entomol* **38**:289–297. doi:10.1603/0022-2585-38.2.289

Lundström JO, Pfeffer M. 2010. Phylogeographic structure and evolutionary history of Sindbis virus. *Vector Borne Zoonotic Dis* **10**:889–907. doi:10.1089/vbz.2009.0069

Luteijn MJ, Ketting RF. 2013. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nature Reviews Genetics* **14**:523–534. doi:10.1038/nrg3495

Luyet P-P, Falguières T, Pons V, Pattnaik AK, Gruenberg J. 2008. The ESCRT-I subunit TSG101 controls endosome-to-cytosol release of viral RNA. *Traffic* **9**:2279–2290. doi:10.1111/j.1600-0854.2008.00820.x

L'vov DK, Skvortsova TM, Gromashevskiĭ VL, Berezina LK, Iakovlev VI. 1985. [Isolation of the causative agent of Karelian fever from Aedes sp. mosquitoes]. *Vopr Virusol* **30**:311–313.

Lwande OW, Näslund J, Lundmark E, Ahlm K, Ahlm C, Bucht G, Evander M. 2019. Experimental Infection and Transmission Competence of Sindbis Virus in Culex torrentium and Culex pipiens Mosquitoes from Northern Sweden. *Vector Borne Zoonotic Dis* **19**:128–133. doi:10.1089/vbz.2018.2311

Lwande OW, Obanda V, Bucht G, Mosomtai G, Otieno V, Ahlm C, Evander M. 2015. Global emergence of Alphaviruses that cause arthritis in humans. *Infect Ecol Epidemiol* **5**:29853. doi:10.3402/iee.v5.29853

Lyles, D.S., and Rupprecht, C.E. 2007. Rhabdoviridae.In Fields Virology, Knipe, D.M., and Howley, P.M., Eds. pp. 1364–1394.

Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. 2011. Successive Increases in the Resistance of Drosophila to Viral Infection through a Transposon Insertion Followed by a Duplication. *PLoS Genet* **7**. doi:10.1371/journal.pgen.1002337

Majumdar S, Rio DC. 2015. P transposable elements in Drosophila and other eukaryotic organisms. *Microbiol Spectr* **3**:MDNA3-0004–2014. doi:10.1128/microbiolspec.MDNA3-0004-2014

Makarova JA, Shkurnikov MU, Wicklein D, Lange T, Samatov TR, Turchinovich AA, Tonevitsky AG. 2016. Intracellular and extracellular microRNA: An update on localization and biological role. *Prog Histochem Cytochem* **51**:33–49. doi:10.1016/j.proghi.2016.06.001

Malik HS, Henikoff S, Eickbush TH. 2000. Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* **10**:1307–1318.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. *Cell* **137**:522–535. doi:10.1016/j.cell.2009.03.040

Manousis T, Moore NF. 1987. Cricket Paralysis Virus, a Potential Control Agent for the Olive Fruit Fly, Dacus oleae Gmel. *Appl Environ Microbiol* **53**:142–148.

Marques JT, Imler J-L. 2016. The diversity of insect antiviral immunity: insights from viruses. *Curr Opin Microbiol* **32**:71–76. doi:10.1016/j.mib.2016.05.002

Marsano RM, Leronni D, D'Addabbo P, Viggiano L, Tarasco E, Caizzi R. 2012. Mosquitoes LTR Retrotransposons: A Deeper View into the Genomic Sequence of Culex quinquefasciatus. *PLOS ONE* **7**:e30770. doi:10.1371/journal.pone.0030770

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**:10–12. doi:10.14806/ej.17.1.200

Maruyama K, Hartl DL. 1991. Evolution of the Transposable Element Mariner in Drosophila Species. *Genetics* **128**:319–329.

Mathur K, Anand A, Dubey SK, Sanan-Mishra N, Bhatnagar RK, Sunil S. 2016. Analysis of chikungunya virus proteins reveals that non-structural proteins nsP2 and nsP3 exhibit RNA interference (RNAi) suppressor activity. *Sci Rep* **6**. doi:10.1038/srep38065

Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, Glassford WJ, Herre M, Redmond SN, Rose NH, Weedall GD, Wu Y, Batra SS, Brito-Sierra CA, Buckingham SD, Campbell CL, Chan S, Cox E, Evans BR, Fansiri T, Filipović I, Fontaine A, Gloria-Soria A, Hall R, Joardar VS, Jones AK, Kay RGG, Kodali VK, Lee J, Lycett GJ, Mitchell SN, Muehling J, Murphy MR, Omer AD, Partridge FA, Peluso P, Aiden AP, Ramasamy V, Rašić G, Roy S, Saavedra-Rodriguez K, Sharan S, Sharma A, Smith ML, Turner J, Weakley AM, Zhao Z, Akbari OS, Black WC, Cao H, Darby AC, Hill CA, Johnston JS, Murphy TD, Raikhel AS, Sattelle DB, Sharakhov IV, White BJ, Zhao L, Aiden EL, Mann RS, Lambrechts L, Powell JR, Sharakhova MV, Tu Z, Robertson HM, McBride CS, Hastie AR, Korlach J, Neafsey DE, Phillippy AM, Vosshall LB. 2018. Improved reference genome of Aedes aegypti informs arbovirus vector control. *Nature* **563**:501. doi:10.1038/s41586-018-0692-z

McClintock B. 1950. The Origin and Behavior of Mutable Loci in Maize. *Proc Natl Acad Sci U S A* **36**:344–355.

Medzhitov R, Preston-Hurlburt P, Janeway CA. 1997. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* **388**:394–397. doi:10.1038/41131

Medzhitov R, Schneider DS, Soares MP. 2012. Disease Tolerance as a Defense Strategy. *Science* **335**:936–941. doi:10.1126/science.1214935

Melancon P, Garoff H. 1987. Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease. *J Virol* **61**:1301–1309.

Mellroth P, Steiner H. 2006. PGRP-SB1: an N-acetylmuramoyl L-alanine amidase with antibacterial activity. *Biochem Biophys Res Commun* **350**:994–999. doi:10.1016/j.bbrc.2006.09.139

Meng X, Khanuja BS, Ip YT. 1999. Toll receptor-mediated Drosophila immune response requires Dif, an NF-kappaB factor. *Genes Dev* **13**:792–797.

Merenciano M, Ullastres A, de Cara M a. R, Barrón MG, González J. 2016. Multiple Independent Retroelement Insertions in the Promoter of a Stress Response Gene Have Variable Molecular

and Functional Effects in Drosophila. *PLoS Genet* **12**:e1006249. doi:10.1371/journal.pgen.1006249
Merkling SH, Bronkhorst AW, Kramer JM, Overheul GJ, Schenck A, Van Rij RP. 2015. The Epigenetic Regulator G9a Mediates Tolerance to RNA Virus Infection in Drosophila. *PLoS Pathog* **11**. doi:10.1371/journal.ppat.1004692
Merkling SH, van Rij RP. 2015. Analysis of resistance and tolerance to virus infection in Drosophila. *Nat Protocols* **10**:1084–1097. doi:10.1038/nprot.2015.071
Merkling SH, van Rij RP. 2013. Beyond RNAi: antiviral defense strategies in Drosophila and mosquito. *J Insect Physiol* **59**:159–170. doi:10.1016/j.jinsphys.2012.07.004
Mével-Ninio M, Pelisson A, Kinder J, Campos AR, Bucheton A. 2007. The flamenco Locus Controls the gypsy and ZAM Retroviruses and Is Required for Drosophila Oogenesis. *Genetics* **175**:1615–1624. doi:10.1534/genetics.106.068106
Mierlo JT van, Bronkhorst AW, Overheul GJ, Sadanandan SA, Ekström J-O, Heestermans M, Hultmark D, Antoniewski C, Rij RP van. 2012. Convergent Evolution of Argonaute-2 Slicer Antagonism in Two Distinct Insect RNA Viruses. *PLOS Pathogens* **8**:e1002872. doi:10.1371/journal.ppat.1002872
Mierlo JT van, Overheul GJ, Obadia B, Cleef KWR van, Webster CL, Saleh M-C, Obbard DJ, Rij RP van. 2014. Novel Drosophila Viruses Encode Host-Specific Suppressors of RNAi. *PLOS Pathogens* **10**:e1004256. doi:10.1371/journal.ppat.1004256
Miesen P, Girardi E, van Rij RP. 2015. Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in Aedes aegypti mosquito cells. *Nucleic Acids Res* **43**:6545–6556. doi:10.1093/nar/gkv590
Miesen P, Ivens A, Buck AH, van Rij RP. 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. doi:10.1371/journal.pntd.0004452
Miesen P, Joosten J, van Rij RP. 2016b. PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector Mosquitoes. *PLOS Pathogens* **12**:e1006017. doi:10.1371/journal.ppat.1006017
Mills RE, Bennett EA, Iskow RC, Devine SE. 2007. Which transposable elements are active in the human genome? *Trends Genet* **23**:183–191. doi:10.1016/j.tig.2007.02.006
Mirkovic-Hösle M, Förstemann K. 2014. Transposon Defense by Endo-siRNAs, piRNAs and Somatic pilRNAs in Drosophila: Contributions of Loqs-PD and R2D2. *PLoS One* **9**. doi:10.1371/journal.pone.0084994
Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC. 2005. Slicer function of Drosophila Argonautes and its involvement in RISC formation. *Genes Dev* **19**:2837–2848. doi:10.1101/gad.1370605
Modolo L, Lerat E. 2015. UrQt: an efficient software for the Unsupervised Quality trimming of NGS data. *BMC Bioinformatics* **16**. doi:10.1186/s12859-015-0546-8
Mondotte JA, Saleh M-C. 2018. Antiviral Immune Response and the Route of Infection in Drosophila melanogaster. *Adv Virus Res* **100**:247–278. doi:10.1016/bs.aivir.2017.10.006
Monsanto-Hearne V, Johnson KN. 2018. Wolbachia-mediated protection of Drosophila melanogaster against systemic infection with its natural viral pathogen Drosophila C virus does not involve changes in levels of highly abundant miRNAs. *Journal of General Virology* **99**:827–831. doi:10.1099/jgv.0.001064
Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM. 2012. Production of Virus-Derived Ping-Pong-Dependent piRNA-like Small RNAs in the Mosquito Soma. *PLoS Pathog* **8**. doi:10.1371/journal.ppat.1002470
Morley VJ, Mendiola SY, Turner PE. 2015. Rate of novel host invasion affects adaptability of evolving RNA virus lineages. *Proceedings of the Royal Society B: Biological Sciences* **282**:20150801. doi:10.1098/rspb.2015.0801
Mueller S, Gausson V, Vodovar N, Deddouche S, Troxler L, Perot J, Pfeffer S, Hoffmann JA, Saleh M-C, Imler J-L. 2010. RNAi-mediated immunity provides strong protection against the negative-strand RNA vesicular stomatitis virus in Drosophila. *Proc Natl Acad Sci USA* **107**:19390–19395. doi:10.1073/pnas.1014378107

Mugnier N, Gueguen L, Vieira C, Biémont C. 2008. The heterochromatic copies of the LTR retrotransposons as a record of the genomic events that have shaped the Drosophila melanogaster genome. *Gene* **411**:87–93. doi:10.1016/j.gene.2008.01.010

Muñoz-López M, García-Pérez JL. 2010. DNA Transposons: Nature and Applications in Genomics. *Curr Genomics* **11**:115–128. doi:10.2174/138920210790886871

Myles KM, Pierro DJ, Olson KE. 2004. Comparison of the transmission potential of two genetically distinct Sindbis viruses after oral infection of Aedes aegypti (Diptera: Culicidae). *J Med Entomol* **41**:95–106. doi:10.1603/0022-2585-41.1.95

Myles KM, Wiley MR, Morazzani EM, Adelman ZN. 2008. Alphavirus-derived small RNAs modulate pathogenesis in disease vector mosquitoes. *Proc Natl Acad Sci USA* **105**:19938–19943. doi:10.1073/pnas.0803408105

Nag DK, Brecher M, Kramer LD. 2016. DNA forms of arboviral RNA genomes are generated following infection in mosquito cell cultures. *Virology* **498**:164–171. doi:10.1016/j.virol.2016.08.022

Nag DK, Kramer LD. 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. doi:10.1099/jgv.0.000945

Nainu F, Tanaka Y, Shiratsuchi A, Nakanishi Y. 2015. Protection of Insects against Viral Infection by Apoptosis-Dependent Phagocytosis. *J Immunol* **195**:5696–5706. doi:10.4049/jimmunol.1500613

Naitza S, Rossé C, Kappler C, Georgel P, Belvin M, Gubb D, Camonis J, Hoffmann JA, Reichhart JM. 2002. The Drosophila immune defense against gram-negative infection requires the death protein dFADD. *Immunity* **17**:575–581.

Nayak A, Berry B, Tassetto M, Kunitomi M, Acevedo A, Deng C, Krutchinsky A, Gross J, Antoniewski C, Andino R. 2010. Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in Drosophila. *Nat Struct Mol Biol* **17**:547–554. doi:10.1038/nsmb.1810

Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z, Megy K, Grabherr M, Ren Q, Zdobnov EM, Lobo NF, Campbell KS, Brown SE, Bonaldo MF, Zhu J, Sinkins SP, Hogenkamp DG, Amedeo P, Arensburger P, Atkinson PW, Bidwell S, Biedler J, Birney E, Bruggner RV, Costas J, Coy MR, Crabtree J, Crawford M, Debruyn B, Decaprio D, Eiglmeier K, Eisenstadt E, El-Dorry H, Gelbart WM, Gomes SL, Hammond M, Hannick LI, Hogan JR, Holmes MH, Jaffe D, Johnston JS, Kennedy RC, Koo H, Kravitz S, Kriventseva EV, Kulp D, Labutti K, Lee E, Li S, Lovin DD, Mao C, Mauceli E, Menck CFM, Miller JR, Montgomery P, Mori A, Nascimento AL, Naveira HF, Nusbaum C, O'leary S, Orvis J, Pertea M, Quesneville H, Reidenbach KR, Rogers Y-H, Roth CW, Schneider JR, Schatz M, Shumway M, Stanke M, Stinson EO, Tubio JMC, Vanzee JP, Verjovski-Almeida S, Werner D, White O, Wyder S, Zeng Q, Zhao Q, Zhao Y, Hill CA, Raikhel AS, Soares MB, Knudson DL, Lee NH, Galagan J, Salzberg SL, Paulsen IT, Dimopoulos G, Collins FH, Birren B, Fraser-Liggett CM, Severson DW. 2007. Genome sequence of Aedes aegypti, a major arbovirus vector. *Science* **316**:1718–1723. doi:10.1126/science.1138878

Niklasson B, Espmark A, LeDuc JW, Gargan TP, Ennis WA, Tesh RB, Main AJ. 1984. Association of a Sindbis-like virus with Ockelbo disease in Sweden. *Am J Trop Med Hyg* **33**:1212–1217.

Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, Bonnefond L, Matsumoto N, Nishizawa T, Nakanaga K, Aoki J, Ishitani R, Siomi H, Siomi MC, Nureki O. 2012. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* **491**:284–287. doi:10.1038/nature11509

Obbard DJ, Gordon KHJ, Buck AH, Jiggins FM. 2009. The evolution of RNAi as a defence against viruses and transposable elements. *Philos Trans R Soc Lond B Biol Sci* **364**:99–115. doi:10.1098/rstb.2008.0168

O'Brien J, Hayder H, Zayed Y, Peng C. 2018. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)* **9**. doi:10.3389/fendo.2018.00402

Ohanessian A, Echalier G. 1967. Multiplication of Drosophila Hereditary Virus (σ Virus) in Drosophila Embryonic Cells cultivated in vitro. *Nature* **213**:1049. doi:10.1038/2131049a0

Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K. 2013. DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. *Genes Dev* **27**:1656–1661. doi:10.1101/gad.221515.113

Okamura K, Balla S, Martin R, Liu N, Lai EC. 2008a. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat Struct Mol Biol* **15**:581–590. doi:10.1038/nsmb.1438

Okamura K, Chung W-J, Ruby JG, Guo H, Bartel DP, Lai EC. 2008b. The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**:803–806. doi:10.1038/nature07015

Olivieri D, Senti K-A, Subramanian S, Sachidanandam R, Brennecke J. 2012. The cochaperone shutdown defines a group of biogenesis factors essential for all piRNA populations in Drosophila. *Mol Cell* **47**:954–969. doi:10.1016/j.molcel.2012.07.021

Olivieri D, Sykora MM, Sachidanandam R, Mechtler K, Brennecke J. 2010. An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. *EMBO J* **29**:3301–3317. doi:10.1038/emboj.2010.212

Olovnikov I, Ryazansky S, Shpiz S, Lavrov S, Abramov Y, Vaury C, Jensen S, Kalmykova A. 2013. De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. *Nucleic Acids Res* **41**:5757–5768. doi:10.1093/nar/gkt310

Ou JH, Rice CM, Dalgarno L, Strauss EG, Strauss JH. 1982. Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. *Proc Natl Acad Sci U S A* **79**:5235–5239.

Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. 2019. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet* **20**:89–108. doi:10.1038/s41576-018-0073-3

Padmanabha D, Baker KD. 2014. Drosophila gains traction as a repurposed tool to investigate metabolism. *Trends Endocrinol Metab* **25**:518–527. doi:10.1016/j.tem.2014.03.011

Palatini U, Miesen P, Carballar-Lejarazu R, Ometto L, Rizzo E, Tu Z, van Rij RP, 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. doi:10.1186/s12864-017-3903-3

Panda D, Rose PP, Hanna SL, Gold B, Hopkins KC, Lyde RB, Marks MS, Cherry S. 2013. Genome-wide RNAi screen identifies SEC61A and VCP as conserved regulators of Sindbis virus entry. *Cell Rep* **5**:1737–1748. doi:10.1016/j.celrep.2013.11.028

Paradkar PN, Trinidad L, Voysey R, Duchemin J-B, Walker PJ. 2012. Secreted Vago restricts West Nile virus infection in Culex mosquito cells by activating the Jak-STAT pathway. *Proc Natl Acad Sci USA* **109**:18915–18920. doi:10.1073/pnas.1205231109

Park J-E, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. 2011. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**:201–205. doi:10.1038/nature10198

Paul P, Chakraborty A, Sarkar D, Langthasa M, Rahman M, Bari M, Singha RS, Malakar AK, Chakraborty S. 2018. Interplay between miRNAs and human diseases. *J Cell Physiol* **233**:2007–2018. doi:10.1002/jcp.25854

Pearson MN, Rohrmann GF. 2002. Transfer, incorporation, and substitution of envelope fusion proteins among members of the Baculoviridae, Orthomyxoviridae, and Metaviridae (insect retrovirus) families. *J Virol* **76**:5301–5304.

Pélisson A, Song SU, Prud'homme N, Smith PA, Bucheton A, Corces VG. 1994. Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. *EMBO J* **13**:4401–4411.

Pélisson A, Teysset L, Chalvet F, Kim A, Prud'homme N, Terzian C, Bucheton A. 1997. About the origin of retroviruses and the co-evolution of the gypsy retrovirus with the Drosophila flamenco host gene. *Genetica* **100**:29–37.

Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S. 2013. Transposition driven genomic heterogeneity in the Drosophila brain. *Science* **340**. doi:10.1126/science.1231965

Petersen M, Armisén D, Gibbs RA, Hering L, Khila A, Mayer G, Richards S, Niehuis O, Misof B. 2019. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evol Biol* **19**. doi:10.1186/s12862-018-1324-9

Petit M, Mongelli V, Frangeul L, Blanc H, Jiggins F, Saleh M-C. 2016. piRNA pathway is not required for antiviral defense in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **113**:E4218–E4227. doi:10.1073/pnas.1607952113

Phillips A, Mossel E, Sanchez-Vargas I, Foy B, Olson K. 2010. Alphavirus transducing system: tools for visualizing infection in mosquito vectors. *J Vis Exp*. doi:10.3791/2363

Pietilä MK, Hellström K, Ahola T. 2017. Alphavirus polymerase and RNA replication. *Virus Res* **234**:44–57. doi:10.1016/j.virusres.2017.01.007

Piontkivska H, Matos LF, Paul S, Scharfenberg B, Farmerie WG, Miyamoto MM, Wayne ML. 2016. Role of Host-Driven Mutagenesis in Determining Genome Evolution of Sigma Virus (DMelSV; Rhabdoviridae) in Drosophila melanogaster. *Genome Biol Evol* **8**:2952–2963. doi:10.1093/gbe/evw212

Plus N, Croizier G, Jousset FX, David J. 1975. Picornaviruses of laboratory and wild Drosophila melanogaster: geographical distribution and serotypic composition. *Ann Microbiol (Paris)* **126**:107–117.

Poirier EZ, Goic B, Tomé-Poderti L, Frangeul L, Boussier J, Gausson V, Blanc H, Vallet T, Loyd H, Levi LI, Lanciano S, Baron C, Merkling SH, Lambrechts L, Mirouze M, Carpenter S, Vignuzzi M, Saleh M-C. 2018. Dicer-2-Dependent Generation of Viral DNA from Defective Genomes of RNA Viruses Modulates Antiviral Immunity in Insects. *Cell Host Microbe* **23**:353-365.e8. doi:10.1016/j.chom.2018.02.001

Pouteau S, Grandbastien M-A, Boccara M. 1994. Microbial elicitors of plant defence responses activate transcription of a retrotransposon. *The Plant Journal* **5**:535–542. doi:10.1046/j.1365-313X.1994.05040535.x

Pritham EJ, Putliwala T, Feschotte C. 2007. Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene* **390**:3–17. doi:10.1016/j.gene.2006.08.008

Prud'homme N, Gans M, Masson M, Terzian C, Bucheton A. 1995. Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. *Genetics* **139**:697–711.

Qi H, Watanabe T, Ku H-Y, Liu N, Zhong M, Lin H. 2011. The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. *J Biol Chem* **286**:3789–3797. doi:10.1074/jbc.M110.193888

Quesneville H, Bergman CM, Andrieu O, Autard D, Nouaud D, Ashburner M, Anxolabehere D. 2005. Combined evidence annotation of transposable elements in genome sequences. *PLoS Comput Biol* **1**:166–175. doi:10.1371/journal.pcbi.0010022

Ramirez JL, Dimopoulos G. 2010. The Toll immune signaling pathway control conserved anti-dengue defenses across diverse Ae. aegypti strains and against multiple dengue virus serotypes. *Dev Comp Immunol* **34**:625–629. doi:10.1016/j.dci.2010.01.006

Rancès E, Ye YH, Woolfit M, McGraw EA, O'Neill SL. 2012. The relative importance of innate immune priming in Wolbachia-mediated dengue interference. *PLoS Pathog* **8**:e1002548. doi:10.1371/journal.ppat.1002548

Reed LJ, Muench H. 1938. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS. *Am J Epidemiol* **27**:493–497. doi:10.1093/oxfordjournals.aje.a118408

Reinganum C. 1975. The isolation of cricket paralysis virus from the emperor gum moth, Antheraea eucalypti Scott, and its infectivity towards a range of insect species. *Intervirology* **5**:97–102. doi:10.1159/000149886

Rice CM, Levis R, Strauss JH, Huang HV. 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J Virol* **61**:3809–3819.
Riegler M, Sidhu M, Miller WJ, O'Neill SL. 2005. Evidence for a Global Wolbachia Replacement in Drosophila melanogaster. *Current Biology* **15**:1428–1433. doi:10.1016/j.cub.2005.06.069
Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. 2011. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biology* **12**:R22. doi:10.1186/gb-2011-12-3-r22
Robertson HM, Engels WR. 1989. Modified P elements that mimic the P cytotype in Drosophila melanogaster. *Genetics* **123**:815–824.
Robine N, Lau NC, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower MD, Lai EC. 2009. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr Biol* **19**:2066–2076. doi:10.1016/j.cub.2009.11.064
Roche S, Bressanelli S, Rey FA, Gaudin Y. 2006. Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. *Science* **313**:187–191. doi:10.1126/science.1127683
Roche S, Rey FA, Gaudin Y, Bressanelli S. 2007. Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. *Science* **315**:843–848. doi:10.1126/science.1135710
Roekring S, Flegel TW, Malasit P, Kittayapong P. 2006. Challenging successive mosquito generations with a densonucleosis virus yields progressive survival improvement but persistent, innocuous infections. *Dev Comp Immunol* **30**:878–892. doi:10.1016/j.dci.2005.12.006
Rose JK. 1980. Complete intergenic and flanking gene sequences from the genome of vesicular stomatitis virus. *Cell* **19**:415–421.
Rose PP, Hanna SL, Spiridigliozzi A, Wannissorn N, Beiting DP, Ross SR, Hardy RW, Bambina SA, Heise MT, Cherry S. 2011. Natural Resistance-associated Macrophage Protein (NRAMP) is a cellular receptor for Sindbis virus in both insect and mammalian hosts. *Cell Host Microbe* **10**:97–104. doi:10.1016/j.chom.2011.06.009
Ross RJ, Weiner MM, Lin H. 2014. PIWI proteins and PIWI–interacting RNAs in the soma. *Nature* **505**:353–359. doi:10.1038/nature12987
Rozanov MN, Koonin EV, Gorbalenya AE. 1992. Conservation of the putative methyltransferase domain: a hallmark of the "Sindbis-like" supergroup of positive-strand RNA viruses. *J Gen Virol* **73 (Pt 8)**:2129–2134. doi:10.1099/0022-1317-73-8-2129
Rozhkov NV, Hammell M, Hannon GJ. 2013. Multiple roles for Piwi in silencing Drosophila transposons. *Genes & Development* **27**:400–412. doi:10.1101/gad.209767.112
Rubach JK, Wasik BR, Rupp JC, Kuhn RJ, Hardy RW, Smith JL. 2009. Characterization of purified Sindbis Virus nsP4 RNA-dependent RNA Polymerase activity in vitro. *Virology* **384**:201–208. doi:10.1016/j.virol.2008.10.030
Rubin E, Lithwick G, Levy AA. 2001. Structure and evolution of the hAT transposon superfamily. *Genetics* **158**:949–957.
Rupp JC, Sokoloski KJ, Gebhart NN, Hardy RW. 2015. Alphavirus RNA synthesis and non-structural protein functions. *J Gen Virol* **96**:2483–2500. doi:10.1099/jgv.0.000249
Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi MC. 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. *Nature* **461**:1296–1299. doi:10.1038/nature08501
Saito K, Ishizu H, Komai M, Kotani H, Kawamura Y, Nishida KM, Siomi H, Siomi MC. 2010. Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. *Genes Dev* **24**:2493–2498. doi:10.1101/gad.1989510
Saito K, Ishizuka A, Siomi H, Siomi MC. 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol* **3**:e235. doi:10.1371/journal.pbio.0030235
Saito K, Sakaguchi Y, Suzuki Takeo, Suzuki Tsutomu, Siomi H, Siomi MC. 2007. Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev* **21**:1603–1608. doi:10.1101/gad.1563607
Saldaña MA, Etebari K, Hart CE, Widen SG, Wood TG, Thangamani S, Asgari S, Hughes GL. 2017. Zika virus alters the microRNA expression profile and elicits an RNAi response in Aedes aegypti

mosquitoes. *PLOS Neglected Tropical Diseases* **11**:e0005760. doi:10.1371/journal.pntd.0005760
Saleh M-C, Tassetto M, van Rij RP, Goic B, Gausson V, Berry B, Jacquier C, Antoniewski C, Andino R. 2009. Antiviral immunity in Drosophila requires systemic RNA interference spread. *Nature* **458**:346–350. doi:10.1038/nature07712
Saleh M-C, van Rij RP, Hekele A, Gillis A, Foley E, O'Farrell PH, Andino R. 2006. The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nat Cell Biol* **8**:793–802. doi:10.1038/ncbl439
Salonen A, Vasiljeva L, Merits A, Magden J, Jokitalo E, Kääriäinen L. 2003. Properly folded nonstructural polyprotein directs the semliki forest virus replication complex to the endosomal compartment. *J Virol* **77**:1691–1702. doi:10.1128/jvi.77.3.1691-1702.2003
Sang R, Arum S, Chepkorir E, Mosomtai G, Tigoi C, Sigei F, Lwande OW, Landmann T, Affognon H, Ahlm C, Evander M. 2017. Distribution and abundance of key vectors of Rift Valley fever and other arboviruses in two ecologically distinct counties in Kenya. *PLoS Negl Trop Dis* **11**. doi:10.1371/journal.pntd.0005341
Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. *J Virol* **84**:9733–9748. doi:10.1128/JVI.00694-10
Sansone CL, Cohen J, Yasunaga A, Xu J, Osborn G, Subramanian H, Gold B, Buchon N, Cherry S. 2015. Microbiota-dependent priming of antiviral intestinal immunity in Drosophila. *Cell Host Microbe* **18**:571–581. doi:10.1016/j.chom.2015.10.010
Sanz MA, Madan V, Carrasco L, Nieva JL. 2003. Interfacial domains in Sindbis virus 6K protein. Detection and functional characterization. *J Biol Chem* **278**:2051–2057. doi:10.1074/jbc.M206611200
Sariola M, Saraste J, Kuismanen E. 1995. Communication of post-Golgi elements with early endocytic pathway: regulation of endoproteolytic cleavage of Semliki Forest virus p62 precursor. *J Cell Sci* **108 (Pt 6)**:2465–2475.
Schluckebier G, O'Gara M, Saenger W, Cheng X. 1995. Universal catalytic domain structure of AdoMet-dependent methyltransferases. *J Mol Biol* **247**:16–20. doi:10.1006/jmbi.1994.0117
Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**:863–864. doi:10.1093/bioinformatics/btr026
Schnettler E, Donald CL, Human S, Watson M, Siu RWC, McFarlane M, Fazakerley JK, Kohl A, Fragkoudis R. 2013a. Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *Journal of General Virology* **94**:1680–1689. doi:10.1099/vir.0.053850-0
Schnettler E, Ratinier M, Watson M, Shaw AE, McFarlane M, Varela M, Elliott RM, Palmarini M, Kohl A. 2013b. RNA interference targets arbovirus replication in Culicoides cells. *J Virol* **87**:2441–2454. doi:10.1128/JVI.02848-12
Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, Vlak JM, Kohl A, Khromykh AA, Pijlman GP. 2012. Noncoding flavivirus RNA displays RNA interference suppressor activity in insect and Mammalian cells. *J Virol* **86**:13486–13500. doi:10.1128/JVI.01104-12
Scott JC, Brackney DE, Campbell CL, Bondu-Hawkins V, Hjelle B, Ebel GD, Olson KE, Blair CD. 2010. Comparison of dengue virus type 2-specific small RNAs from RNA interference-competent and -incompetent mosquito cells. *PLoS Negl Trop Dis* **4**:e848. doi:10.1371/journal.pntd.0000848
Scotti PD, Dearing S, Mossop DW. 1983. Flock House virus: a nodavirus isolated from Costelytra zealandica (White) (Coleoptera: Scarabaeidae). *Arch Virol* **75**:181–189.
Senti K-A, Brennecke J. 2010. The piRNA Pathway: Guardian of the Genome –A Fly's Perspective. *Trends Genet* **26**:499–509. doi:10.1016/j.tig.2010.08.007
Sessegolo C, Burlet N, Haudry A. 2016. Strong phylogenetic inertia on genome size and transposable element content among 26 species of flies. *Biol Lett* **12**. doi:10.1098/rsbl.2016.0407

Shaw AE, Veronesi E, Maurin G, Ftaich N, Guiguen F, Rixon F, Ratinier M, Mertens P, Carpenter S, Palmarini M, Terzian C, Arnaud F. 2012. Drosophila melanogaster as a Model Organism for Bluetongue Virus Replication and Tropism. *Journal of Virology* **86**:9015–9024. doi:10.1128/JVI.00131-12

Shirako Y, Strauss JH. 1994. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J Virol* **68**:1874–1885.

Shirako Y, Strauss JH. 1990. Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology* **177**:54–64.

Shpiz S, Ryazansky S, Olovnikov I, Abramov Y, Kalmykova A. 2014. Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the drosophila germline. *PLoS Genet* **10**:e1004138. doi:10.1371/journal.pgen.1004138

Sienski G, Dönertas D, Brennecke J. 2012. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**:964–980. doi:10.1016/j.cell.2012.10.040

Silverman N, Zhou R, Stöven S, Pandey N, Hultmark D, Maniatis T. 2000. A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity. *Genes Dev* **14**:2461–2471.

Siomi MC, Sato K, Pezic D, Aravin AA. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. *Nature Reviews Molecular Cell Biology* **12**:nrm3089. doi:10.1038/nrm3089

Stabell M, Eskeland R, Bjørkmo M, Larsson J, Aalen RB, Imhof A, Lambertsson A. 2006. The Drosophila G9a gene encodes a multi-catalytic histone methyltransferase required for normal development. *Nucleic Acids Res* **34**:4609–4621. doi:10.1093/nar/gkl640

Stewart M. 2007. Molecular mechanism of the nuclear protein import cycle. *Nature Reviews Molecular Cell Biology* **8**:195–208. doi:10.1038/nrm2114

Strauss EG, Rice CM, Strauss JH. 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133**:92–110.

Strauss EG, Rice CM, Strauss JH. 1983. Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proc Natl Acad Sci U S A* **80**:5271–5275.

Strauss JH, Strauss EG. 1994. The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* **58**:491–562.

Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, Zhu L, Padhukasahasram B, Karlins E, Davis S, Jones PG, Quignon P, Johnson GS, Parker HG, Fretwell N, Mosher DS, Lawler DF, Satyaraj E, Nordborg M, Lark KG, Wayne RK, Ostrander EA. 2007. A single IGF1 allele is a major determinant of small size in dogs. *Science* **316**:112–115. doi:10.1126/science.1137045

Suzuki Y, Frangeul L, Dickson LB, Blanc H, Verdier Y, Vinh J, Lambrechts L, Saleh M-C. 2017. Uncovering the Repertoire of Endogenous Flaviviral Elements in Aedes Mosquito Genomes. *J Virol* **91**. doi:10.1128/JVI.00571-17

Swevers L, Liu J, Smagghe G. 2018. Defense Mechanisms against Viral Infection in Drosophila: RNAi and Non-RNAi. *Viruses* **10**. doi:10.3390/v10050230

Tabachnick WJ. 1994. Genetics of Insect Vector Competence for Arboviruses In: Harris KF, editor. Advances in Disease Vector Research, Advances in Disease Vector Research. New York, NY: Springer New York. pp. 93–108. doi:10.1007/978-1-4612-2590-4_4

Tanji T, Ip YT. 2005. Regulators of the Toll and Imd pathways in the Drosophila innate immune response. *Trends in Immunology* **26**:193–198. doi:10.1016/j.it.2005.02.006

Tassetto M, Kunitomi M, Andino R. 2017. Circulating Immune Cells Mediate a Systemic RNAi-Based Adaptive Antiviral Response in Drosophila. *Cell* **169**:314-325.e13. doi:10.1016/j.cell.2017.03.033

Taylor RM, Hurlbut HS, Work TH, Kingston JR, Frothingham TE. 1955. Sindbis virus: a newly recognized arthropodtransmitted virus. *Am J Trop Med Hyg* **4**:844–862.

Teixeira L, Ferreira Á, Ashburner M. 2008. The Bacterial Symbiont Wolbachia Induces Resistance to RNA Viral Infections in Drosophila melanogaster. *PLoS Biol* **6**. doi:10.1371/journal.pbio.1000002

Teninges D, Bras F, Dezélée S. 1993. Genome organization of the sigma rhabdovirus: six genes and a gene overlap. *Virology* **193**:1018–1023. doi:10.1006/viro.1993.1219
Teninges D, Ohanessian A, Richard-Molard C, Contamine D. 1979. Isolation and Biological Properties of Drosophila X Virus. *Journal of General Virology* **42**:241–254. doi:10.1099/0022-1317-42-2-241
Terzian C, Pélisson A, Bucheton A. 2001. Evolution and phylogeny of insect endogenous retroviruses. *BMC Evol Biol* **1**:3. doi:10.1186/1471-2148-1-3
Tingström O, Wesula Lwande O, Näslund J, Spyckerelle I, Engdahl C, Von Schoenberg P, Ahlm C, Evander M, Bucht G. 2016. Detection of Sindbis and Inkoo Virus RNA in Genetically Typed Mosquito Larvae Sampled in Northern Sweden. *Vector Borne Zoonotic Dis* **16**:461–467. doi:10.1089/vbz.2016.1940
Tomari Y, Du T, Zamore PD. 2007. Sorting of Drosophila Small Silencing RNAs. *Cell* **130**:299–308. doi:10.1016/j.cell.2007.05.057
Touret F, Guiguen F, Greenland T, Terzian C. 2014a. In between: gypsy in Drosophila melanogaster reveals new insights into endogenous retrovirus evolution. *Viruses* **6**:4914–4925. doi:10.3390/v6124914
Touret F, Guiguen F, Terzian C. 2014b. Wolbachia Influences the Maternal Transmission of the gypsy Endogenous Retrovirus in Drosophila melanogaster. *mBio* **5**. doi:10.1128/mBio.01529-14
Treiber CD, Waddell S. 2017. Resolving the prevalence of somatic transposition in Drosophila. *eLife* **6**:e28297. doi:10.7554/eLife.28297
Tsai CW, McGraw EA, Ammar E-D, Dietzgen RG, Hogenhout SA. 2008. Drosophila melanogaster Mounts a Unique Immune Response to the Rhabdovirus Sigma virus. *Appl Environ Microbiol* **74**:3251–3256. doi:10.1128/AEM.02248-07
Tsujimoto H, Anderson MAE, Myles KM, Adelman ZN. 2018. Identification of Candidate Iron Transporters From the ZIP/ZnT Gene Families in the Mosquito Aedes aegypti. *Front Physiol* **9**. doi:10.3389/fphys.2018.00380
Tüfekci KU, Oner MG, Meuwissen RLJ, Genç S. 2014. The role of microRNAs in human diseases. *Methods Mol Biol* **1107**:33–50. doi:10.1007/978-1-62703-748-8_3
Ullastres A, Merenciano M, González J. 2019. Natural transposable element insertions drive expression changes in genes underlying Drosophila immune response. *bioRxiv* 655225. doi:10.1101/655225
Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. 2006. A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline. *Science* **313**:320–324. doi:10.1126/science.1129333
van Cleef KWR, van Mierlo JT, Miesen P, Overheul GJ, Fros JJ, Schuster S, Marklewitz M, Pijlman GP, Junglen S, van Rij RP. 2014. Mosquito and Drosophila entomobirnaviruses suppress dsRNA- and siRNA-induced RNAi. *Nucleic Acids Res* **42**:8732–8744. doi:10.1093/nar/gku528
Van den Broeck D, Maes T, Sauer M, Zethof J, De Keukeleire P, D'hauw M, Van Montagu M, Gerats T. 1998. Transposon Display identifies individual transposable elements in high copy number lines. *Plant J* **13**:121–129. doi:10.1046/j.1365-313X.1998.00004.x
van Rij RP, Saleh M-C, Berry B, Foo C, Houk A, Antoniewski C, Andino R. 2006. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. *Genes Dev* **20**:2985–2995. doi:10.1101/gad.1482006
Varjak M, Dietrich I, Sreenu VB, Till BE, Merits A, Kohl A, Schnettler E. 2018a. Spindle-E Acts Antivirally Against Alphaviruses in Mosquito Cells. *Viruses* **10**. doi:10.3390/v10020088
Varjak M, Leggewie M, Schnettler E. 2018b. The antiviral piRNA response in mosquitoes? *J Gen Virol* **99**:1551–1562. doi:10.1099/jgv.0.001157
Varjak M, Maringer K, Watson M, Sreenu VB, Fredericks AC, Pondeville E, Donald CL, Sterk J, Kean J, Vazeille M, Failloux A-B, Kohl A, Schnettler E. 2017. Aedes aegypti Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral Responses. *mSphere* **2**. doi:10.1128/mSphere.00144-17
Vasudevan S. 2012. Posttranscriptional upregulation by microRNAs. *Wiley Interdiscip Rev RNA* **3**:311–330. doi:10.1002/wrna.121

Venkataraman S, Prasad BVLS, Selvarajan R. 2018. RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses* **10**. doi:10.3390/v10020076
Vidal NM, Ludwig A, Loreto ELS. 2009. Evolution of Tom, 297, 17.6 and rover retrotransposons in Drosophilidae species. *Mol Genet Genomics* **282**:351–362. doi:10.1007/s00438-009-0468-0
Vieira C, Biémont C. 2004. Transposable element dynamics in two sibling species: Drosophila melanogaster and Drosophila simulans. *Genetica* **120**:115–123.
Vieira C, Lepetit D, Dumont S, Biémont C. 1999. Wake up of transposable elements following Drosophila simulans worldwide colonization. *Mol Biol Evol* **16**:1251–1255. doi:10.1093/oxfordjournals.molbev.a026215
Vodovar N, Bronkhorst AW, van Cleef KWR, Miesen P, Blanc H, van Rij RP, Saleh M-C. 2012. Arbovirus-Derived piRNAs Exhibit a Ping-Pong Signature in Mosquito Cells. *PLoS ONE* **7**:e30861. doi:10.1371/journal.pone.0030861
Vogel E, Santos D, Mingels L, Verdonckt T-W, Broeck JV. 2019. RNA Interference in Insects: Protecting Beneficials and Controlling Pests. *Front Physiol* **9**. doi:10.3389/fphys.2018.01912
Volff J-N. 2006. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays* **28**:913–922. doi:10.1002/bies.20452
von Bonsdorff CH, Harrison SC. 1975. Sindbis virus glycoproteins form a regular icosahedral surface lattice. *J Virol* **16**:141–145.
Wahlberg JM, Boere WA, Garoff H. 1989. The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to low pH during virus maturation. *J Virol* **63**:4991–4997.
Wang X-H, Aliyari R, Li W-X, Li H-W, Kim K, Carthew R, Atkinson P, Ding S-W. 2006. RNA interference directs innate immunity against viruses in adult Drosophila. *Science* **312**:452–454. doi:10.1126/science.1125694
Wang Y, Jin B, Liu P, Li J, Chen X, Gu J. 2018. piRNA Profiling of Dengue Virus Type 2-Infected Asian Tiger Mosquito and Midgut Tissues. *Viruses* **10**. doi:10.3390/v10040213
Wayne ML, Blohm GM, Brooks ME, Regan KL, Brown BY, Barfield M, Holt RD, Bolker BM. 2011. The prevalence and persistence of sigma virus, a biparentally transmitted parasite of Drosophila melanogaster. *Evol Ecol Res* **13**:323–345.
Webster CL, Waldron FM, Robertson S, Crowson D, Ferrari G, Quintana JF, Brouqui J-M, Bayne EH, Longdon B, Buck AH, Lazzaro BP, Akorli J, Haddrill PR, Obbard DJ. 2015. The Discovery, Distribution, and Evolution of Viruses Associated with Drosophila melanogaster. *PLoS Biol* **13**. doi:10.1371/journal.pbio.1002210
Weiss B, Nitschko H, Ghattas I, Wright R, Schlesinger S. 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. *Journal of Virology* **63**:5310–5318.
Wessler SR. 1996. Plant retrotransposons: Turned on by stress. *Current Biology* **6**:959–961. doi:10.1016/S0960-9822(02)00638-3
Whitfield ZJ, Dolan PT, Kunitomi M, Tassetto M, Seetin MG, Oh S, Heiner C, Paxinos E, Andino R. 2017. The Diversity, Structure, and Function of Heritable Adaptive Immunity Sequences in the Aedes aegypti Genome. *Curr Biol* **27**:3511-3519.e7. doi:10.1016/j.cub.2017.09.067
Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH. 2007. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* **8**:973–982. doi:10.1038/nrg2165
Wu Q, Luo Y, Lu R, Lau N, Lai EC, Li W-X, Ding S-W. 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *PNAS* **107**:1606–1611. doi:10.1073/pnas.0911353107
Xu J, Cherry S. 2014. Viruses and Antiviral Immunity in Drosophila. *Dev Comp Immunol* **42**. doi:10.1016/j.dci.2013.05.002
Yamaguchi M, Yoshida H. 2018. Drosophila as a Model Organism. *Adv Exp Med Biol* **1076**:1–10. doi:10.1007/978-981-13-0529-0_1
Yamashiro H, Siomi MC. 2018. PIWI-Interacting RNA in Drosophila: Biogenesis, Transposon Regulation, and Beyond. *Chem Rev* **118**:4404–4421. doi:10.1021/acs.chemrev.7b00393

Yampolsky LY, Webb C, Shabalina SA, Kondrashov AS. 1999. Rapid accumulation of a vertically transmitted parasite triggered by relaxation of natural selection among hosts.
Yan Z, Hu HY, Jiang X, Maierhofer V, Neb E, He L, Hu Y, Hu H, Li N, Chen W, Khaitovich P. 2011. Widespread expression of piRNA-like molecules in somatic tissues. *Nucleic Acids Res* **39**:6596–6607. doi:10.1093/nar/gkr298
Yashiro R, Murota Y, Nishida KM, Yamashiro H, Fujii K, Ogai A, Yamanaka S, Negishi L, Siomi H, Siomi MC. 2018. Piwi Nuclear Localization and Its Regulatory Mechanism in Drosophila Ovarian Somatic Cells. *Cell Reports* **23**:3647–3657. doi:10.1016/j.celrep.2018.05.051
Yin H, Lin H. 2007. An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. *Nature* **450**:304–308. doi:10.1038/nature06263
Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ. 2015. Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* **350**:339–342. doi:10.1126/science.aab0700
Zacks MA, Paessler S. 2010. Encephalitic alphaviruses. *Vet Microbiol* **140**:281–286. doi:10.1016/j.vetmic.2009.08.023
Zaidman-Rémy A, Poidevin M, Hervé M, Welchman DP, Paredes JC, Fahlander C, Steiner H, Mengin-Lecreulx D, Lemaitre B. 2011. Drosophila Immunity: Analysis of PGRP-SB1 Expression, Enzymatic Activity and Function. *PLoS One* **6**. doi:10.1371/journal.pone.0017231
Zambon RA, Nandakumar M, Vakharia VN, Wu LP. 2005. The Toll pathway is important for an antiviral response in Drosophila. *Proc Natl Acad Sci USA* **102**:7257–7262. doi:10.1073/pnas.0409181102
Zambon RA, Vakharia VN, Wu LP. 2006. RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. *Cellular Microbiology* **8**:880–889. doi:10.1111/j.1462-5822.2006.00688.x
Zettervall C-J, Anderl I, Williams MJ, Palmer R, Kurucz E, Ando I, Hultmark D. 2004. A directed screen for genes involved in Drosophila blood cell activation. *Proc Natl Acad Sci USA* **101**:14192–14197. doi:10.1073/pnas.0403789101
Zhang X, Fugère M, Day R, Kielian M. 2003. Furin processing and proteolytic activation of Semliki Forest virus. *J Virol* **77**:2981–2989. doi:10.1128/jvi.77.5.2981-2989.2003
Zuo L, Wang Z, Tan Y, Chen X, Luo X. 2016. piRNAs and Their Functions in the Brain. *Int J Hum Genet* **16**:53–60.

ANNEXES

Article en preparation pour une soumission au journal *Nature (Letter).*

The siRNA and piRNA pathways interplay and modulate transposable element transcription upon viral infection in *Drosophila*

Marlène Roy[1,2,3], Barbara Viginier[2], Edouard Saint-Michel[3], Frédérick Arnaud[2], Maxime Ratinier[2] and Marie Fablet[1]

[1]Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon; Université Lyon 1; CNRS; UMR 5558, Villeurbanne 69622, France

[2] EPHE, PSL Research University, INRA, Univ Lyon, Université Claude Bernard Lyon1, UMR754, IVPC, F-69007, Lyon, France

[3]LabEx Ecofect, Université de Lyon, Lyon, France

Abstract

Transposable elements (TEs) are genomic parasites that are found in all genomes, some of which display structure similarity to certain viruses. In insects, TEs are controlled by the piRNA pathway in gonads, and the siRNA pathway is dedicated to TE somatic control and defense against viruses. These two small interfering RNA pathways are led by distinct molecular effectors and described as independent. However, similarities in structure and control mechanisms across TEs and viruses suggest that an interplay may exist. Using Sindbis virus (SINV) in *Drosophila*, here we show that viral infection affects TE transcript amounts via modulation of the piRNA and siRNA repertoires. Our data suggest that the modulation results from shared RNA substrates and, potentially, common trans-regulators of the small RNA pathways. These results are of fundamental importance in evolutionary genomics, suggesting that viral epidemics or chronic infections can impact TE activity and thus the rate of subsequent genetic diversification.

The boundary between TEs and viruses is sometimes faint as illustrated by the gypsy endogenous retrovirus (ERV) in *Drosophila*. Gypsy is vertically transmitted following Mendel's laws, but also proved to have infectious properties[1]. In *Drosophila* ovaries, gypsy is silenced by the well-described flamenco piRNA cluster[2]. The activity and copy numbers of such ERVs are widely variable within species, as we have shown using the tirant ERV of *Drosophila simulans*[3]. Here, we used Makindu, a *D. simulans* strain that we previously deeply characterized for its transposable elements (TEs)[3–5]. We first treated Makindu flies with

tetracycline to get rid of the *Wolbachia* endosymbiont, which is known to interfere with viral replication[6,7]. Then, we infected female adults with Sindbis virus (SINV) by intrathoracic injections. SINV is an arbovirus of the *Togaviridae* family (single stranded RNA genome of positive polarity), which naturally infects mosquitoes[8]. We assessed SINV replication in this *D. simulans* strain through TCID50 assays. SINV indeed replicated in carcasses (whole body without ovaries) (Fig. 1a), but was not associated with fly mortality. This led to the production of SINV-directed 21 nt-long small RNAs (Extended data Fig. 1). However, SINV barely replicated in ovaries (Fig. 1a), the tissue where the piRNA pathway is described to control TEs in *Drosophila*. These results were confirmed by immunostaining using an anti-SINV capsid protein antibody (Fig. 1b,c).

In order to assess TE transcriptional activity and control across conditions, we decided to perform RNA-seq and small RNA-seq from samples extracted six days post infection (6dpi), which corresponds to the exponential phase of viral replication in carcasses (Fig. 1a). Samples were produced from ovary and carcass tissues. As reported[9], we used a small RNA preparation protocol that ensures the isolation of small RNAs involved in nucleoproteic complexes, and therefore excludes mRNA degradation products. We separated reads corresponding to 23-30 nt-long RNAs and 21 nt-long RNAs, and considered them to be piRNAs and siRNAs, respectively. siRNAs are found in all parts of the fly body, while piRNAs are considered to be restricted to the germline in *Drosophila*[10]. However, somatic piRNAs are recurrently described[11] (and even with evidence of efficient ping-pong[12] production), but are potentially restricted to certain cell types or tissues (for instance, somatic piRNAs are found in *Drosophila melanogaster* heads[13] but absent in thorax[10]). Here, the analysis of read counts for 237 annotated TE families and subfamilies revealed a significant global reduction of TE transcript read counts in carcasses upon infection compared to the non-infected samples (mock) (Fig. 1d). These results were confirmed using RT-qPCR (Extended data Fig. 2). The global decrease in TE transcript amounts was associated with a significant increase in TE-directed 23-30 nt (Fig. 1e) and 21 nt-long small RNAs (Fig. 1f). On the contrary, in ovaries, no significant shift in TE transcript amounts upon infection could be detected (Fig. 1g). However, in this organ, we found that TE-directed small RNAs were notably modulated: TE-directed 23-30 nt small RNAs were less abundant (Fig. 1h) while TE-directed 21 nt small RNA mounts increased (Fig. 1i). This mirror-like response between both small interfering RNA classes may explain the resulting absence of modulation of TE transcript amounts. Altogether, our results show for the first time that TE transcripts and TE-directed small RNAs are modulated upon viral infection in *Drosophila*.

However, our RNA-seq analysis also revealed that the Makindu strain was chronically infected by Nora virus (Extended data Fig. 3), which may interfere with SINV infection. Therefore, in order to control for chronic infections, and go deeper into the mechanisms allowing such a modulation in TE transcript amounts upon infection, we decided to switch to the *D. melanogaster* model, which has available a wealth of mutants. We treated all strains with tetracycline to get rid of Wolbachia, and bleached the embryos to remove viral particles

coming from chronic infections[14]. We infected w^{1118} adult females with SINV by intrathoracic injections, and retrieved comparable results to those we observed using Makindu carcasses, as described below, although using a smaller viral dose. Similar to what we observed in Makindu, SINV replicated in carcasses (Fig. 2a), TE transcript amounts decreased upon infection (Fig. 2b, black dots), and this TE modulation was associated with an increase in TE-directed 23-30 nt small RNAs (Fig. 2c, black dots) and 21 nt small RNAs (Fig. 2d, black dots).

It was previously demonstrated that viral response through siRNAs depended on the systemic diffusion of dsRNA viral intermediates, which then triggers the production of siRNAs[15]. The systemic diffusion relies on the dsRNA uptake pathway, of which CG4572 is an effector[15,16]. As we did for the w^{1118} strain, we infected $CG4572^{-/-}$ female adults with SINV by intrathoracic injections. The used $CG4572^{c05963}$ mutation was obtained from a w^{1118} genetic background[15]; therefore, the $CG4572^{-/-}$ strain harbors the same pool of TE insertions as w^{1118}. The analysis revealed that the modulation in TE-directed 23-30 nt and 21 nt small RNAs upon infection virtually disappeared in the $CG4572^{-/-}$ mutant carcasses, compared to wild-type (Fig. 2c,d). This indicates that the TE modulation relies on the dsRNA uptake pathway in carcasses. Therefore, upon infection, when the dsRNA uptake pathway allows systemic diffusion of viral RNAs, we propose that this machinery also opportunistically carries TE dsRNA intermediates and allows their diffusion. This triggers the systemic production of viral siRNAs as well as TE-directed siRNAs, which subsequently leads to the reduction of TE transcript amounts in carcasses.

These results suggest that the TE transcript modulation we observe may first rely on TE-directed siRNA increase. Therefore, we next investigated the consequences of the impairment of siRNA production, using a mutant of *dcr2*, which is responsible for siRNA formation. As expected, SINV replicated more abundantly in this mutant (Fig. 2a). Surprisingly, this abundant SINV replication was associated with a high production of SINV-directed 23-30 nt small RNAs, however only in the sense orientation (Extended data Fig. 4). In the $dcr2^{-/-}$ genotype, SINV infection led to an increase in TE transcript amounts (Fig. 2b, red dots), and a decrease in TE-directed 23-30 nt RNAs (Fig. 2c, red dots). We propose that the reduction in TE-directed 23-30 nt RNAs is due to the piRNA machinery being saturated by the abundant viral RNAs, as observed above. In addition, in the infection context, we observed that the $dcr2^{-/-}$ mutant also displayed a strong repression of the *piwi* gene (Fig. 2e; log2 Fold Change = -2.12; p-value = 0.001), associated with a strong induction of *dcr2* transcription (which, however, does not lead to more Dcr-2 function since the coding sequence is mutated) compared to w^{1118} (Fig. 2e; log2 Fold Change = 2.80; p-value = 6.6e-34).

Considering all these results, we propose that the observed cross-talk between the siRNA and piRNA pathways in carcasses is due to TE-directed siRNAs and piRNAs sharing RNA precursors, at least in part (Fig. 2f), as strengthened by the strong positive correlation between TE-directed siRNAs and piRNAs that we find in all conditions (Extended data Fig. 5). Indeed, it was previously shown that TEs produce sense and antisense transcripts (allowing the

formation of dsRNAs) and that mixed ping-pong pairs of piRNAs and siRNAs exist[17,18]. In addition, when the siRNA pathway is less efficient (as exemplified in the extreme by the dcr2$^{-/-}$ mutant), we propose that there exist certain so-far unidentified host factors that are able both to stimulate the siRNA pathway (increasing the expression of *dcr2*) and repress the piRNA pathway (decreasing the expression of *piwi*). We suggest that such fine-tuning would have been selected to allow to deploy the RNA interference efforts towards viral control upon infection because it represents a more acute threat in carcasses compared to TE activity.

In ovaries, similar to our observations in Makindu, SINV barely replicated in w^{1118} and CG4572$^{-/-}$. However, viral particles were detected in dcr2$^{-/-}$ (Fig. 3a). While TE transcripts were also barely modulated in w^{1118} (Fig. 3b, black dots), TE-directed small RNAs were less abundant upon infection (Fig. 3c,d, black dots), which contrasted with Makindu patterns (Fig. 1h,i). CG4572$^{-/-}$ and w^{1118} patterns largely overlapped (Fig. 3c,d), suggesting that the dsRNA uptake pathway does not determine TE-directed small RNA amounts in ovaries. On the contrary, TE-directed 23-30 nt RNA amounts increased in the dcr2$^{-/-}$ mutant (Fig. 3c, red dots), a result that we do not fully explain so far, also considering that genes of the piRNA and siRNA pathways were not differentially expressed upon infection (Extended data Fig. 6). Nevertheless, what is very interesting from an evolutionary perspective and deserves future investigation, is that TE transcript amounts were virtually stable in ovaries upon infection both in Makindu and w^{1118}, while TE-directed small RNAs showed completely different patterns (strong mirror-like pattern in Makindu vs weak reduction of both classes in w^{1118}). This suggests that there is a strong selective pressure on the maintenance of TE transcriptome stability in ovaries, however achieved by distinct mechanisms in *D. simulans* and *D. melanogaster*. In addition, future investigations are needed to explore the impacts of the modulation of the ovarian piRNA amounts that are transmitted to the next generation in *D. simulans*.

Here, we demonstrate that SINV infection affects TE transcript amounts in the soma via a modulation of TE-directed small RNA amounts. While the piRNA and siRNA pathways are described as relying on independent effectors, we propose that an interplay exists between both pathways because they share their RNA substrates (at least in part). This is the first demonstration of the effects of viral infection on TE activity. The next step is clearly to investigate whether transposition rates are affected by these variations in TE transcript amounts. These results are of fundamental importance because they indicate that viral infections, and epidemics in particular, may affect mutation rates through the modulation of TE activity, and therefore the speed of genetic diversification, as was reported to be particularly important in the brain, for instance[13].

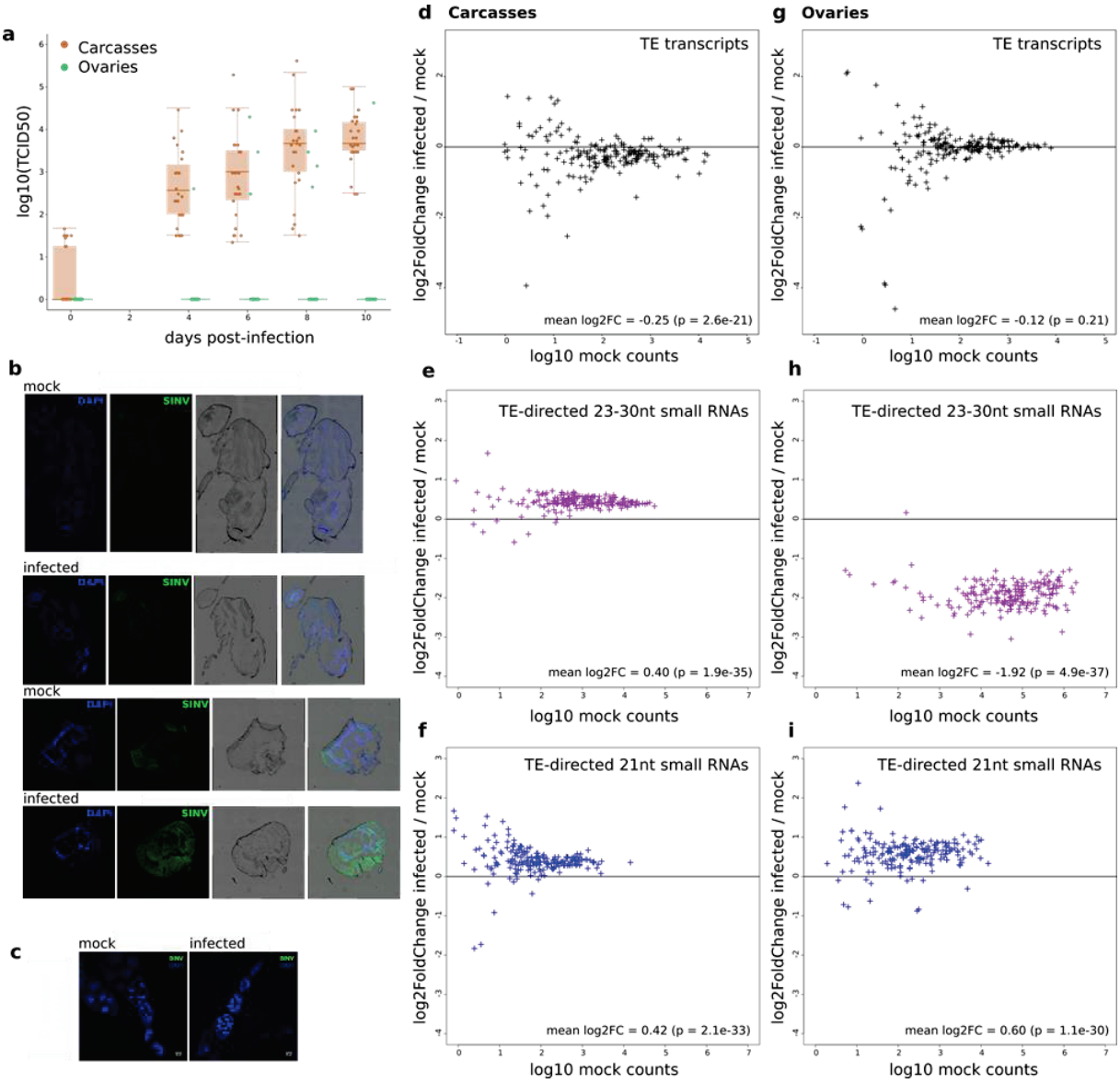


Figure 1 | SINV infection in *D. simulans* Makindu flies. a, Viral titration assessed by TCID50 assays in ovaries (green) and carcasses (brown). **b,** SINV tropism in carcasses as revealed by an anti-capsid antibody. SINV is mainly detected in heads and thorax. **c,** SINV could not be detected in ovaries using immunostaining. **d, g,** TE transcript modulation upon SINV infection, expressed as log2 (infected counts / mock counts) for each annotated TE family, in carcasses and ovaries, respectively. **e, h,** TE-directed 23-30 nt long small RNA modulation upon SINV infection, in carcasses and ovaries, respectively, and **f, i,** TE-directed 21 nt long small RNA modulation upon SINV infection, in carcasses and ovaries, respectively. Statistical significance of shifts assessed by comparing infected counts to mock counts using Wilcoxon paired tests.

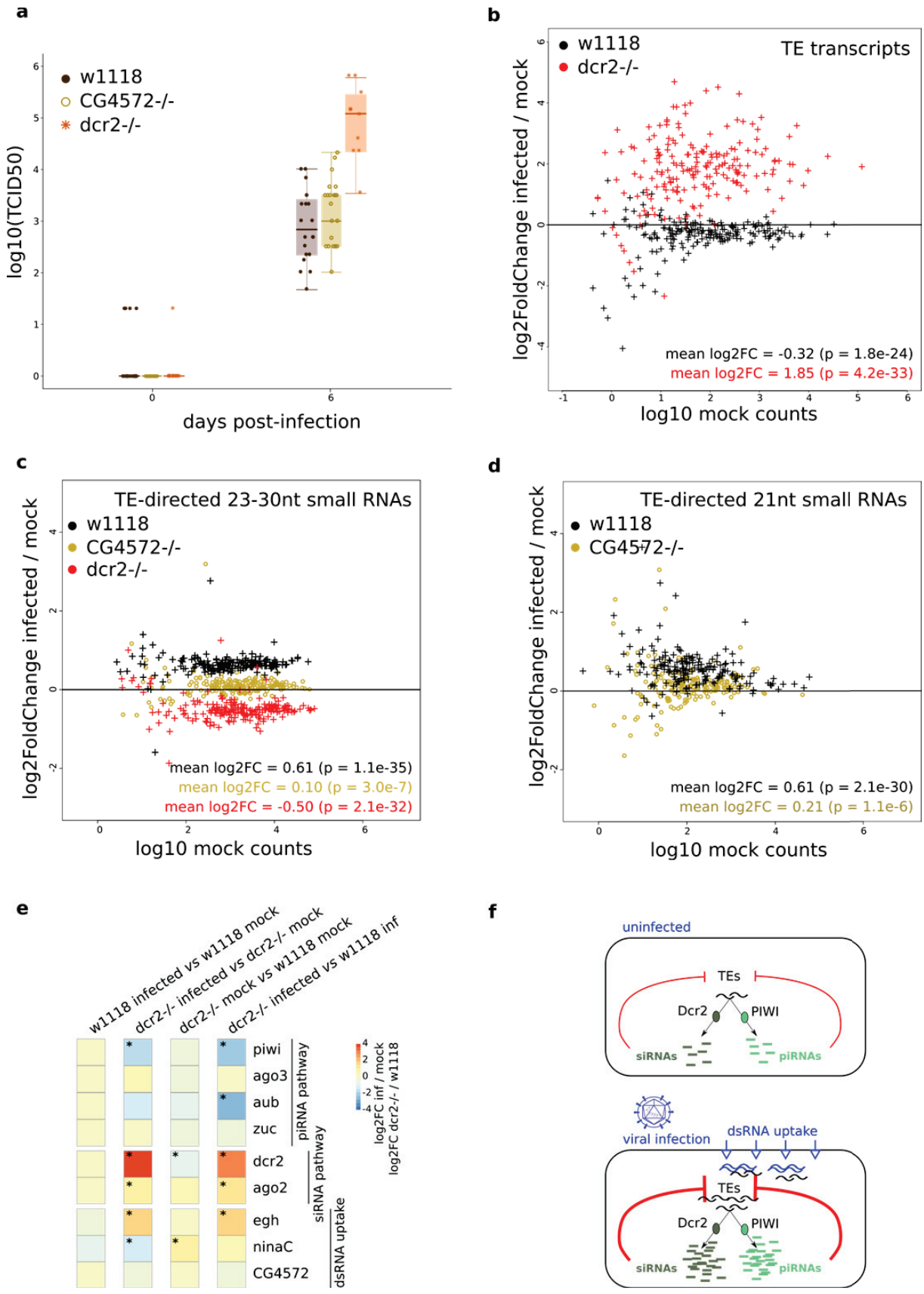


Figure 2 | SINV infection in *D. melanogaster*, carcasses. a, Viral titration assessed by TCID50 assays in the three strains. **b,** TE transcript modulation upon infection expressed as log2 (infected counts / mock counts) for each annotated TE family. **c,** TE-directed 23-30 nt long

small RNA modulation upon infection. **d,** TE-directed 21 nt small RNA modulation upon infection. Statistical significance of shifts assessed by comparing infected counts to mock counts using Wilcoxon paired tests. **e,** Modulation of transcript levels for a subset of genes involved in the piRNA, siRNA, and dsRNA uptake pathways, as assessed by DESeq2. Significant shifts (adjusted p-value < 0.05) are indicated by stars. **f,** Model for TE modulation in carcasses upon infection.

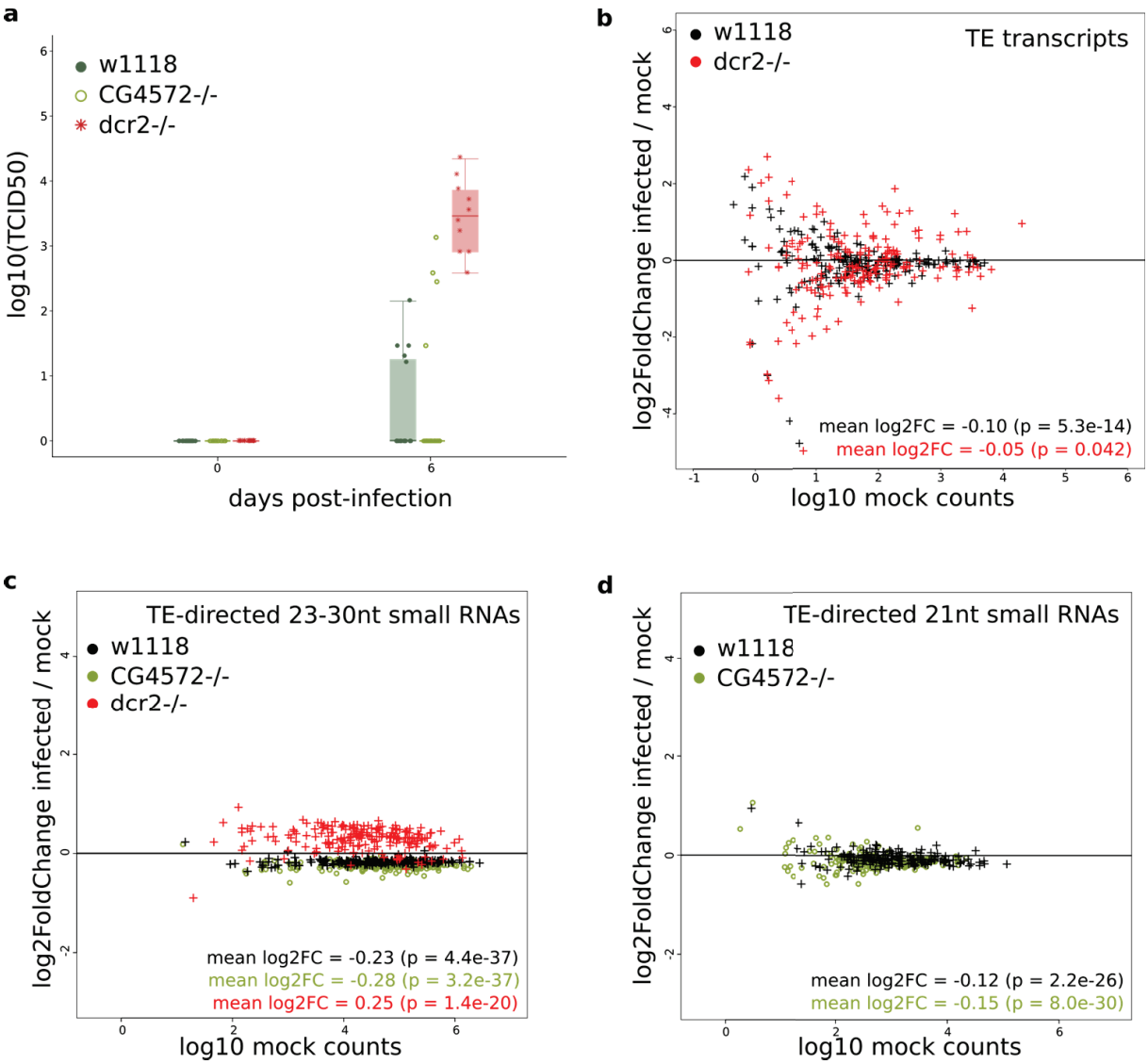


Figure 3 | SINV infection in *D. melanogaster*, ovaries. a, Viral titration assessed by TCID50 assays in the three strains. The signal detected in the dcr2$^{-/-}$ mutant results from viral particles surrounding the ovaries, but not entering the tissue, as assessed by immunostaining (Extended data Fig. 7). **b,** TE transcript modulation upon infection expressed as log2 (infected counts / mock counts) for each annotated TE family. **c,** TE-directed 23-30 nt long small RNA modulation upon infection. **d,** TE-directed 21 nt small RNA modulation upon infection. Statistical significance of shifts assessed by comparing infected counts to mock counts using Wilcoxon paired tests.

Acknowledgments

This work is dedicated to the memory of Christophe Terzian, who initiated this project but left us too early.

This work was performed using the computing facilities of the CC LBBE/PRABI, the Arthro-tools platform, and the Nanodrop 2000 Thermofisher of IGFL (managed by Benjamin Gillet, who we thank for his availability). We thank Carla Saleh, who kindly provided us with *D. melanogaster* strains. Sequencing was performed by the GenomEast platform, a member of the "France Génomique"consortium (ANR-10-INBS-0009).

We thank Cristina Vieira, Rita Rebollo, Matthieu Boulesteix, and Hélène Henri for helpul discussions. We thank Judit Salces-Ortiz, Justine Picarle, Vincent Merel, and BF2I for technical help. We also thank Sévérine Chambeyron, François Leulier, and Philippe Marianneau for helpful discussions.

This work was funded by Labex ECOFECT (grant ERMIT) and ANR (grant ExHyb).

References

1. Kim, A. et al. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 91, 1285–1289 (1994).

2. Pélisson, A. et al. Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. EMBO J. 13, 4401–4411 (1994).

3. Akkouche, A. et al. tirant, a newly discovered active endogenous retrovirus in Drosophila simulans. J. Virol. 86, 3675–3681 (2012).

4. Akkouche, A. et al. Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. EMBO Rep. 14, 458–464 (2013).

5. Fablet, M. et al. Dynamic Interactions Between the Genome and an Endogenous Retrovirus: Tirant in Drosophila simulans Wild-Type Strains. G3 Bethesda Md 9, 855–865 (2019).

6. Teixeira, L., Ferreira, A. & Ashburner, M. The bacterial symbiont Wolbachia induces resistance to RNA viral infections in Drosophila melanogaster. PLoS Biol. 6, e2 (2008).

7. Bhattacharya, T., Newton, I. L. G. & Hardy, R. W. Wolbachia elevates host methyltransferase expression to block an RNA virus early during infection. PLoS Pathog. 13, e1006427 (2017).

8. Taylor, R. M., Hurlbut, H. S., Work, T. H., Kingston, J. R. & Frothingham, T. E. Sindbis virus: a newly recognized arthropodtransmitted virus. Am. J. Trop. Med. Hyg. 4, 844–862 (1955).

9. Grentzinger, T. et al. A user-friendly chromatographic method to purify small regulatory RNAs. Methods San Diego Calif 67, 91–101 (2014).

10. Lewis, S. H. et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. Nat. Ecol. Evol. 2, 174–181 (2018).

11. Mirkovic-Hösle, M. & Förstemann, K. Transposon defense by endo-siRNAs, piRNAs and somatic pilRNAs in Drosophila: contributions of Loqs-PD and R2D2. PloS One 9, e84994 (2014).

12. Brennecke, J. et al. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103 (2007).

13. Perrat, P. N. et al. Transposition-driven genomic heterogeneity in the Drosophila brain. Science 340, 91–95 (2013).

14. Merkling, S. H. & van Rij, R. P. Analysis of resistance and tolerance to virus infection in Drosophila. Nat. Protoc. 10, 1084–1097 (2015).

15. Saleh, M.-C. et al. Antiviral immunity in Drosophila requires systemic RNA interference spread. Nature 458, 346–350 (2009).

16. Saleh, M.-C. et al. The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat. Cell Biol. 8, 793–802 (2006).

17. Shpiz, S., Ryazansky, S., Olovnikov, I., Abramov, Y. & Kalmykova, A. Euchromatic transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the drosophila germline. PLoS Genet. 10, e1004138 (2014).

18. Olovnikov, I. et al. De novo piRNA cluster formation in the Drosophila germ line triggered by transgenes containing a transcribed transposon fragment. Nucleic Acids Res. 41, 5757–5768 (2013).

19. Tomari, Y., Du, T. & Zamore, P. D. Sorting of Drosophila small silencing RNAs. Cell 130, 299–308 (2007).

20. Riegler, M., Sidhu, M., Miller, W. J. & O'Neill, S. L. Evidence for a global Wolbachia replacement in Drosophila melanogaster. Curr. Biol. CB 15, 1428–1433 (2005).

21. Goic, B. et al. RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model Drosophila. Nat. Immunol. 14, 396–403 (2013).

22. Reed, L. J. & Muench, H. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS. Am. J. Epidemiol. 27, 493–497 (1938).

23. Greiser-Wilke, I., Moenning, V., Kaaden, O. R. & Figueiredo, L. T. Most alphaviruses share a conserved epitopic region on their nucleocapsid protein. J. Gen. Virol. 70 (Pt 3), 743–748 (1989).

24. Modolo, L. & Lerat, E. UrQt: an efficient software for the Unsupervised Quality trimming of NGS data. BMC Bioinformatics 16, 137 (2015).

25. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36 (2013).

26. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. 25, 2078–2079 (2009).

27. Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L. & Pachter, L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol. 12, R22 (2011).

28. Lerat, E., Fablet, M., Modolo, L., Lopez-Maestre, H. & Vieira, C. TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. Nucleic Acids Res. 45, e17 (2017).

29. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

30. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12 (2011).

31. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. Bioinforma. Oxf. Engl. 27, 863–864 (2011).

Methods

Fly strains and husbandry

We used the *D. simulans* wild-type strain Makindu, and the following *D. melanogaster* strains: w^{1118} (wild-type, kindly provided by M-C Saleh), *Dcr-2*$^{[L811fsX]}$ mutant[19] (kindly provided by M-C Saleh), and the *CG4572*^{c05963} mutant stock (Bloomington Drosophila Stock Center). Flies were reared on a standard corn medium at 25°C as isofemale lines. To eliminate the endosymbiotic bacteria *Wolbachia*, flies were reared during two generations on standard agarose medium containing 0.25 mg/mL tetracycline hypochloride (Sigma-Aldrich). The absence of *Wolbachia* was validated using *Wolbachia* 16S primers (5' TTGTAGCCTGCTATGGTATAACT 3' and 5' GAATAGGTATGATTTTCATGT 3'), and *Wolbachia wsp* primers (5' TGGTCCAATAAGTGATGAAGAAAC 3' and 5' AAAAATTAAACGCTACTCCA 3'), as described previously[20]. As described[14], *D. melanogaster* eggs were dechorionated to eliminate chronic infections with 50% (vol/vol) household bleach solution for 10 min, washed with water, and transferred to fresh medium during two generations before experiments. The subsequent absence of the viruses was confirmed using RT-PCR and primers already described[14]. All experiments were performed using 3-6 day-old female flies in accordance with previous studies[15,21].

Virus production and titration

SINV stocks were produced by reverse genetics on BHK-21 cell line. Briefly, viral RNA was in vitro transcribed using mMessage mMachine SP6 kit (Thermo Fisher Scientific) and 10 µg RNA was electroporated in BHK-21 cells using a Gene Pulser XcellTM Electroporation System (Biorad). After cytopathic effects visualizations, culture medium containing SINV infectious particles were collected and centrifuged. Viral titrations were determined in PFU by a plaque-assay on Vero E6 cell line.

Fly infection, dissection and viral titration

Flies were reared on standard medium at 25°C. Flies were individually intrathoracically injected (Nanoject II apparatus; Drummond Scientific) with 32.2 nL containing 2,300 SINV infectious particles (PFU) (w^{1118}, *Dcr-2*$^{[L811fsX]}$ and *CG4572*^{c05963} mutants) in the BHK-21 cell culture medium (DMEM). In parallel, control flies were mock infected under the same conditions. The Makindu strain was intrathoracically injected with 207 nL containing 14,900 SINV PFU. Fly mortality at day one was attributed to damage due to the injection procedure, and these flies were excluded from further analyses. Flies were maintained at 25°C and transferred to fresh medium every two to three days. For RNA-seq and small RNA-seq, 30 flies

and 50 flies, respectively, were injected and the pairs of ovaries were dissected at 6 days post-infection.

For viral titration, pairs of ovaries and carcasses (whole body without ovaries) were crushed in a fresh grinding solution containing DMEM, 4% SVF (Gibco), 2.5 µg/mL B amphotericin (PAN Biotecj GmbH), 100 U/mL nystatin (Gibco), 50g/mL gentamicin (Gibco), 50 g/mL penicillin/streptomycin (Gibco). Viral titers in flies were determined by end-point dilution analysis on individual flies (n =10) on Vero E6 cell line using the Reed and Muench's method and viral titers were expressed as log10 50% tissue culture infective doses (TCID50/mL)[22]. SINV injection experiments were performed independently, in triplicate, and using three different stocks of SINV.

Immunostaining experiments

Immunostaining experiments were performed at 6 dpi from dissected ovaries and slices of infected heads and whole flies. Ovary pairs were fixed for 4 hours in 4% Paraformaldehyde solution under stirring at room temperature and permeabilized with PBS with Triton X-100 0.5% for 1 hour. Whole flies and heads were fixed for 4 hours in 2% Paraformaldehyde solution under stirring at room temperature. Then, they were incubated for 16 hours in 12% sucrose solution under stirring at 4°C, arranged in OCT blocs, and kept at -80°C. Slices of 14 µm thick sections were made using a cryostat (LEICA CM3050S) at -21°C. Immunostaining experiments including a second PFA fixation (0.5 % for 20 minutes) and permeabilization (PBS with Triton X-100 0.3% for 30 minutes) for slices were performed using a SINV anti-capsid[23] and a secondary A488 goat anti-mouse (A-11029, Invitrogen). All acquisitions were performed using LSM 510 META confocal microscopy (Carl Zeiss) and analyzed with ZEN software.

RNA-seq

Total RNA from 30 pairs of ovaries or 30 carcasses were extracted at 6 days post-infection using Trizol (Thermo Fisher Scientific) coupled with the RNeasy Mini Kit (Qiagen), according to the manufacturer's protocol. Purified RNA (10 µg) was treated using DNase TurboDNAse (Ambion DNAfree kit), followed by the validation of RNA quality using a Bioanalyzer (Agilent). Librairies were prepared after polyadenylated RNA selection. Purified RNAs were used for library preparation using the TruSeq RNA Sample Prep Kit v2. Enriched libraries were paired-end sequenced on a Hiseq 4000 apparatus (Illumina). Sequencing adaptors were removed using UrQt24 and trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned on reference genes retrieved from FlyBase (ftp://ftp.flybase.net/releases/FB2019_01/dsim_r2.02/fasta/dsim-all-gene-r2.02.fasta.gz and ftp.flybase.net/releases/FB2019_01/dmel_r6.26/fasta/dmel-all-gene-r6.26.fasta.gz) using

TopHat[25]. Alignments were then sorted using SAMtools26 and reads were counted using eXpress[27]. In parallel, we counted reads corresponding to TEs using the TEcount module of TEtools[28], and the list of sequences available at ftp://pbil.univ-lyon1.fr/pub/datasets/Roy2019/. Read counts obtained for genes on one hand and TEs on the other hand were then concatenated to make the complete read count table. It was analyzed at once using the DESeq2 R package (version 1.18)[29], which model internally corrects for library size, and provides normalization.

Data were deposited under the accession number PRJNA540249.

Small RNA-seq

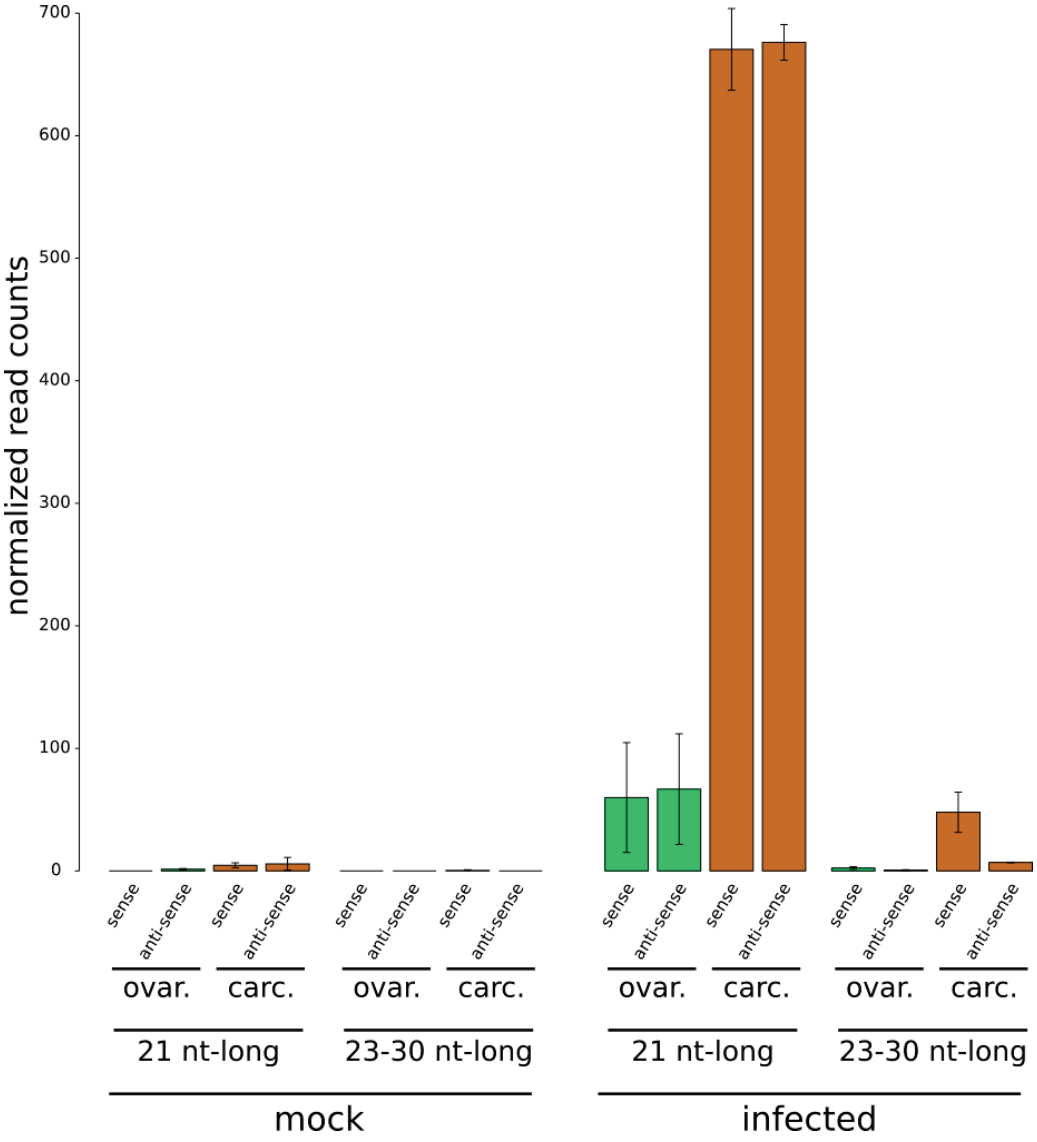
Small RNA were isolated from 50 pairs of ovaries or 50 carcasses using a HiTrap Q HP anion exchange chromatography column (GE Healthcare Life Sciences), as previously described[9]. Pairs of ovaries or carcasses were crushed on HQ 0.1 solution (Hepes 20 mM, KOH 20 mM, EDTA 0.2 mM, MgCl2 1,5 mM, Glycerol autoclaved 10 %, Ack 0,1M, Pefabloc® SC 0.1 mg/ml DTT 1 mM) and the supernantant was added on the HiTrap column, followed by a high Ack concentration solution (HQ 4 M). Small RNAs were purified using phenol/chlorophorm isoamyl alcohol extraction (Sigma-Aldrich) and quantified using a Qubit microRNA Assay kit (Thermo Fisher Scientific). Size selection of small RNA (18-50 bp) was performed on gel at the GenomEast platform. Purified small RNAs were used for library preparation (SQ00/SIL-04-SR), and sequenced on a Hiseq 4000 apparatus (Illumina). We first removed adapter sequences using cutadatp[30] -a TGGAATTCTCGGGTGCCAAGGAACTCCAGTCACTTA. Read count numbers were normalized to miRNA read count numbers (miRNA sequences retrieved from FlyBase: dsim-all-miRNA-r2.02.fasta.gz for Makindu, and dmel-all-miRNA-r6.25.fasta.gz for *D. melanogaster* samples). Using PRINSEQ lite version 0.20.431, we filtered reads of size 23 to 30 nt and considered these as piRNAs, and reads of size 21 nt and considered these as siRNAs. Using a modified version of the TEcount module of TEtools[5], we mapped the reads against a list of TE sequences (available at ftp://pbil.univ-lyon1.fr/pub/datasets/Roy2019/) either in the sense or antisense directions. We then computed the sums of sense and antisense counts.

Data were deposited under the accession number PRJNA540249.

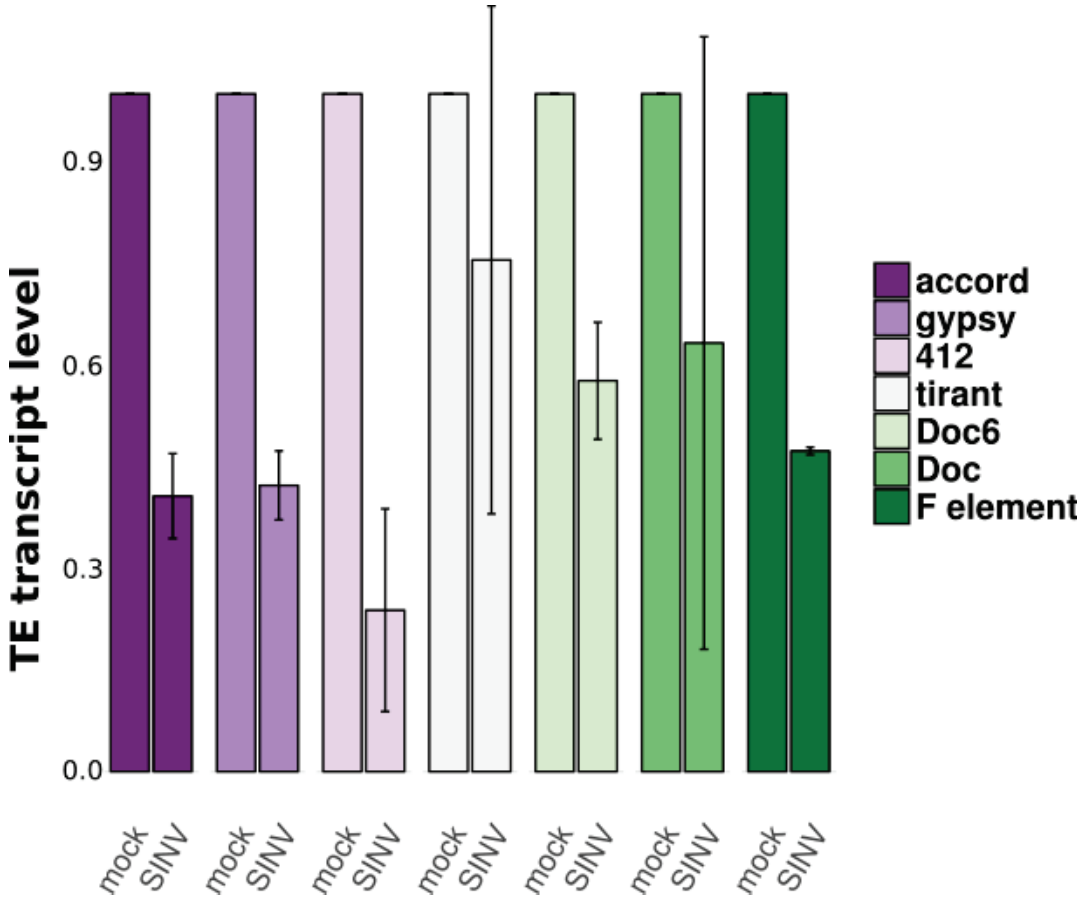
Statistical analyses

Statistical analyses were performed using R. We tested whether the modulations of TE sequences upon infection were different from zero using Wilcoxon paired test comparing TE normalized counts in the infected condition versus mock condition, for the considered 237 families. Differential expressions of genes were tested using the DESeq2 R package, and adjusted p-values were considered.

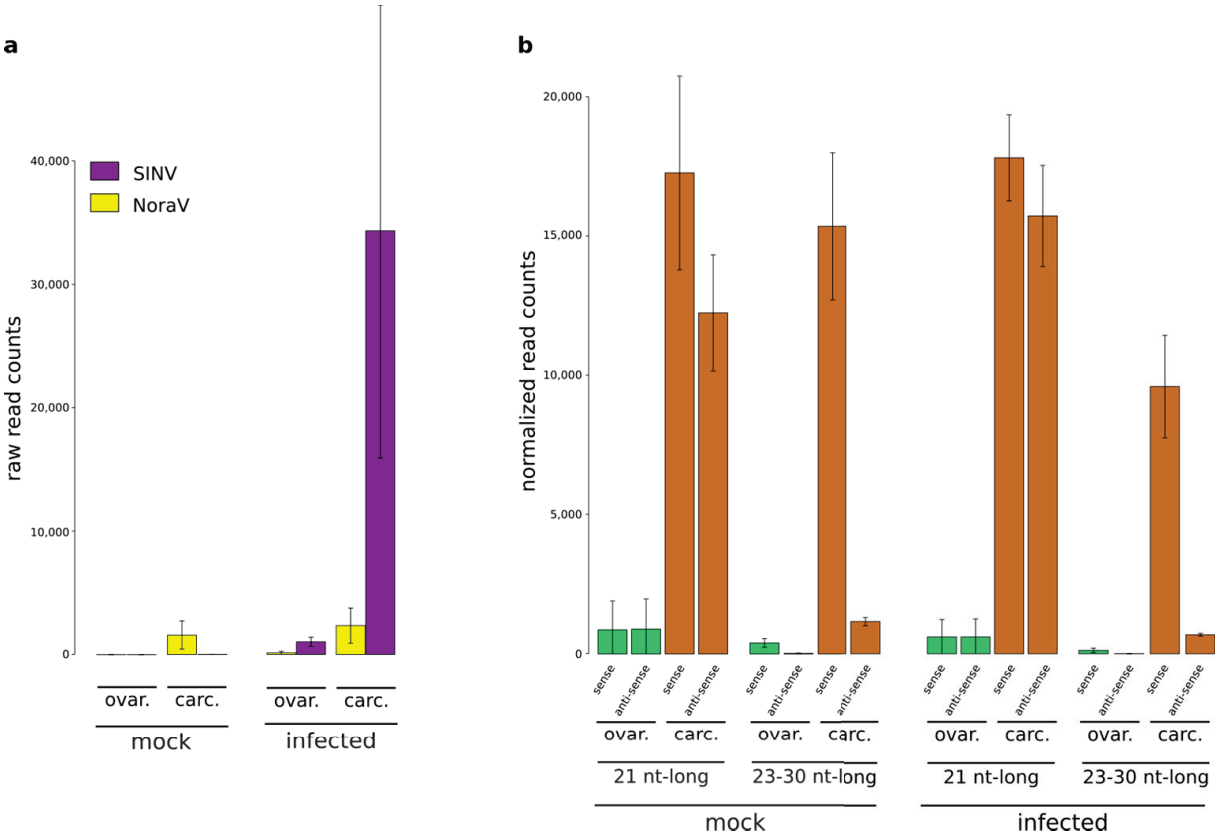
Supplementary data



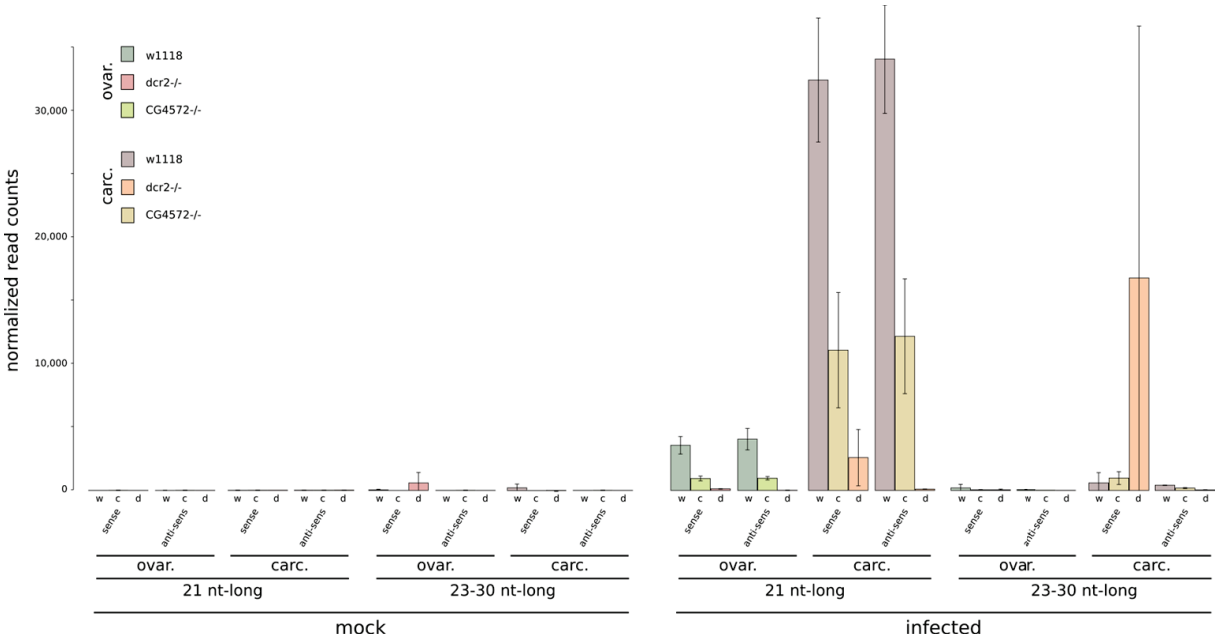
Extended data Figure 1 | SINV-directed small RNAs in Makindu samples. As expected, SINV-directed 21 nt-long small RNAs are detected in carcasses, and in ovaries to a lesser extent, upon SINV infection.



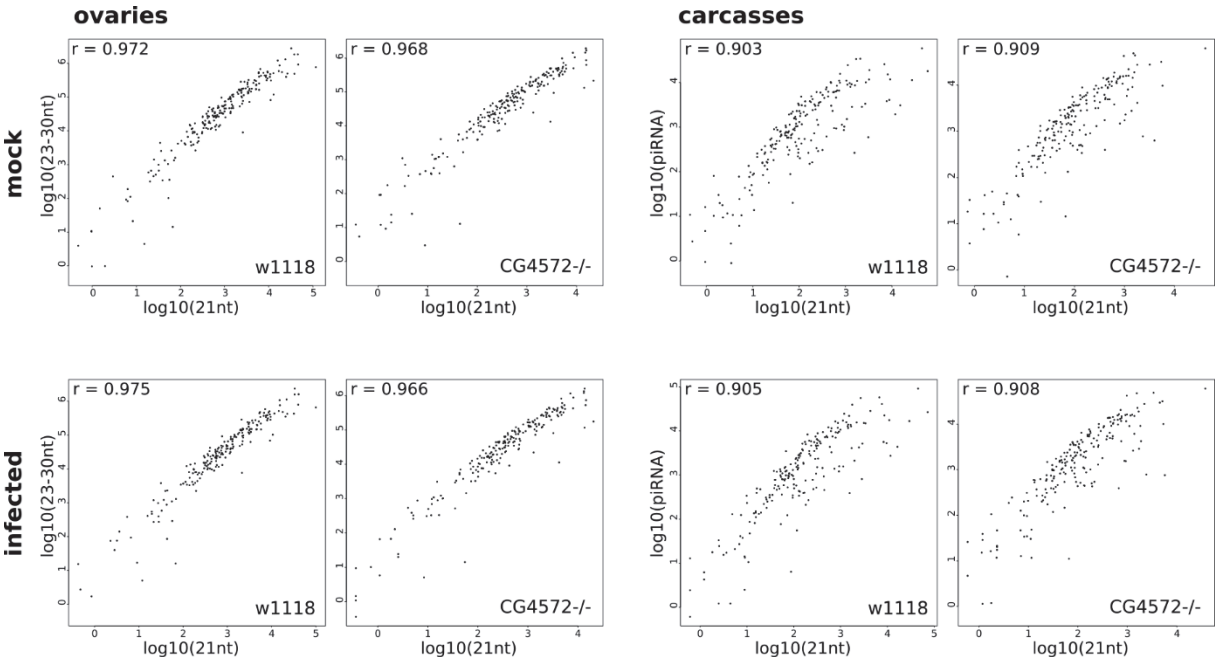
Extended data Figure 2 | TE transcription levels in carcasses upon SINV infection in Makindu. TE transcription levels, as assessed by RT-qPCR, are expressed as fold changes relative to *rp49* transcript levels. Reductions in TE transcript amounts are observed upon SINV infection for all TE families.



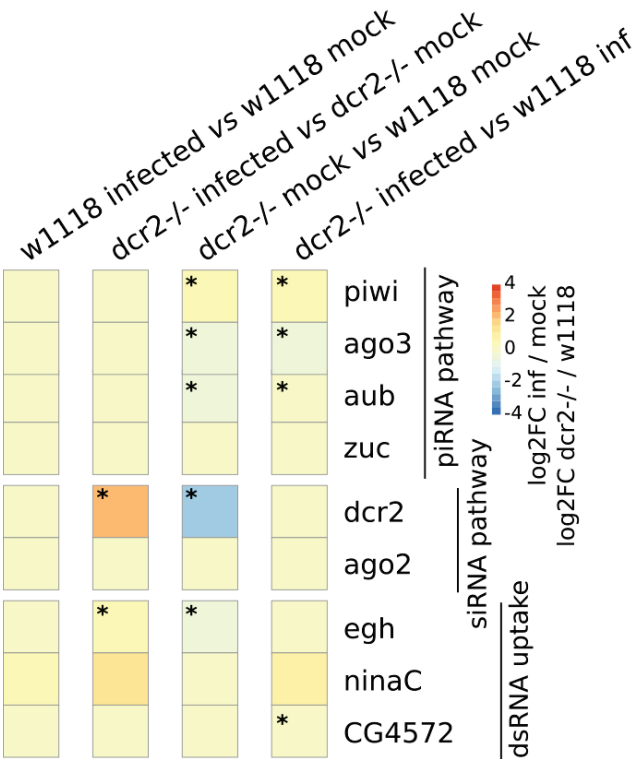
Extended data Figure 3 | Nora virus chronic infection in Makindu flies. a. RNA counts in Makindu samples. As expected, SINV reads are only detected in the infected condition. On the contrary, Nora virus reads are found in carcass samples of both conditions. **b.** Nora virus-directed small RNAs. Nora virus-directed small RNAs are detected in all conditions, essentially in carcasses. “Mock” and “infected” correspond to SINV infection conditions.



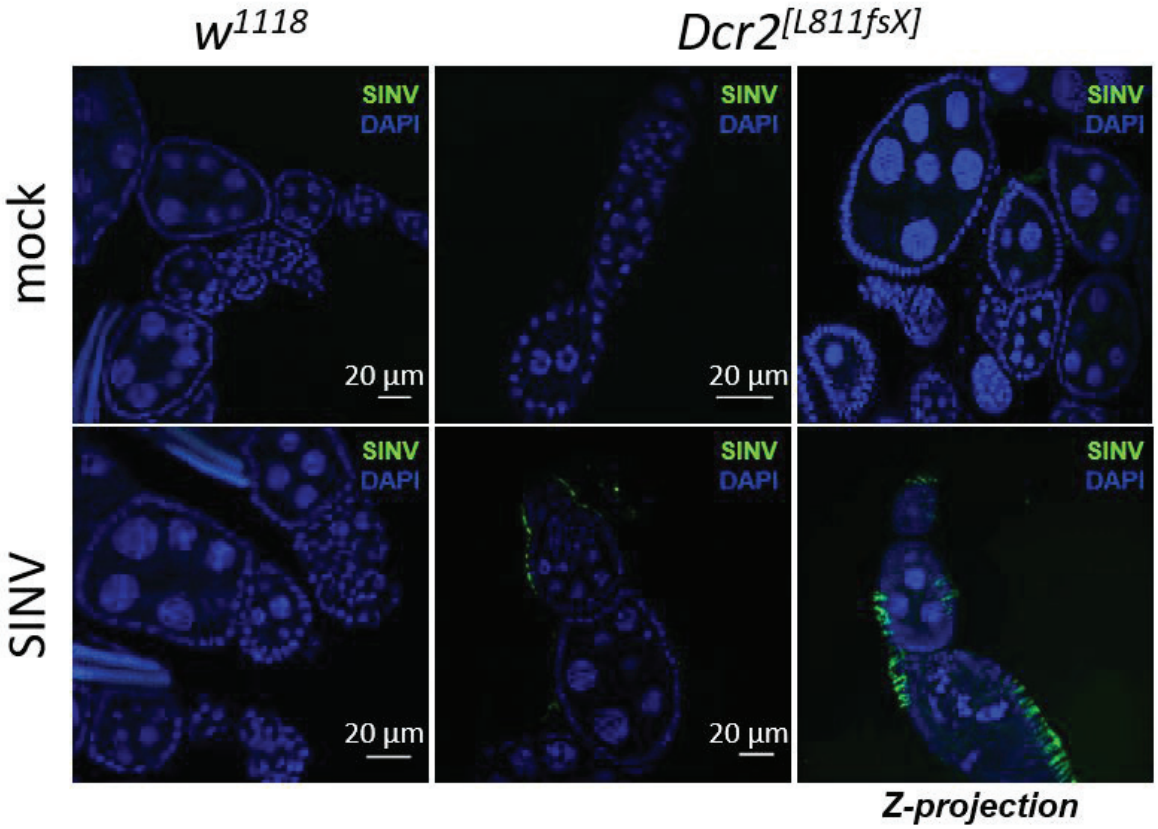
Extended data Figure 4 | SINV-directed small RNA production in *D. melanogaster* samples. w: w^{1118}; c: $CG4572^{-/-}$; d: $dcr2^{-/-}$. As expected, no signal is detected in mock conditions. In infected conditions, as expected, SINV-directed 21 nt RNAs are produced in ovaries, and essentially in carcasses in w^{1118} and $CG4572^{-/-}$ genotypes. In $dcr2^{-/-}$, an unexpected strong production of SINV-directed 23-30 nt RNAs is detected in carcasses.



Extended data Figure 5 | TE-directed 21 nt and 23-30 nt-long small RNAs are positively correlated in all conditions. Pearson correlation coefficients are indicated in the upper left corners.



Extended data Figure 6 | Differential expression of genes of the piRNA, siRNA and dsRNA uptake pathways in ovaries. Stars indicate significant differential expression as assessed by DESeq2 adjusted p-values inferior to 0.05.



Extended data Figure 7 | SINV was not detected inside ovarian structures. SINV tropism in ovaries as revealed by an anti-capsid antibody. SINV could not be detected in w^{1118} ovaries using immunostaining but was only detected at the periphery of $dcr2^{-/-}$ mutant ovaries. Z- projection acquisitions were performed for $dcr2^{-/-}$ mutants.

Methods for extended data

RT-qPCR

Total RNA of 30 pairs of ovaries or carcasses (whole body without ovaries) were extracted at 6 dpi by Trizol (Thermo Fisher Scientific) coupled with the RNeasy Mini Kit method (Qiagen), according to the manufacturer's protocol, in order to quantify a set of TE families (*accord*, *gypsy*, *412*, *tirant*, *Doc6*, *Doc*, *F element*). Purified RNA was treated using Turbo DNAse (Ambion DNAfree kit). Reverse transcription was performed from 200 ng (carcasses) or 500 ng (ovaries) of total RNA using SuperScript IV VILO Master Mix (Thermo Fisher Scientific) according to the manufacturer's instructions. qPCR was performed using the LightCycler 480 SYBR Green I Master, according to the manufacturer's protocol. r*p49* and *adh* were used for normalization purposes. Primers are recorded in the table below.

accord	Fw 5’ TTCACCCGTCGAAAGACTTC 3’ Rv 5’ GCCGTGAAAGAGTTCGAAAG 3’
gypsy	Fw 5’ GGCTCATTGCCGTTAAACAT 3’ Rv 5’ TCTTCCTTCTTTCGCTGAGG 3’
412	Fw 5’ TTGATGGGCAAAAGATCCAT 3’ Rv 5’ TTGCTGGAATTGTCGTTTCA 3’
tirant	Fw 5’ ACA CGT TCC CTG AAC AGA CG 3’ Rv 5’ GAA CGT TAC CAA TCC GAG CA 3’
Doc6	Fw 5’ ACATCATGTGCCCATTGTGC 3’ Rv 5’ AAATTCGGCATGGGCATGTC 3’
Doc	Fw 5’ GCACCTGGTTTCGAAGTGAT 3’ Rv 5’ GTAGCCAGGCGATTTTCTTG 3’
F element	Fw 5’ ACGTCGCTTACGCAGAGAAT 3’ Rv 5’ GTTTTGTGCCTGTTGGTGTG 3’
rp49	Fw 5’ CGGATCGATATGCTAAGCTGT 3’ Rv 5’ GCG CTT GTT CGA TCC GTA 3’
adh	Fw 5’ CCGTGGTCAACTTCACCAGCTC 3’ Rv 5’ TCCAACCAGGAGTTGAACTTGTGC 3’

Article en début de préparation.

Transposable element reactivation during viral infections in *Drosophila*

Marlène Roy[1,2], Barbara Viginier[2], Frédérick Arnaud[2], Maxime Ratinier[2] and Marie Fablet[1]

[1]Laboratoire de Biométrie et Biologie Evolutive, Université de Lyon; Université Lyon 1; CNRS; UMR 5558, Villeurbanne 69622, France,

[2] EPHE, PSL Research University, INRA, Univ Lyon, Université Claude Bernard Lyon1, UMR754, IVPC, F-69007, Lyon, France

Article en préparation pour une soumission au *Journal of Virology*.

Characterization of the virulence of the two main Toscana virus lineages

Marlène Roy[2] and Sandra Lacote[1,3], Sophie Desloire[2], Jérémy Welsch[4], Johann Vulin[1], Noémie Aurine[5], Justine Girard[2], Etienne Ossona De Mendez[2], Bertrand Pain[5], Cyrille Mathieu[4], Philippe Marianneau[3], Frédérick Arnaud[2], Maxime Ratinier[2]

[1]ANSES-Laboratoire de Lyon, Unité Virologie, Lyon, France.

[2]EPHE, PSL Research University, INRA, Univ Lyon, Université Claude Bernard Lyon1, UMR754, IVPC, F-69007, Lyon, France.

[3]ANSES-Laboratoire de Lyon, Plateforme d'Expérimentation Animale, Lyon, France.

[4]CIRI, International Center for Infectiology Research, CIRI, International Center for Infectiology Research, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Univ Lyon, F-69007, LYON, France.

[5]Univ Lyon, Université Lyon 1, INSERM, INRA, Stem Cell and Brain Research Institute, U1208, USC1361, Bron, France.

Abstract

Using a variety of *in vitro* (such as induced Pluripotent Stem Cells, iPSCs, derived neural cells), *ex vivo* and *in vivo* models of infection, we determined that two Toscana virus (TOSV) strains, representatives of the two main known lineages (A and B) of this virus, show marked differences in term of fitness and tropism. Thus, levels of viremia and humoral response in the blood of TOSV-A infected IFNAR$^{-/-}$ mice are delayed in time than those of TOSV-B. Moreover, despite higher level of replication, TOSV-B was not detected in the brain of the infected animals (0/6) while TOSV-A has been found in this organ of 4 mice out of six. Altogether, the data obtained in this study opens new avenues to elucidate viral determinants involved in viral pathogenesis and transmission.

Self and Nonself from a Genomic Perspective: Transposable Elements

Marie Fablet, Judit Salces-Ortiz, Bianca Fraga Menezes, Marlène Roy and Cristina Vieira

Abstract Transposable elements (TEs) are sequences that can move and multiply along the chromosomes. Considered for a long time as genomic parasites, they are now acknowledged as key players of genome function and evolution. Accordingly, the presence of TEs in a genome may affect the chromatin structure of the regions in which they are inserted. TEs allow us to revisit self and nonself distinction at the genomic level, through the complex relationships they display with the genome and the epigenome and their interaction with the environment.

1 Introduction

Genomes are not mere strings of genes. They are also scattered with—apparently—dispensable sequences, which we know as transposable elements (TEs) due to their ability to move from one locus to another along the chromosomes. They were discovered in the 1950s thanks to the pioneering work of Barbara McClintock (1950), who was awarded the Nobel Prize in 1983. The historic study of TE research is a textbook case of how ideas evolve within the scientific community. First considered as "junk DNA," devoid of scientific interest, they are now rec-

M. Fablet (✉) · C. Vieira (✉)
43 boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex, France
e-mail: marie.fablet@univ-lyon1.fr

C. Vieira
e-mail: cristina.vieira@univ-lyon1.fr

M. Fablet · J. Salces-Ortiz · B.F. Menezes · M. Roy · C. Vieira
Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558, Université Lyon 1, Université de Lyon, Villeurbanne, France
e-mail: judit.salces-ortiz@univ-lyon1.fr

B.F. Menezes
e-mail: bianca.menezes@univ-lyon1.fr

M. Roy
e-mail: marlene.roy@etu.univ-lyon1.fr

© Springer International Publishing AG 2017

P. Pontarotti (ed.), *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts*,
DOI 10.1007/978-3-319-61569-1_6

ognized as key functional and evolutionary components of genomes (Biémont 2010).

TEs—also known as mobile elements—are generally shorter than 10 kb and are able to mobilize, i.e., to *transpose*, and multiply within genomes thanks to the enzymatic machinery they encode. TE copies resulting from transposition events of a given TE share sequence similarity with each other and constitute TE families. Based on their transposition mechanisms, TE families are broadly organized into two classes (a more comprehensive classification is proposed in Wicker et al. (2007)) (Fig. 1). Class II TEs are also called *transposons* and are mobilized through a *cut and paste* mechanism. They encode one enzyme called *transposase* and are bordered by Terminal Inverted Repeats (TIRs). Class I TEs are termed *retrotransposons* and get mobilized through a *copy and paste* scenario via an RNA intermediate. Two subclasses of retrotransposons are distinguished. LTR retrotransposons are bordered by direct Long Terminal Repeats (LTRs), which include regulatory sequences, and display three Open Reading Frames (ORFs): *gag*—encodes proteins of the capsid, *pol*—encodes the retrotransposition enzymes, i.e., reverse transcriptase, RNase-H, integrase, protease, and sometimes *env*—encodes proteins involved in the formation of an envelope. Endogenous retroviruses (ERVs) are sequences of viral origin integrated into the genome and transmitted from parent to progeny, like genes. They are included into the LTR retrotransposon subclass. The other retrotransposon subclass is made of non-LTR retrotransposons. Among them, LINEs (Long Interspersed Nuclear Elements) display two ORFs, sharing similarity with *gag* and *pol*, and are terminated by a polyA tail, and SINEs (Short Interspersed Nuclear Elements) are non-autonomous elements, meaning they do not encode their own machinery and are mobilized by LINEs. SINEs are non-coding and result from accidental transposition of RNA polymerase III transcripts (such as transfer RNAs) (Wicker et al. 2007).

What makes the transition from a genomic parasite to a domesticated sequence? This is a sensitive question around TEs. We intend to answer it throughout this

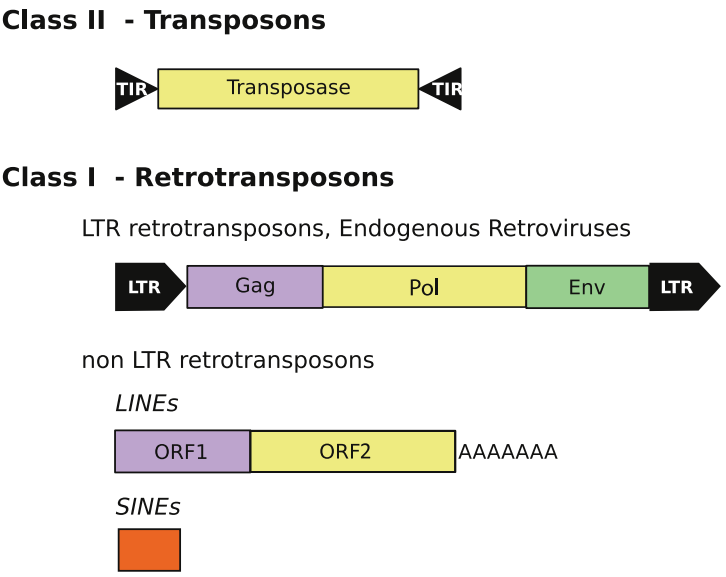


Fig. 1 Structures of the archetypes of TE classes and subclasses. TIR: Terminal Inverted Repeat. LTR: Long Terminal Repeat. ORF: Open Reading Frame. SINEs are short non-coding sequences that derive from RNA polymerase III transcripts

chapter by examining ways the host genome controls TEs, the effects of TE insertions, and the way they can be fully integrated into the host genome functioning and evolution. The long-lasting interaction between genomes and TEs makes the latter considered sometimes as genomic nonself and sometimes as genomic self.

2 TEs as Nonself: Genomes Fight Against TEs

TEs are a threat to genome integrity because of both their mobility and their repetitive nature. Indeed, TEs may be harmful when a new insertion disrupts the coding sequence of a gene or regulatory sequences. They are also harmful when ectopic recombination occurs between two distant copies of the same TE family (Petrov et al. 2003), which leads to deleterious chromosomal rearrangements. Population genetic models were proposed and considered the deleterious effects of TE insertions and the deleterious consequences of genomic rearrangements (Charlesworth and Langley 1989; Langley et al. 1988). The elimination of TEs is thus expected in the long-term process of evolution. However, what we know from genomic analyses is that is it generally not the case. The resolution of this conundrum was allowed by the discovery of defense mechanisms against TE expression that can act rapidly, at a shorter evolutionary timescale. We will dedicate this section to these mechanisms.

2.1 Genomes Control TEs Through the PiRNA Pathway

We have known for almost 15 years that TEs are controlled by RNA interference mechanisms via the Piwi-interacting RNA (piRNA) class of small RNAs, which were originally called rasiRNAs (Aravin et al. 2007; Klenov et al. 2007; Saito et al. 2006; Vagin et al. 2006). Most of our knowledge comes from fly studies which will be presented below, but similar mechanisms also occur in other taxa.

piRNAs are 23–30-nt-long single-stranded RNA molecules (Fig. 2). They are loaded onto proteins of the PIWI subfamily (Argonaute family): Aubergine (Aub), Argonaute 3 (Ago3), and Piwi (Meister 2013; Siomi et al. 2011), which display RNase-H activity (Jinek and Doudna 2009). In *Drosophila melanogaster* ovaries, antisense piRNAs are loaded onto Aub and target sense TE transcripts. The latter are then sliced into sense piRNAs, which are loaded onto Ago3 and target antisense TE transcripts. This leads to the so-called *ping-pong loop* (Brennecke et al. 2007), which occurs in the *nuage*, a dense cytoplasmic region located near the nucleus. Such piRNAs are termed secondary piRNAs. Secondary piRNAs loaded onto Ago3 show a 10 nt sequence overlap at their 5′ end with secondary piRNAs loaded onto Aub, which is called the *ping-pong signature* (Brennecke et al. 2007).

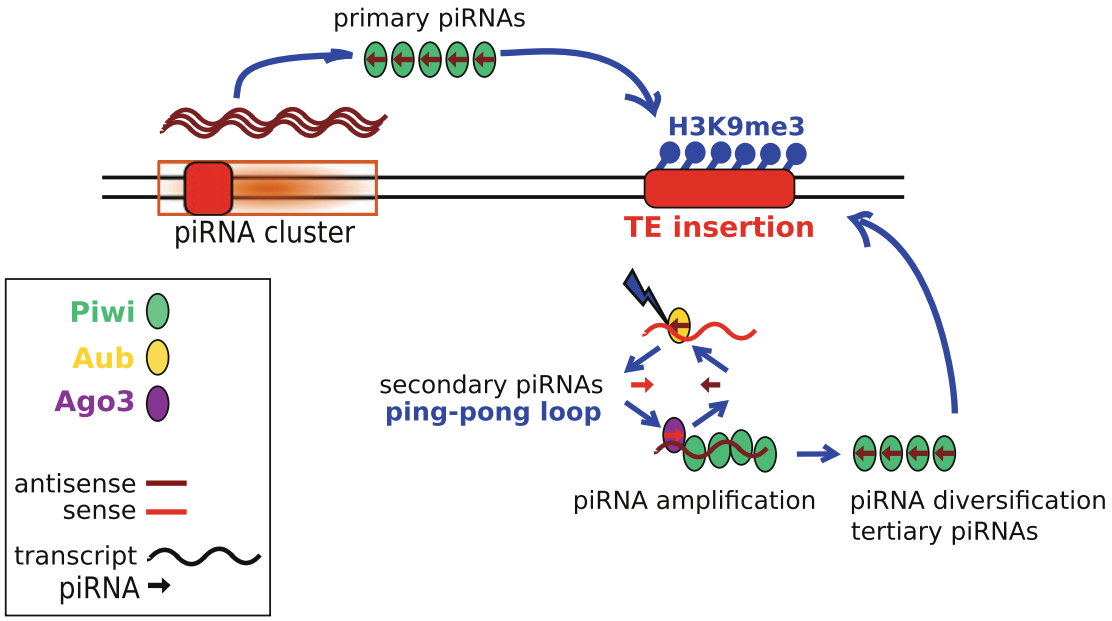


Fig. 2 piRNA pathway. Antisense piRNAs are loaded onto Aub and target sense TE transcripts. The latter are then sliced into sense piRNAs, which are loaded onto Ago3 and target antisense TE transcripts. This leads to the *ping-pong loop*, producing secondary piRNAs. Piwi-bound piRNAs go to the nucleus and promote the transcription of piRNA clusters. This generates primary piRNAs, which target TE insertion sites triggering their heterochromatinization (H3K9me3 histone marks). In addition, tertiary piRNAs result from Piwi-induced phasing and allow the diversification of piRNA sequences

Antisense piRNAs may also be loaded onto Piwi and go back to the nucleus. They then promote the transcription of particular loci that are called piRNA clusters. These loci are heterochromatic and, in *Drosophila*, are predominantly filled with TE sequences. The best known *Drosophila* cluster is *flamenco* (X chromosome, ~180 kb), which controls the *gypsy* endogenous retrovirus (Pélisson et al. 1994; Prud'homme et al. 1995) among others. The transcription of piRNA clusters leads to the production of long precursors that are cleaved into primary piRNAs. These are loaded onto Piwi and target TE insertion sites triggering their heterochromatinization via the recruitment of histone-modifying enzymes (Akkouche et al. 2013; Le Thomas et al. 2013; Rozhkov et al. 2013a; Sienski et al. 2012). Such a mechanism ensures TE control both, at the transcriptional and posttranscriptional levels.

In addition, it was recently discovered that Piwi-bound piRNAs displayed sequence *phasing*, meaning that the 3′ ends of such piRNAs are immediately followed by the 5′ end of the next Piwi-bound piRNAs (Han et al. 2015; Mohn et al. 2015). Such phased piRNAs are considered tertiary piRNAs (Siomi and Siomi 2015) and go back to the nucleus and achieve transcriptional TE silencing. These phased Piwi-bound piRNAs are mostly generated from Ago3-bound secondary piRNAs (Wang et al. 2015) (Fig. 2). While the ping-pong loop provides secondary piRNAs in high amounts, those phased tertiary piRNAs allow increases in the diversity of target sequences (Han et al. 2015).

In *Drosophila* ovaries, secondary piRNAs are produced only in germline cells while primary piRNAs are produced both in germline and somatic cells of the ovary (Malone et al. 2009). Which piRNA cluster is transcribed also depends on the cell type. Feeding of the ping-pong loop can be provided by the maternal transmission of Aub-bound secondary piRNAs. In addition, while it was considered for a long time that the *Drosophila* piRNA pathway took place exclusively in the ovaries, it is now known that an efficient secondary piRNA pathway also takes place in the fat body (Jones et al. 2016), widening the potential roles of piRNA regulation.

2.2 Genomes and TEs: A Red Queen Relationship

It is common today to consider the piRNA pathway as an immune pathway acting at the genomic level (Aravin et al. 2007; Fablet 2014; Malone and Hannon 2009; Senti and Brennecke 2010; Siomi et al. 2011). Such a parallel is even more relevant when performing an evolutionary analysis of the genes involved in this system. Classical host-pathogen interactions are often described using the *Red Queen* metaphor, after Lewis Carroll's novel (van Valen 1973). It illustrates the fact that both partners of this antagonistic relationship are constantly evolving. For instance, concerning antiviral immunity in *Drosophila*, which is achieved through the siRNA pathway, Obbard et al. (2006, 2009) showed that genes involved in the siRNA pathway displayed strong signatures of recurrent positive selection, which is distinctive of rapid, constant evolution. It is noteworthy that genes involved in the piRNA pathway (GIPPs) also showed strong, significant signatures of recurrent positive selection in the same study (Obbard et al. 2009). Other works, using different samples and methods, confirmed this general trend of rapid sequence evolution for GIPPs (Fablet et al. 2014; Kolaczkowski et al. 2011; Obbard et al. 2009b). In addition, GIPPs also display rapid expression rate evolution (Fablet et al. 2014). All these elements suggest that the Red Queen metaphor also applies to the genome (at the level of GIPPs) versus TE interaction, meaning that the relationship that host genomes develop toward TEs could be considered host-pathogen-like, at least to some extent. However, the molecular mechanisms of the interaction and the rapid evolution are still a topic of further investigation.

Some authors even go further and apply the hygiene hypothesis to the case of TEs and GIPPs. The hygiene hypothesis states that, in *Homo sapiens* populations, the increase of pathogen abundance during the neolithic period constituted a selective pressure for an efficient immune system, which is now assumed to be responsible for an increase in the incidence of autoimmune diseases in the aseptic-ized post-industrial era (Rook 2012; Strachan 1989). Blumenstiel et al. (2016) propose that genomic autoimmunity could result from off-targets of a too efficient piRNA pathway and would reinforce the evolutionary tension between GIPPs and

the genome. Nevertheless, there is still a need for empirical data to sustain this hypothesis.

2.3 Acquired Genomic Immunity Is Provided by PiRNA Clusters

The piRNA pathway is considered as an acquired immune pathway since it provides protection only against those TE families that have one copy inserted into a piRNA cluster (Malone et al. 2009; Zanni et al. 2013). Indeed, a clear causal relationship between the presence of the TE family within a cluster and the repression of its mobility was demonstrated in the case of the *gypsy* and *ZAM* retrotransposons, which are controlled by the *flamenco* cluster (Zanni et al. 2013). Zanni et al. (2013) also suggested that although their chromosomal locations are conserved across closely related species (Malone et al. 2009), piRNA clusters such as *flamenco* are highly dynamic regions. Indeed, insertions, deletions, and rearrangements frequently occur, even at the intraspecific scale. Again, this rapid evolution fits with the above-mentioned Red Queen hypothesis. As a consequence, the dynamics of transposition can also be rapidly evolving, since a specific TE family can reacquire activity if ever the piRNA cluster looses the corresponding insertion.

TEs are a threat to the genome at the time they invade it, for instance, after a horizontal transfer event or after reactivation of an old family. Theoretical work on *Drosophila* indicates that they remain a threat up until a copy inserts within a piRNA cluster, which triggers control through the piRNA pathway (Lu and Clark 2010). Indeed, once the piRNA pathway controls a TE family, the different copies of these families are neither able to transpose—because they are transcriptionally and posttranscriptionally silenced by piRNAs—nor involved in ectopic recombination, since recombination is less likely to occur in heterochromatinized regions. Khurana et al. (2011) experimentally demonstrated that *P* element introduction into a naive genome first leads to a high transpositional, potentially deleterious activity. However, as the host ages, control of the *P* element can be established. The authors propose that this control is possible due to the insertion of a *P* copy within a piRNA cluster (Khurana et al. 2011) at some point along the life of the host individual. Similarly, Rozhkov et al. (2013b) introduced a *Penelope* TE into a naive *D. melanogaster* genome and found that after some generations, *Penelope* control was established due to insertions into piRNA clusters.

We propose that a TE family is recognized as genetically nonself as long as it is not inserted into a piRNA cluster. Once silenced, the TE family can enter a new relationship with the host genome, eventually leading to genomic assimilation into self.

3 TEs as Self: Genomes Recruit TEs

3.1 TEs as Raw Material for Genome Evolution

TE mobilization is known to be responsible for deleterious mutations or to be involved in the first steps of cancer and other diseases in humans (Belancio et al. 2009; Hancks and Kazazian 2016; Hancks and Kazazian 2012; Kim et al. 2007). They can also be at the origin of phenotypic variation, without deleterious effects, as for instance, body size diversity in dogs (Sutter et al. 2007). The historic example is that of Barbara McClintock, which led her to the discovery of TEs: Variation in pigmentation patterns in maize kernels is due to the activity of the *Ac* and *Ds* transposons (McClintock 1950). Other examples of color variation due to transposable elements have been recorded in plants. For instance, color variation in maize pericarp can be due to the insertion of an *Ac* transposon into the *p1* gene (Zhang and Peterson 2005), and petal color variation in *Anthirrhinum* is associated with different insertions of the *Tam* transposon on both sides of the *nivea* gene, responsible for flower pigmentation (Uchiyama et al. 2013). In animals, TE insertions are associated with coat color variability in dogs and cats (Clark et al. 2006; David et al. 2014), or the classical example of the agouti color in mice (Morgan et al. 1999).

TE mobilization contributes to raw material for genome evolution. While TE insertions into genes lead to deleterious mutations, most of the time, the insertion of a TE in the vicinity of genes will increase the possibilities and modalities of gene expression. For example, the insertion of a *Doc* TE within the *CHKov* gene leads to alternative transcripts that provide *Drosophila* with pesticide resistance (Aminetzach et al. 2005). On a more global scale, Vieira et al. (1999) observed an increase of the global TE content in derived strains of *D. melanogaster* as strains were further away from the ancestral, African area. They proposed that such an increase in TE amounts could be adaptive in the frame of the colonization of new environments. As reviewed by Casacuberta and González (2013), TEs may play major roles in environmental adaptation. Similarly, but at a shorter timescale, Perrat et al. (2013) found that transposition occurred in memory-relevant neurons in the *Drosophila* brain and proposed that this genomic heterogeneity was a conserved feature of the brain potentially producing behavioral variability between individuals. Prior work on rodents and humans (Coufal et al. 2009; Muotri et al. 2005) also suggested that transposon-mediated genomic heterogeneity may be a conserved characteristic of certain neurons.

TEs are made up of coding and regulatory sequences. Long-term, safe relationships between the host genome and TEs may lead to the recruitment of TE sequences, so that such insertions are considered as *domesticated* (Miller et al. 1999; Volff 2006). For instance, at least 25% of human gene promoters derive from TEs (Jordan et al. 2003). The most famous cases of domestication are *rag1* and *rag2* genes—involved in the vertebrate V(D)J recombination giving birth to immune cell receptors (Agrawal et al. 1998), *syncytin* genes—remnants of an ERV

env gene, allowing placenta formation (Blaise et al. 2003; Mi et al. 2000), and *TART* and *HeT-A*—non-LTR retrotransposons insuring telomere maintenance in *Drosophila* (Biessmann et al. 1990; Levis et al. 1993). New examples continue to be published. For instance, the extensively taught industrial melanism mutation of peppered moth is now known to be due to a TE insertion (Van't Hof et al. 2016).

3.2 Genomes Recruit TE Regulatory Sequences

At the time they invade a genome, most TEs possess all necessary sequences for their expression and mobilization. For instance, they all display promoter sequences —even the non-autonomous SINEs—and transcription termination signals. Most TEs also contain transcription factor binding sites (Chuong et al. 2016; Zemojtel and Vingron 2012). As mentioned previously, TEs are the targets of silencing mechanisms from chromatin modifications to DNA methylation (Slotkin and Martienssen 2007) that minimize transposition potential. Such epigenetic modifications also modulate the expression of neighboring genes (Hollister and Gaut 2009; Rebollo et al. 2011) by the spreading of chromatin marks. This abundant source of regulatory sequences may be hijacked by the genome for its own purpose. For instance, TEs may be a source of promoter sequences for duplicated genes (e.g., Fablet et al. 2009). In addition, the fact of having the same repetitive sequences all along the chromosomes may transform TEs into conductors of regulatory networks (Chuong et al. 2017; Rebollo et al. 2012). For instance, Chuong et al. (2016) recently demonstrated that *MER41.AIM2*, an old human endogenous retrovirus, behaved as an enhancer and provided susceptibility to interferon in the inflammatory response to infection.

The number of reported cases of TEs being involved in gene regulation is exponentially increasing. What we mention in this chapter is merely the tip of the iceberg. Among the most remarkable examples, TEs may be involved in sex determination. In melon, a TE insertion in the promoter of a transcription factor spreads DNA methylation and results in the transition from male to female flowers (Martin et al. 2009). The determination of new sex chromosomes was also shown to be driven by TEs and the epigenetic marks that are associated with the neo-Y chromosome in *Drosophila miranda* (Zhou et al. 2013).

3.3 Genomes Recruit TE Coding Sequences

While it is rather common to find TEs recruited for their regulatory regions, it is less frequent, although more spectacular, to encounter TEs recruited for their coding sequences. The V(D)J recombination system in immune cells of vertebrates has been suggested for a long time to result from the domestication of a putative *RAG* transposon (Agrawal et al. 1998). It is only very recently that an active *ProtoRAG*

TE was actually found in the lancelet genome and is considered as a relative of the long-sought *RAG* TE (Huang et al. 2016).

The mouse *Fv1* gene is involved in virus resistance and derives from a TE *gag* gene (Yan et al. 2009). *Gag* genes of LTR retrotransposons and endogenous retroviruses encode capsid proteins, which, in the case of *Fv1*, interacts with exogenous virus capsid proteins, blocking progression of the viral cycle (Hilditch et al. 2011).

Hoen and Bureau (2015) performed a comprehensive search of novel genes derived from TEs in the *Arabidopsis thaliana* model plant genome and found dozens of such genes. This particular new relationship between the genome and TEs implies that each of these domesticated TEs had lost its mobilization ability and is now expressed—while non-domesticated TEs are generally silenced.

Kokošar and Kordiš (2013) thoroughly studied such phenomena in mammalian genomes and found that it had been particularly frequent in the ancestor of placental mammals. They propose a scenario for the transition from TEs to domesticated genes. They suggest that nucleotide changes first allow neofunctionalization, along with exonization of TE domains. The recruitment of *cis*-regulatory regions, such as promoters, is the next compulsory step in the acquisition of functionality. It can be achieved through different mechanisms such as the recruitment of bidirectional promoters, or promoter capture via the evolution of 5′-UTR exon/intron structures (Fablet et al. 2009; Kaessmann et al. 2009; Kokošar and Kordiš 2013).

4 TEs as Nonself Sensors

From genomic parasites to genomic full components, from genomic nonself to genomic self, TEs present complex relationships with the host genome. TEs have been shaping genomes for millions of years, and they may play a key role in the evolution of gene regulatory networks. This may explain the wide range of rapid phenotypic evolution and origin of morphological novelties found within and among populations (Cowley and Oakey 2013). In that sense, TEs are eventually used by the genomes as nonself sensors.

4.1 TEs as Sensors of Divergent Genomes

In the 1960s and 1970s, studies on *D. melanogaster* populations revealed the phenomenon of *hybrid dysgenesis*, which refers to aberrant phenotypic traits observed in the F1 of crosses between particular strains or natural populations. This was attributed to differences in TE content between the parental lines (Kidwell 1977; Picard 1976). Owing to the very recent development of the epigenomic field, we now know that the observed disruption of genomic stability in hybrids is the consequence of a high rate of TE mobilization, which itself is the result of the destabilization of epigenetic regulatory networks (Chambeyron et al. 2008; Jensen et al. 2008; Todeschini et al. 2010).

Since these pioneering studies of hybrid dysgenesis in *Drosophila*, other cases of TE reactivation during divergent crosses were recorded in other species of *Drosophila*, mice, and wallabies (Labrador et al. 1999; Lopez-Maestre et al. 2017; Metcalfe et al. 2007; O'Neill et al. 1998; Vela et al. 2014). The challenge is now to understand whether TEs are a primary contribution to population isolation, as for the *P* element in *D. melanogaster*, which was acquired by horizontal transfer, or whether they are only participating in the hybrid incompatibility once speciation has occurred (Craddock 2016; Rebollo et al. 2010). Both processes are probably involved. For example, when crossing closely related *Drosophila* species such as *D. mojavensis* and *D. arizonae*, a well-suited model to study speciation, authors showed that only very few TEs are misexpressed in hybrids (Carnelossi et al. 2014; Lopez-Maestre et al. 2017). Increasing the evolutionary distance between species will lead to a genome-wide misregulation of TEs, which has been attributed to divergence in the GIPP proteins, that fail to control TEs (García Guerreiro (personal communication), Kelleher et al. 2012). Thus, it appears that TEs are sensitive to genomic divergence and therefore fundamental guardians of genomic self.

4.2 TEs as Sensors of the Abiotic Environment

As previously mentioned, TEs bear transcription binding sites, which can contribute to making them sensitive to the environment. For instance, they harbor heat shock responsive elements (Pietzenuk et al. 2016), ecdysone binding sites (Micard et al. 1988), etc. An interesting example is the *Bari* insertion upstream of *D. melanogaster Jheh2* and *Jheh3* genes that adds antioxidant response elements and provides oxidative stress resistance to individuals bearing it (Guio et al. 2014).

TEs are also the targets of epigenetic modifications, through histone modifications or DNA methylation, depending on the organisms (Lister et al. 2008; Rebollo et al. 2011; Sienski et al. 2012). We may notice that it is a way for the genome to distinguish genomic self from genomic nonself. Although we do not understand the exact molecular mechanisms, we know that epigenetic modifications are sensitive to the environment. This TE/epigenetics/environment tripartite relationship makes TEs important players of evolvability (Fablet and Vieira 2011), promoting phenotypic diversity and environmental adaptation (Song and Cao 2017) (Fig. 3). The agouti coat color of mice is an emblematic illustration. Variation in coat color in mice can be due to variability in the methylation level of a TE insertion upstream of the *A* gene responsible for coat color (Morgan et al. 1999). Waterland and Jirtle (2003) showed that a methyl-rich diet provided to the mother could impact the methylation level of this particular TE and had direct effect on the coat color of the progeny. Evolutionary implications are wide. For instance, we may imagine that, in the case of melon sexual determination mentioned above (Martin et al. 2009), populations placed in environments differing for their methyl richness would display different sex ratios, therefore impacting the effective population size and the efficiency of natural selection.

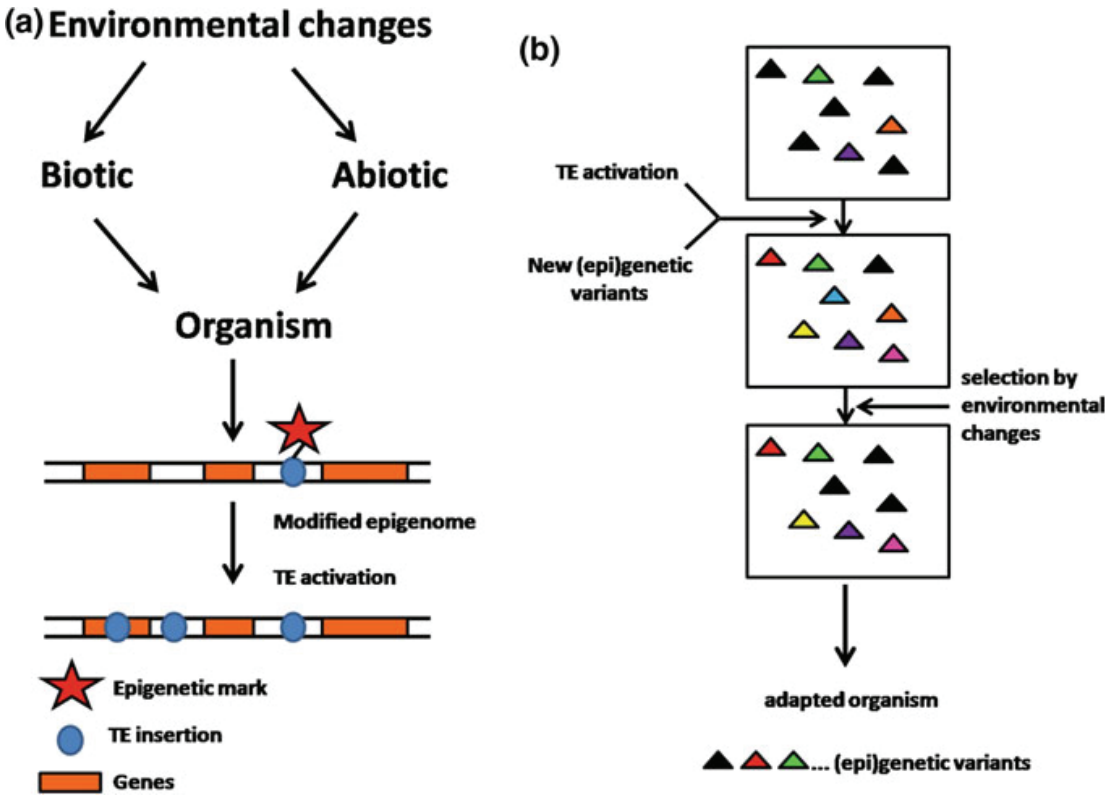


Fig. 3 TE, epigenome, and environment relationship. **a** Environmental changes could be perceived by the organisms through the epigenome and impair the control of transposition, leading to TE-mediated insertional mutagenesis and the induction of phenotypic variability. **b** (Epi)genetic variability increases due to induced TE activity in response to environmental changes on which selection could act providing favorable mutations to promote adaptation

Such sensitivity to environmental conditions is referred to as phenotypic plasticity, i.e., the ability of a genotype to express different phenotypes according to the environment. As explained above, TEs appear as major actors of plasticity, and therefore play a significant role in the adaptive potential of species (Fig. 3) (Casacuberta and González 2013; Fablet and Vieira 2011).

4.3 TEs as Sensors of the Biotic Environment

TEs may even behave as sensors of the biotic environment. For instance, the tobacco *Tnt1* TE is known to be reactivated in case of fungal infection (Melayah et al. 2001). Response to fungal infection in *A. thaliana* is controlled by the targeted demethylation of TE sequences located within promoters of particular genes resulting in downregulation of many stress response genes (Le et al. 2014). This sensitivity to pathogens may even be directly involved in immune pathways. This is the case when TE insertions display antiviral properties, such as the above-mentioned *Fv1* example. This is also true when TEs provide interferon-sensitive sequences, as in the case of *MER41.AIM2* (Chuong et al. 2016). The authors of this study suggest that such an

involvement of TEs from the endogenous retrovirus class in immunity is not unexpected. Indeed, they propose that it results from ancient viral adaptations allowing exploitation of immune pathways.

Such an idea is illustrated in the results of Karijolich et al. (2015). In mice, during murine gammaherpesvirus 68 infection, transcription of the *B2* SINE is enhanced and activates the antiviral NF-κB pathway. However, at the same time, this SINE expression is hijacked by the virus and stimulates viral replication (Karijolich et al. 2015).

5 Conclusion

The relationship between a TE family and the host genome is complex (Fig. 4). Firstly, because it evolves with time from the moment the TE family invades the genome to the moment the interaction is stable. Secondly, as TEs are repeated, the different copies of a given family may not have the same fate. One copy may be

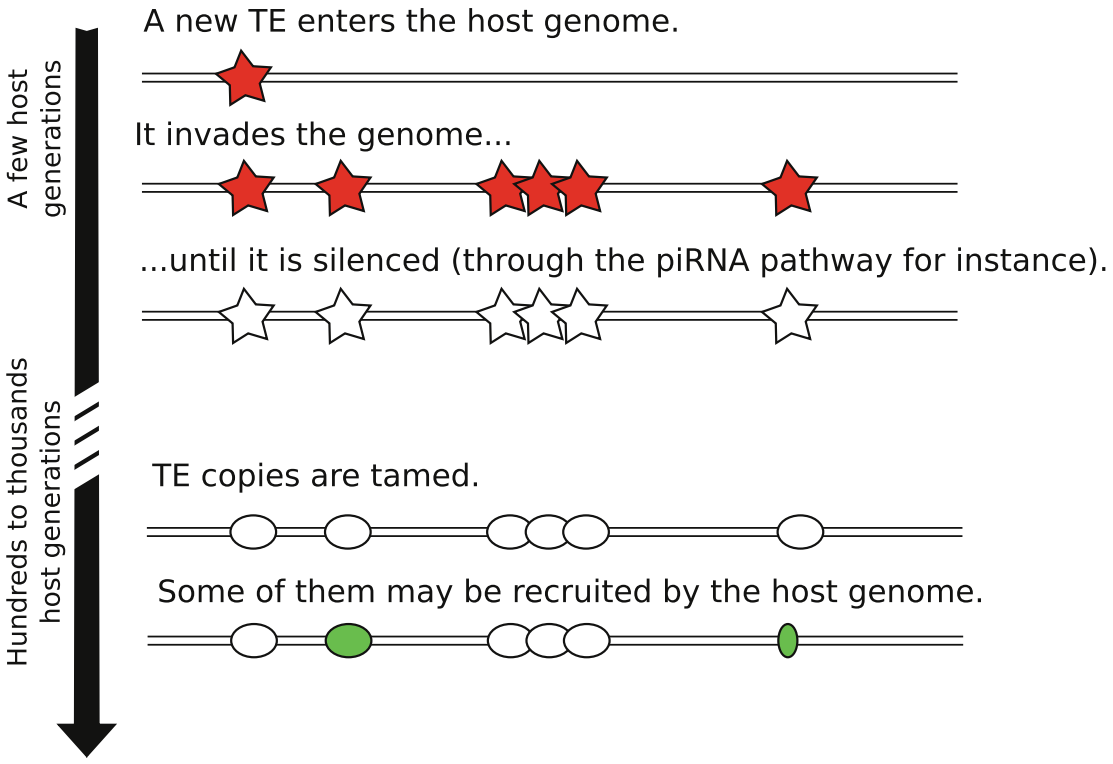


Fig. 4 A model for host genome/TE interactions. A new TE can enter the genome, for instance, due to horizontal transmission or to crossing between divergent genomes, as in the case of hybrid dysgenesis. At that time, it is recognized as genomic nonself. This TE will proliferate into the host genome up until one copy integrates into a piRNA cluster, which sets up transcriptional and posttranscriptional silencing. Controlled TE copies are the targets of heterochromatic marks, so are still identified as distinct from the canonical genome. However, they are no longer harmful. Some TE copies, or at least some parts of TE copies (e.g., transcription factor binding sites), may be recruited by the host genome. These copies are indicated in *green* in the figure. They now fulfill genomic functions. They are said to be domesticated and make part of genomic self

domesticated, while the others are silenced by the host genome. Such ambivalent behavior leads to complex and very sensitive interaction mechanisms—mostly through epigenetic pathways, and allows a wealth of implications of TEs in host genome evolution. Only the comprehensive knowledge of exact TE insertion sites in their epigenomic context will allow us to disentangle self from nonself in each TE-genome relationship.

References

Agrawal A, Eastman QM, Schatz DG (1998) Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394:744–751. doi:10.1038/29457

Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, Braman V, Chambeyron S, Vieira C (2013) Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. EMBO Rep 14:458–464. doi:10.1038/embor.2013.38

Aminetzach YT, Macpherson JM, Petrov DA (2005) Pesticide resistance via transposition-mediated adaptive gene truncation in Drosophila. Science 309:764–767. doi:10.1126/science.1112699

Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318:761–764. doi:10.1126/science.1146484

Belancio VP, Deininger PL, Roy-Engel AM (2009) LINE dancing in the human genome: transposable elements and disease. Genome Med. 1:97. doi:10.1186/gm97

Biémont C (2010) A brief history of the status of transposable elements: from junk DNA to major players in evolution. Genetics 186:1085–1093. doi:10.1534/genetics.110.124180

Biessmann H, Mason JM, Ferry K, d'Hulst M, Valgeirsdottir K, Traverse KL, Pardue ML (1990) Addition of telomere-associated HeT DNA sequences "heals" broken chromosome ends in Drosophila. Cell 61:663–673

Blaise S, de Parseval N, Bénit L, Heidmann T (2003) Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc. Natl. Acad. Sci. U. S. A. 100:13013–13018. doi:10.1073/pnas.2132646100

Blumenstiel JP, Erwin AA, Hemmer LW (2016) What Drives Positive Selection in the Drosophila piRNA Machinery? The Genomic Autoimmunity Hypothesis. Yale J. Biol. Med. 89:499–512

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128:1089–1103. doi:10.1016/j.cell.2007.01.043

Carnelossi EAG, Lerat E, Henri H, Martinez S, Carareto CMA, Vieira C (2014) Specific activation of an I-like element in Drosophila interspecific hybrids. Genome Biol. Evol. 6:1806–1817. doi:10.1093/gbe/evu141

Casacuberta E, González J (2013) The impact of transposable elements in environmental adaptation. Mol Ecol 22:1503–1517. doi:10.1111/mec.12170

Chambeyron S, Popkova A, Payen-Groschêne G, Brun C, Laouini D, Pelisson A, Bucheton A (2008) piRNA-mediated nuclear accumulation of retrotransposon transcripts in the Drosophila female germline. Proc. Natl. Acad. Sci. U. S. A. 105:14964–14969. doi:10.1073/pnas.0805943105

Charlesworth B, Langley CH (1989) The population genetics of Drosophila transposable elements. Annu Rev Genet 23:251–287. doi:10.1146/annurev.ge.23.120189.001343

Chuong EB, Elde NC, Feschotte C (2016) Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science 351:1083–1087. doi:10.1126/science.aad5497

Chuong EB, Elde NC, Feschotte C (2017) Regulatory activities of transposable elements: from conflicts to benefits. Nat Rev Genet 18:71–86. doi:10.1038/nrg.2016.139

Clark LA, Wahl JM, Rees CA, Murphy KE (2006) Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. Proc. Natl. Acad. Sci. U. S. A. 103:1376–1381. doi:10.1073/pnas.0506940103

Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH (2009) L1 retrotransposition in human neural progenitor cells. Nature 460:1127–1131. doi:10.1038/nature08248

Cowley M, Oakey RJ (2013) Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet 9:e1003234. doi:10.1371/journal.pgen.1003234

Craddock EM (2016) Profuse evolutionary diversification and speciation on volcanic islands: transposon instability and amplification bursts explain the genetic paradox. Biol. Direct 11:44. doi:10.1186/s13062-016-0146-1

David VA, Menotti-Raymond M, Wallace AC, Roelke M, Kehler J, Leighty R, Eizirik E, Hannah SS, Nelson G, Schäffer AA, Connelly CJ, O'Brien SJ, Ryugo DK (2014) Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. G3 Bethesda Md 4:1881–1891. doi:10.1534/g3.114.013425

Fablet M (2014) Host control of insect endogenous retroviruses: small RNA silencing and immune response. Viruses 6:4447–4464. doi:10.3390/v6114447

Fablet M, Akkouche A, Braman V, Vieira C (2014) Variable expression levels detected in the Drosophila effectors of piRNA biogenesis. Gene 537:149–153. doi:10.1016/j.gene.2013.11.095

Fablet M, Bueno M, Potrzebowski L, Kaessmann H (2009) Evolutionary origin and functions of retrogene introns. Mol Biol Evol 26:2147–2156. doi:10.1093/molbev/msp125

Fablet M, Vieira C (2011) Evolvability, epigenetics and transposable elements. BioMol Concepts 2:333–341

Guio L, Barrón MG, González J (2014) The transposable element Bari-Jheh mediates oxidative stress response in Drosophila. Mol Ecol 23:2020–2030. doi:10.1111/mec.12711

Han BW, Wang W, Li C, Weng Z, Zamore PD (2015) Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science 348:817–821

Hancks DC, Kazazian HH (2016) Roles for retrotransposon insertions in human disease. Mob DNA 7. doi:10.1186/s13100-016-0065-9

Hancks DC, Kazazian HH Jr (2012) Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. Molecular and genetic bases of disease 22:191–203. doi:10.1016/j.gde.2012.02.006

Hilditch L, Matadeen R, Goldstone DC, Rosenthal PB, Taylor IA, Stoye JP (2011) Ordered assembly of murine leukemia virus capsid protein on lipid nanotubes directs specific binding by the restriction factor, Fv1. Proc. Natl. Acad. Sci. U. S. A. 108:5771–5776. doi:10.1073/pnas.1100118108

Hoen DR, Bureau TE (2015) Discovery of novel genes derived from transposable elements using integrative genomic analysis. Mol Biol Evol 32:1487–1506. doi:10.1093/molbev/msv042

Hollister JD, Gaut BS (2009) Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res 19:1419–1428. doi:10.1101/gr.091678.109

Huang S, Tao X, Yuan S, Zhang Y, Li P, Beilinson HA, Zhang Y, Yu W, Pontarotti P, Escriva H, Le Petillon Y, Liu X, Chen S, Schatz DG, Xu A (2016) Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. Cell 166:102–114. doi:10.1016/j.cell.2016.05.032

Jensen PA, Stuart JR, Goodpaster MP, Goodman JW, Simmons MJ (2008) Cytotype regulation of P transposable elements in Drosophila melanogaster: repressor polypeptides or piRNAs? Genetics 179:1785–1793. doi:10.1534/genetics.108.087072

Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. Nature 457:405–412. doi:10.1038/nature07755

Jones BC, Wood JG, Chang C, Tam AD, Franklin MJ, Siegel ER, Helfand SL (2016) A somatic piRNA pathway in the Drosophila fat body ensures metabolic homeostasis and normal lifespan. Nat. Commun. 7:13856. doi:10.1038/ncomms13856

Jordan IK, Rogozin IB, Glazko GV, Koonin EV (2003) Origin of a substantial fraction of human regulatory sequences from transposable elements. Trends Genet. TIG 19:68–72

Kaessmann H, Vinckenbosch N, Long M (2009) RNA-based gene duplication: mechanistic and evolutionary insights. Nat Rev Genet 10:19–31. doi:10.1038/nrg2487

Karijolich J, Abernathy E, Glaunsinger BA (2015) Infection-Induced Retrotransposon-Derived Noncoding RNAs Enhance Herpesviral Gene Expression via the NF-κB Pathway. PLoS Pathog 11:e1005260. doi:10.1371/journal.ppat.1005260

Kelleher ES, Edelman NB, Barbash DA (2012) Drosophila interspecific hybrids phenocopy piRNA-pathway mutants. PLoS Biol 10:e1001428. doi:10.1371/journal.pbio.1001428

Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, Nowosielska A, Li C, Zamore PD, Weng Z, Theurkauf WE (2011) Adaptation to P element transposon invasion in Drosophila melanogaster. Cell 147:1551–1563. doi:10.1016/j.cell.2011.11.042

Kidwell MG (1977) Reciprocal differences in female recombination associated with hybrid dysgenesis in Drosophila melanogaster. Genet Res 30:77–88

Kim D-S, Huh J-W, Kim H-S (2007) Transposable elements in human cancers by genome-wide EST alignment. Genes Genet. Syst. 82:145–156

Klenov MS, Lavrov SA, Stolyarenko AD, Ryazansky SS, Aravin AA, Tuschl T, Gvozdev VA (2007) Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. Nucleic Acids Res 35:5430–5438. doi:10.1093/nar/gkm576

Kokošar J, Kordiš D (2013) Genesis and regulatory wiring of retroelement-derived domesticated genes: a phylogenomic perspective. Mol Biol Evol 30:1015–1031. doi:10.1093/molbev/mst014

Kolaczkowski B, Hupalo DN, Kern AD (2011) Recurrent adaptation in RNA interference genes across the Drosophila phylogeny. Mol Biol Evol 28:1033–1042. doi:10.1093/molbev/msq284

Labrador M, Farré M, Utzet F, Fontdevila A (1999) Interspecific hybridization increases transposition rates of Osvaldo. Mol Biol Evol 16:931–937

Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B (1988) On the role of unequal exchange in the containment of transposable element copy number. Genet Res 52:223–235

Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF (2013) Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 27:390–399. doi:10.1101/gad.209841.112

Le T-N, Schumann U, Smith NA, Tiwari S, Au PCK, Zhu Q-H, Taylor JM, Kazan K, Llewellyn DJ, Zhang R, Dennis ES, Wang M-B (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in Arabidopsis. Genome Biol 15:458. doi:10.1186/s13059-014-0458-3

Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen FM (1993) Transposons in place of telomeric repeats at a Drosophila telomere. Cell 75:1083–1093

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133:523–536. doi:10.1016/j.cell.2008.03.029

Lopez-Maestre H, Carnelossi EAG, Lacroix V, Burlet N, Mugat B, Chambeyron S, Carareto CMA, Vieira C (2017) Identification of misexpressed genetic elements in hybrids between Drosophila-related species. Sci. Rep. 7:40618. doi:10.1038/srep40618

Lu J, Clark AG (2010) Population dynamics of PIWI-interacting RNAs (piRNAs) and their targets in Drosophila. Genome Res 20:212–227. doi:10.1101/gr.095406.109

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137:522–535. doi:10.1016/j.cell.2009.03.040

Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. Cell 136:656–668. doi:10.1016/j.cell.2009.01.045

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A (2009) A transposon-induced epigenetic change leads to sex determination in melon. Nature 461:1135–1138. doi:10.1038/nature08498

McClintock B (1950) The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. U. S. A. 36:344–355

Meister G (2013) Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14:447–459. doi:10.1038/nrg3462

Melayah D, Bonnivard E, Chalhoub B, Audeon C, Grandbastien MA (2001) The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. Plant J. Cell Mol. Biol. 28:159–168

Metcalfe CJ, Bulazel KV, Ferreri GC, Schroeder-Reiter E, Wanner G, Rens W, Obergfell C, Eldridge MDB, O'Neill RJ (2007) Genomic instability within centromeres of interspecific marsupial hybrids. Genetics 177:2507–2517. doi:10.1534/genetics.107.082313

Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC, McCoy JM (2000) Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature 403:785–789. doi:10.1038/35001608

Micard D, Couderc JL, Sobrier ML, Giraud G, Dastugue B (1988) Molecular study of the retrovirus-like transposable element 412, a 20-OH ecdysone responsive repetitive sequence in Drosophila cultured cells. Nucleic Acids Res 16:455–470

Miller WJ, McDonald JF, Nouaud D, Anxolabéhère D (1999) Molecular domestication–more than a sporadic episode in evolution. Genetica 107:197–207

Mohn F, Handler D, Brennecke J (2015) Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science 348:812–817

Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. Nat Genet 23:314–318. doi:10.1038/15490

Muotri AR, Chu VT, Marchetto MCN, Deng W, Moran JV, Gage FH (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435:903–910. doi:10.1038/nature03663

Obbard DJ, Gordon KHJ, Buck AH, Jiggins FM (2009a) The evolution of RNAi as a defence against viruses and transposable elements. Philos Trans R Soc Lond B Biol Sci 364:99–115. doi:10.1098/rstb.2008.0168

Obbard DJ, Jiggins FM, Halligan DL, Little TJ (2006) Natural selection drives extremely rapid evolution in antiviral RNAi genes. Curr. Biol. CB 16:580–585. doi:10.1016/j.cub.2006.01.065

Obbard DJ, Welch JJ, Kim K-W, Jiggins FM (2009b) Quantifying adaptive evolution in the Drosophila immune system. PLoS Genet 5:e1000698. doi:10.1371/journal.pgen.1000698

O'Neill RJ, O'Neill MJ, Graves JA (1998) Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature 393:68–72. doi:10.1038/29985

Pélisson A, Song SU, Prud'homme N, Smith PA, Bucheton A, Corces VG (1994) Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. EMBO J 13:4401–4411

Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S (2013) Transposition-driven genomic heterogeneity in the Drosophila brain. Science 340:91–95. doi:10.1126/science.1231965

Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE (2003) Size matters: non-LTR retrotransposable elements and ectopic recombination in Drosophila. Mol Biol Evol 20:880–892. doi:10.1093/molbev/msg102

Picard G (1976) Non-mendelian female sterility in Drosophila melanogaster: hereditary transmission of I factor. Genetics 83:107–123

Pietzenuk B, Markus C, Gaubert H, Bagwan N, Merotto A, Bucher E, Pecinka A (2016) Recurrent evolution of heat-responsiveness in Brassicaceae COPIA elements. Genome Biol 17:209. doi:10.1186/s13059-016-1072-3

Prud'homme N, Gans M, Masson M, Terzian C, Bucheton A (1995) Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. Genetics 139:697–711

Rebollo R, Horard B, Hubert B, Vieira C (2010) Jumping genes and epigenetics: Towards new species. Gene 454:1–7. doi:10.1016/j.gene.2010.01.003

Rebollo R, Karimi MM, Bilenky M, Gagnier L, Miceli-Royer K, Zhang Y, Goyal P, Keane TM, Jones S, Hirst M, Lorincz MC, Mager DL (2011) Retrotransposon-induced heterochromatin spreading in the mouse revealed by insertional polymorphisms. PLoS Genet 7:e1002301. doi:10.1371/journal.pgen.1002301

Rebollo R, Romanish MT, Mager DL (2012) Transposable elements: an abundant and natural source of regulatory sequences for host genes. Annu Rev Genet 46:21–42. doi:10.1146/annurev-genet-110711-155621

Rook GAW (2012) Hygiene hypothesis and autoimmune diseases. Clin Rev Allergy Immunol 42:5–15. doi:10.1007/s12016-011-8285-8

Rozhkov NV, Hammell M, Hannon GJ (2013a) Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev 27:400–412. doi:10.1101/gad.209767.112

Rozhkov NV, Schostak NG, Zelentsova ES, Yushenova IA, Zatsepina OG, Evgen'ev MB (2013b) Evolution and dynamics of small RNA response to a retroelement invasion in Drosophila. Mol Biol Evol 30:397–408. doi:10.1093/molbev/mss241

Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev 20:2214–2222. doi:10.1101/gad.1454806

Senti K-A, Brennecke J (2010) The piRNA pathway: a fly's perspective on the guardian of the genome. Trends Genet. TIG 26:499–509. doi:10.1016/j.tig.2010.08.007

Sienski G, Dönertas D, Brennecke J (2012) Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151:964–980. doi:10.1016/j.cell.2012.10.040

Siomi H, Siomi MC (2015) RNA. Phased piRNAs tackle transposons. Science 348:756–757. doi:10.1126/science.aab3004

Siomi MC, Sato K, Pezic D, Aravin AA (2011) PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol 12:246–258. doi:10.1038/nrm3089

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8:272–285. doi:10.1038/nrg2072

Song X, Cao X (2017) Transposon-mediated epigenetic regulation contributes to phenotypic diversity and environmental adaptation in rice. Curr Opin Plant Biol 36:111–118. doi:10.1016/j.pbi.2017.02.004

Strachan DP (1989) Hay fever, hygiene, and household size. BMJ 299:1259–1260

Sutter NB, Bustamante CD, Chase K, Gray MM, Zhao K, Zhu L, Padhukasahasram B, Karlins E, Davis S, Jones PG, Quignon P, Johnson GS, Parker HG, Fretwell N, Mosher DS, Lawler DF, Satyaraj E, Nordborg M, Lark KG, Wayne RK, Ostrander EA (2007) A Single IGF1 Allele Is a Major Determinant of Small Size in Dogs. Science 316:112. doi:10.1126/science.1137045

Todeschini A-L, Teysset L, Delmarre V, Ronsseray S (2010) The epigenetic trans-silencing effect in Drosophila involves maternally-transmitted small RNAs whose production depends on the piRNA pathway and HP1. PLoS ONE 5:e11032. doi:10.1371/journal.pone.0011032

Uchiyama T, Hiura S, Ebinuma I, Senda M, Mikami T, Martin C, Kishima Y (2013) A pair of transposons coordinately suppresses gene expression, independent of pathways mediated by siRNA in Antirrhinum. New Phytol 197:431–440. doi:10.1111/nph.12041

Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. Science 313:320–324. doi:10.1126/science.1129333

van Valen L (1973) A new evolutionary law. Evol Theory 10:71–74

Van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ (2016) The industrial melanism mutation in British peppered moths is a transposable element. Nature 534:102–105. doi:10.1038/nature17951

Vela D, Fontdevila A, Vieira C, García Guerreiro MP (2014) A genome-wide survey of genetic instability by transposition in Drosophila hybrids. PLoS ONE 9:e88992. doi:10.1371/journal.pone.0088992

Vieira C, Lepetit D, Dumont S, Biémont C (1999) Wake up of transposable elements following Drosophila simulans worldwide colonization. Mol Biol Evol 16:1251–1255

Volff J-N (2006) Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. BioEssays News Rev. Mol. Cell. Dev. Biol. 28:913–922. doi:10.1002/bies.20452

Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng Z, Zamore PD (2015) Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. Mol Cell 59:819–830. doi:10.1016/j.molcel.2015.08.007

Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293–5300

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973–982. doi:10.1038/nrg2165

Yan Y, Buckler-White A, Wollenberg K, Kozak CA (2009) Origin, antiviral function and evidence for positive selection of the gammaretrovirus restriction gene Fv1 in the genus Mus. Proc. Natl. Acad. Sci. U. S. A. 106:3259–3263. doi:10.1073/pnas.0900181106

Zanni V, Eymery A, Coiffet M, Zytnicki M, Luyten I, Quesneville H, Vaury C, Jensen S (2013) Distribution, evolution, and diversity of retrotransposons at the flamenco locus reflect the regulatory properties of piRNA clusters. Proc. Natl. Acad. Sci. U. S. A. 110:19842–19847. doi:10.1073/pnas.1313677110

Zemojtel T, Vingron M (2012) P53 binding sites in transposons. Front. Genet. 3:40. doi:10.3389/fgene.2012.00040

Zhang F, Peterson T (2005) Comparisons of Maize pericarp color1 Alleles Reveal Paralogous Gene Recombination and an Organ-Specific Enhancer Region. Plant Cell 17:903–914. doi:10.1105/tpc.104.029660

Zhou Q, Ellison CE, Kaiser VB, Alekseyenko AA, Gorchakov AA, Bachtrog D (2013) The epigenome of evolving Drosophila neo-sex chromosomes: dosage compensation and heterochromatin formation. PLoS Biol 11:e1001711. doi:10.1371/journal.pbio.1001711