



**HAL**  
open science

# Mécanismes d'activation d'un piRNA cluster par des facteurs environnementaux et génétiques chez *Drosophila melanogaster*

Karine Casier

► **To cite this version:**

Karine Casier. Mécanismes d'activation d'un piRNA cluster par des facteurs environnementaux et génétiques chez *Drosophila melanogaster*. Génétique. Sorbonne Université, 2020. Français. NNT : 2020SORUS282 . tel-03412902

**HAL Id: tel-03412902**

**<https://theses.hal.science/tel-03412902>**

Submitted on 3 Nov 2021

**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.

# Sorbonne Université

Ecole doctorale Complexité du Vivant (ED515)

*Laboratoire de Biologie du Développement (UMR7622) – Institut de Biologie Paris Seine  
(IBPS) / Equipe « Epigénétique transgénérationnelle et biologie des petits ARN »*

## **Mécanismes d'activation d'un piRNA cluster par des facteurs environnementaux et génétiques chez *Drosophila melanogaster***

Par Karine Casier

Thèse de doctorat de Génétique

Dirigée par Antoine Boivin et Laure Teyssset

Présentée et soutenue publiquement le 7 septembre 2020

Devant un jury composé de :

Pr. Dominique Higuët	Professeur	Président du jury
Dr. Benjamin Loppin	Directeur de Recherche	Rapporteur
Dr. Mireille Bétermier	Directrice de Recherche	Rapporteur
Dr. Pascale Lesage	Directrice de Recherche	Examineur
Dr. Aurélie Hua-Van	Maitre de Conférences	Examineur
Dr. Jean-René Huynh	Directeur de Recherche	Examineur
Dr. Antoine Boivin	Maitre de Conférences	Co-encadrant
Dr. Laure Teyssset	Maitre de Conférences	Directeur de thèse

*« La réponse à la grande question de la biologie, des piRNA et du reste est...42AB »*

# Remerciements

Je tiens d'abord à remercier les membres du jury pour avoir accepté d'évaluer mes travaux de thèse, en particulier Mireille Bétermier et Benjamin Loppin qui ont accepté d'être rapporteurs du manuscrit.

Un grand merci à toute l'équipe TErBio pour ces bons moments passés, ces repas et bières partagés (mon futur diabète vous remercie lui aussi), ces pots (ça fait longtemps qu'on n'a pas fait de pot, d'ailleurs...non ?), ces fous rires, toutes ces conversations (drôles ou sérieuses), les réunions p'tit-déj...J'ai adoré cette harmonie d'équipe, cette super ambiance de travail, vous allez tous me manquer !

Antoine, merci pour tout. Tu m'as encadrée depuis le début : L3, fin de M2, thèse...et ce, de gré ou de force ! Tu m'as toujours accordé du temps, donné de bons conseils et soutenue, tu m'as appris les valeurs de la science, la rigueur scientifique, ne jamais s'emballer trop vite...ou même trop lentement...en fait, ne jamais s'emballer du tout, toujours se remettre en question, toujours refaire ses manip (10 fois, 100 fois, 1000 fois), toujours faire les bons contrôles (10, 100, 1000 contrôles). Je te considère comme mon « papa scientifique » maintenant. J'ai adoré travailler avec toi, ta bonne humeur, ton sens de l'humour décalé (un peu comme le mien je dois dire), nos discussions littéraires (Lovecraft n'est jamais loin), cinématographiques, vidéo-ludiques et autres questionnements existentiels et métaphysiques (sans trop se prendre au sérieux, tout de même). On va boire une bière ? Une bière Julius, par exemple ?

Laure, merci pour ta bonne humeur, ta bienveillance, tes conseils, ton soutien et ces fous rires. C'est grâce à toi que j'ai pu faire mon premier stage dans l'équipe et par extension cette thèse durant laquelle je me suis littéralement éclatée (dans le bon sens du terme, évidemment !). C'est aussi de t'avoir eu en TD/TP qui m'a donné l'envie d'enseigner et ça a été un vrai plaisir pour moi !

Val...mais qu'aurait été cette thèse sans toi ? Travailler avec toi a été un véritable plaisir. Toujours de bonne humeur, toujours en pleine forme (même après une nuit blanche) et toujours prête à me donner un coup de main...avec ces milliers (ce n'était pas des millions, en fait ?) de mouches à disséquer matin, midi et soir, jusqu'à en rêver la nuit ! On est vraiment sur la même longueur d'onde et j'ai adoré toutes ces discussions qu'on a pu avoir : littérature, cinéma, musique, politique, cuisine et chiffons, et les conversations sans queue ni tête. Autour d'un café (ou deux, ou trois, ou quatre, mais pas cinq, ça rend nerveux), d'un thé, d'un chocolat, d'une bière...ou devant la loupe, pour ces milliers de dissection ! Avec qui vais-je danser sur la table, pour le plaisir ou pour fêter nos papiers acceptés, maintenant ? Si tu étais un personnage de Star Wars, on t'appellerait Dark Val-d'Or ! Merci Val, merci beaucoup.

Stéphane, merci de m'avoir accueillie dans ton équipe la première fois...et la suivante ! J'aurais beaucoup aimé entendre tes imitations de singe ou de vache (ou tout autre animal)

pendant ma thèse. Ces dernières années n'ont pas été faciles pour toi, je souhaite que ton rétablissement se poursuive de la meilleure façon possible. Pour moi, tu es l'image du scientifique passionné, un peu fou-fou et toujours intègre...un modèle que j'aimerais voir perdurer !

Dilyana, on a commencé notre thèse ensemble et on la termine ensemble, quelle belle aventure ! Semée d'embûches, certes, mais mémorable quand même ! Que de souvenirs : congrès (non, plus de sandwich, merci, une soupe sera parfaite), qPCR, rigolades, jardinage, administration (ha ha ha...), inscriptions (surtout pour la 4<sup>ème</sup> année, c'est quand même le niveau ++ !). Merci !

Mira, toujours de bonne humeur, pleine de motivation et toujours en forme ! Merci, je n'oublierais pas ta bonne humeur, tes super tee-shirts ou encore tes délicieux 'nems' au poulet. Et puis, enfin quelqu'un qui comprend ma passion pour le Metal, ouf, je ne suis pas le seul spécimen de ce labo ! Je te souhaite une super thèse, j'espère que tu vas t'épanouir (mais je n'ai pas trop de doutes à ce sujet).

Bruno, 'Lord of the Flies', maître des jeux de société, gestionnaire des escape-games, des repas d'équipe et compagnon de badminton ! Merci pour tout ça...mais aussi pour ta bonne humeur, ta gentillesse, ta disponibilité...et pour m'avoir appris les smashes (ou plutôt d'avoir essayé, uhuh). Je te pardonne ton plus gros défaut : ne pas aimer le chocolat.

Clément, merci pour ta bonne humeur, tes remarques et ton soutien. Avec toi, on sait aussi où trouver les bonnes bouteilles, cachées entre les plantes et les livres !

Julie, on n'aura pas eu beaucoup l'occasion d'interagir...mais j'ai apprécié ta bonne humeur et ton côté fou-fou, je te souhaite une très bonne continuation (en thèse, j'espère !).

Merci à toute l'équipe Peronnet : Héloïse, Delphine, Valérie, Marco, Jean-Michel, Emmanuelle, Hélène, Frédérique...et Sandra ! (Tu resteras une 'Pero' pour moi). Pour les bons moments, les discussions et la bonne ambiance !

Sandra, ça fait déjà deux ans que tu as soutenu ta thèse et tes rires résonnent encore entre les murs du labo, ta bonne humeur est toujours dans l'air ambiante...tu nous as manqués ! On a toutes les deux travaillé sur l'effet de la température et l'effet d'une protéine à domaine Jumonji (pas dans les mêmes processus, mais quand même...). On a passé tellement de bons moments au boulot : les repiquages, les pauses thé/chocolat, les voyages (en particulier celui de Londres ! Tu reprendras bien un petit sandwich ?), les congrès/colloques/ séminaires/ DIF/BIS/JBE, les discussions sur tout et n'importe quoi, et surtout n'importe quoi (les accessoires GFP, par exemple, uhuh) ; et ailleurs avec Pierre : bars, restaurants, jeux, imitations, fous rires...raclettes !

Héloïse, aka Bat Girl, princesse Disney en chef, vivant dans un monde à paillettes peuplé de chats et de licornes...Merci pour tous ces bons moments passés, à rigoler, à se plaindre, à

manger ou à regarder des photos de ton chat (des fois tout en même temps). Je n'oublierai pas nos aventures à Montpellier et à Lausanne : où est-ce qu'on va ? Ça commence à quelle heure ? Ah, il pleut... Au fait, ce n'est pas l'heure du thé par hasard ? Est-ce que tu veux un thé ? Vert, citron, vanille ou exotique ? Tu veux des gâteaux ?...De la raclette ? Je crois même que c'est bientôt l'heure de l'apéro... !

Delphine...Delphine ? Delphiiiiine, tu es là ?? Ah, Delphine !!! J'ai une question ! Ah, qu'est ce que tu l'auras entendu cette phrase...Merci pour avoir été ma Wonder Woman, j'ai toujours pu compter sur toi ! Je voterai pour toi à l'élection de président(e) de l'Univers ! Merci aussi pour ta bonne humeur constante, ton rire communicatif et tes foules d'anecdotes des fois surprenantes, des fois étranges, mais toujours drôles !

Valérie, merci pour nos discussions marrantes dans la pièce à mouche, les échanges sur les films et séries (c'est grâce à toi que j'ai découvert Westworld !), pour ton soutien...je n'oublierai pas toutes ces micro-pauses ou les cafés/thés post-repas accompagnés d'un chocolat (caramel et pointe de sel, s'il vous plait !).

Marco, séducteur de ses dames ! Un score de 234, ça se mérite, quand même ! Merci pour ton soutien, ta gentillesse, tous ces moments de rigolades...et aussi tous les petits gâteaux cachés dans ton tiroir pour nous éviter l'hypoglycémie. Nos conversations à double sens et nos blagues graveleuses vont me manquer.

Merci à l'équipe Gho avec qui nous avons partagé la pièce à mouche et certains pots ! Merci Pénélope pour toutes les rimes en U et les discussions étranges. Tiens, pour la peine, moi aussi j'ai fait un poème :

'La thèse durant, nous nous sommes soutenues  
Nous avons traversé ensemble ce chemin ardu  
Avec ces épreuves parfois tendues  
Mais c'est là que cette aventure se conclue  
Après le manuscrit rendu  
Et la thèse bientôt défendue  
Laisse-moi te dire : merci beaucu' (pour la rime)

Merci aussi à tous les membres de l'UMR7622 !

Merci Ninou d'être là, à mes côtés depuis plus de 10 ans ! De me supporter, de me soutenir dans les bons et les mauvais moments, de m'accompagner dans mes répliques de film ou de jeux vidéo, d'écouter mes diatribes scientifico-politico-métaphysiques et mes blagues douteuses. Et surtout, on n'oublie pas, à 16h mini Zurkon doit vite rentrer prendre son goûter...est-ce qu'on pourra avoir des glaces ? Merci aussi pour avoir relu mon manuscrit...même si tu n'y as rien compris ! Merci également à tous les membres de ma famille, celle qui me soutient et m'encourage, ainsi qu'à mes proches...vous allez pouvoir m'appeler Docteur, si ça vous amuse !

Et pour finir, ne les oublions pas : Merci les mouches !



# Préambule

Il était une fois, la drosophile... Voici un conte qui dure depuis près de 150 ans, commencé à la fin du XIX<sup>ème</sup> siècle par Thomas Morgan, parce qu'il trouvait sans doute qu'elles avaient de beaux yeux.

La drosophile, aussi appelée 'mouche du vinaigre', est un organisme modèle très utilisé dans les laboratoires et si elle rencontre un tel succès, ce n'est évidemment pas sans raison : une petite taille (environ 3 mm), facile à entretenir (nourriture, reproduction, maintien des individus), un cycle de vie court (environ 10 jours à 25°C) et un génome relativement simple (4 paires de chromosomes). Ces caractéristiques en font un modèle de choix dans tous les domaines de la biologie, mais s'il y a bien une discipline pour laquelle la drosophile est particulièrement appréciée c'est celle de la génétique. En 150 ans de travaux, la communauté drosophiliste a pu accumuler de nombreuses connaissances et développer un large éventail d'outils qui facilitent grandement l'avancée des recherches dans ce domaine.

L'histoire de la drosophile s'apparente à une vraie 'success story'. Neuf prix Nobel de physiologie ou médecine à son actif, c'est plus que le nombre de victoires remportées par Lance Armstrong sur le tour de France (et sans dopage, encore). Un courageux triomphe face à l'invasion de l'élément *P*, un autre protagoniste dont nous reparlerons plus tard, entre les années 1950 et 1970. Des milliers d'adeptes à travers le monde, qui lui vouent un culte sans nom dans ces lieux semi-obscur qu'on appelle vulgairement 'pièces à mouche', le grand Cthulhu lui-même ne rencontre pas une telle popularité.

De cette grande saga inachevée dont la drosophile est l'héroïne, intéressons-nous à l'un des chapitres, aussi petit soit-il, qui mentionne l'histoire d'un personnage secondaire et qui débute comme ceci : 'Un beau jour de l'année 2016, une jeune étudiante en génétique allait entreprendre une aventure pleine de surprises et de rebondissements'. Cette thèse raconte l'histoire de cette étudiante qui s'est intéressée à l'effet de l'environnement et de gènes sur la biologie des piRNA clusters. Mais n'en dévoilons pas trop ici, il serait dommage de gâcher le suspense.



# Sommaire

<b>Remerciements</b> .....	<b>3</b>
<b>Préambule</b> .....	<b>7</b>
<b>Sommaire</b> .....	<b>8</b>
<b>Introduction</b> .....	<b>13</b>
<b>I. Les Éléments Transposables</b> .....	<b>15</b>
1. Les différentes classes d'éléments transposables.....	16
2. L'impact des éléments transposables sur le génome.....	19
2.1. Conséquences structurales et fonctionnelles .....	19
2.2. Les éléments transposables, un nid à problèmes !.....	23
2.3. Les éléments transposables : un rôle dans l'évolution ? .....	25
3. Le cas de l'élément <i>P</i> .....	28
3.1. Présentation de l'élément <i>P</i> .....	28
3.2. La dysgénésie des hybrides .....	29
3.3. Régulation de l'élément <i>P</i> .....	30
<b>II. Les PIWI-Interacting RNA</b> .....	<b>33</b>
1. Régulation des ET par les piRNA .....	34
1.1. La lignée germinale de la drosophile .....	34
1.2. Les séquences productrices de piRNA.....	35
- Les piRNA clusters « uni-strand » .....	36
- Les piRNA clusters « dual-strand » .....	38
- Autres sources de piRNA .....	41
1.3. Biogénèse des piRNA chez <i>Drosophila melanogaster</i> .....	44
- Export des ARN précurseurs.....	44
- piRNA primaires et régulation transcriptionnelle .....	48
- piRNA secondaires et régulation post-transcriptionnelle.....	53
1.4. Les piRNA chez d'autres espèces .....	57
2. D'autres fonctions pour les piRNA .....	58
2.1. Régulation de gènes .....	58
2.2. Défense virale.....	60
2.3. Les piRNA, biomarqueurs de maladies ?.....	61

3.	La biologie des piRNA clusters .....	63
3.1.	Des outils d'étude de la biologie des piRNA clusters .....	63
-	Transgènes « cibles » .....	63
-	Le transgène <i>P(lArB)</i> .....	64
-	Le cluster <i>BX2</i> .....	66
3.2.	Paramutation et effets transgénérationnels.....	69
<b>III.</b>	<b>L'épigénétique transgénérationnelle et le rôle des piRNA .....</b>	<b>71</b>
1.	Description de la revue.....	71
2.	Review: Environmentally-Induced Transgenerational Epigenetic Inheritance: Implication of PIWI Interacting RNAs .....	72
<b>IV.</b>	<b>Question biologique.....</b>	<b>95</b>
<b>Résultats</b>	<b>.....</b>	<b>97</b>
<b>I.</b>	<b>Effet de facteurs environnementaux sur l'état épigénétique de <i>BX2</i> .....</b>	<b>98</b>
<b>1.</b>	<b>Effet de la chaleur .....</b>	<b>98</b>
1.1.	Présentation de l'Article I .....	99
1.2.	Article I: Environmentally-induced epigenetic conversion of a piRNA cluster .....	100
2.	Effet du froid .....	156
3.	Autres stress .....	162
3.1.	Sel (NaCl).....	162
3.2.	Cuivre (CuSO <sub>4</sub> ) .....	163
3.3.	Bisphénol A.....	164
<b>II.</b>	<b>Effet de facteurs génétiques dans l'activation de <i>BX2</i> .....</b>	<b>166</b>
1.	Perte de fonction de <i>hfp</i> et activation de <i>BX2</i> .....	166
1.1.	Description du crible génétique.....	166
1.2.	Analyse de la perte de fonction de <i>hfp</i> .....	167
2.	Perte de fonction de <i>JHDM2</i> et activation de <i>BX2</i> .....	171
2.1.	Identification et régulation de <i>JHDM2</i> .....	175
2.2.	<i>hfp</i> et <i>JHDM2</i> dans l'activation de <i>BX2</i> .....	175
3.	Etude de la perte de fonction de <i>JHDM2</i> .....	178
3.1.	<i>JHDM2</i> <i>GLKD</i> et effets pan-génomiques .....	178
3.2.	La perte de fonction de <i>JHDM2</i> affecte-t-elle la génération suivante ?.....	182
<b>III.</b>	<b>La dynamique des piRNA clusters .....</b>	<b>187</b>
1.	Présentation de l'article II .....	187

2. Article II: Modelling early germline immunization by transposable element cooption reveals internal piRNA cluster heterogeneity .....	188
<b>Discussion et perspectives .....</b>	<b>247</b>
<b>I. Comment définir un piRNA cluster .....</b>	<b>248</b>
1. Définition biologique .....	248
2. Définition bioinformatique .....	249
<b>II. Activation d'un piRNA cluster .....</b>	<b>252</b>
<b>III. Effets pan-génomiques .....</b>	<b>254</b>
1. Augmentation de température et production de piRNA .....	254
2. Environnement, stress et piRNA clusters .....	256
3. Inactivation de gènes et production de piRNA .....	257
4. Fonctions de <i>JHDM2</i> .....	259
<b>IV. Être ou ne pas être un piRNA cluster : des facteurs pour spécifier l'identité des piRNA clusters .....</b>	<b>261</b>
<b>V. Dynamique des piRNA clusters et protection du génome .....</b>	<b>262</b>
1. Arrivée d'un ET et invasion du génome .....	263
2. Mise en place et maintien de la régulation .....	264
3. Structures des <i>TAS</i> et régulation du génome .....	267
<b>VI. Comprendre la biologie des piRNA clusters .....</b>	<b>268</b>
<b>Matériel et méthodes .....</b>	<b>271</b>
<b>I. Ressources et outils d'analyses .....</b>	<b>272</b>
<b>II. Génétique .....</b>	<b>272</b>
1. Balanceurs et transgènes .....	272
2. Lignées .....	274
3. Croisement et maintien des stocks .....	277
<b>III. Histochimie .....</b>	<b>277</b>
1. Coloration $\beta$ -galactosidase .....	277
2. Analyse des colorations .....	277
<b>IV. Biologie Moléculaire .....</b>	<b>278</b>
1. Extraction d'ARN .....	278
2. RT-qPCR et RT-PCR .....	278
3. Small RNA-sequencing & RNA-sequencing .....	279
4. RNA-IP .....	279
5. Western Blot .....	280

6. ChIP-qPCR.....	281
7. Liste des amorces .....	283
<b>Bibliographie.....</b>	<b>287</b>
<b>Annexes.....</b>	<b>307</b>
<b>Tables des figures .....</b>	<b>336</b>
<b>Tables des tables .....</b>	<b>338</b>
<b>Résumés .....</b>	<b>340</b>

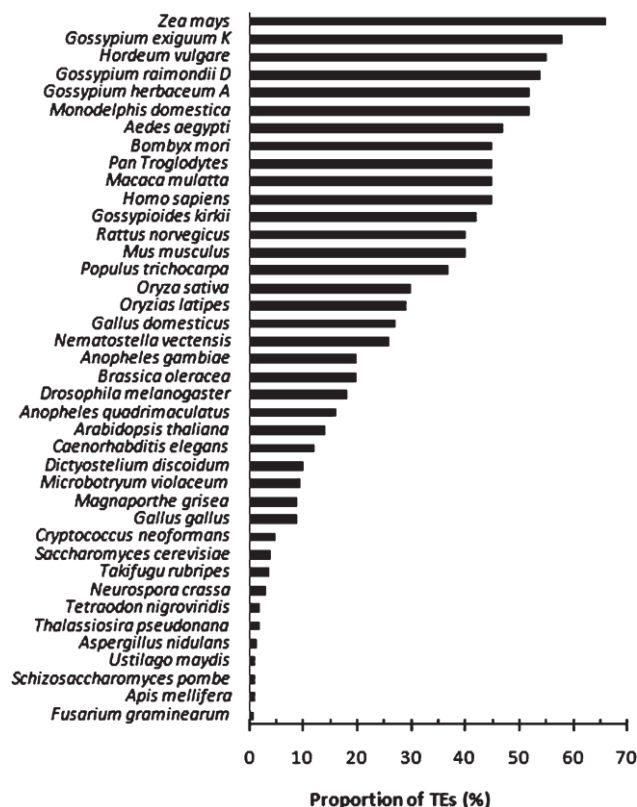


# **INTRODUCTION**



# I. Les Éléments Transposables

Les Éléments Transposables (ET) sont des séquences d'ADN mobiles capables de se déplacer aléatoirement au sein du génome grâce à un mécanisme nommé transposition. Ces séquences particulières ont été découvertes pour la première fois chez le maïs par Barbara McClintock en 1950 (McCLINTOCK, 1950). Le développement des techniques de séquençage à haut débit a permis de mettre en évidence que ces ET constituent une part plus ou moins importante des génomes du monde vivant (Figure 1). Certains génomes contiennent peu de séquences d'ET, par exemple chez la levure *Schizosaccharomyces pombe* où les ET ne représentent pas plus de 5% du génome, alors que d'autres génomes en contiennent énormément, c'est le cas du maïs *Zea mays* où les séquences d'ET constituent près de 70% du génome. On peut noter que chez *Drosophila melanogaster* et chez l'Homme, ces proportions sont respectivement de 21% et de 45% (Biémont, 2010).

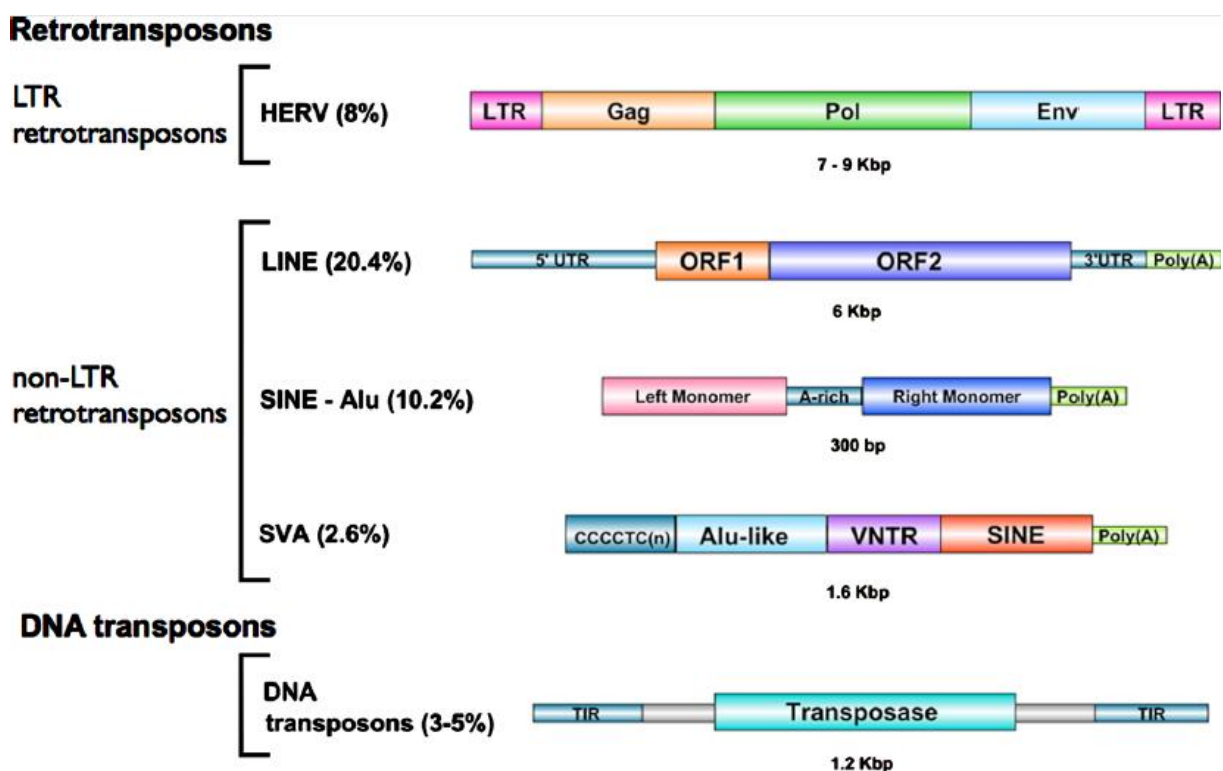


**Figure 1 :** Exemples de la proportion d'ET présents dans le génome de différents organismes. La proportion en ET varie d'une espèce à l'autre, ainsi dans certains génomes on trouve peu d'ET (*Fusarium graminearum*, *Apis mellifera* ou encore *Schizosaccharomyces pombe* avec moins de 5%) alors que pour d'autres les ET en constituent une part très importante (*Zea mays*, *Gossypium exiguum K* ou *Hordeum vulgare* avec 60% et plus) (Biémont, 2010).



# 1. Les différentes classes d'éléments transposables

Les ET peuvent être classés selon différents critères. La méthode de classification communément utilisée est celle proposée par Finnegan (Finnegan, 1989). Elle tient compte de leur structure et de leur mécanisme de transposition. Ainsi, on distingue deux grandes classes d'ET : les ET de classe I (ou rétrotransposons), et les ET de classe II (ou transposons).

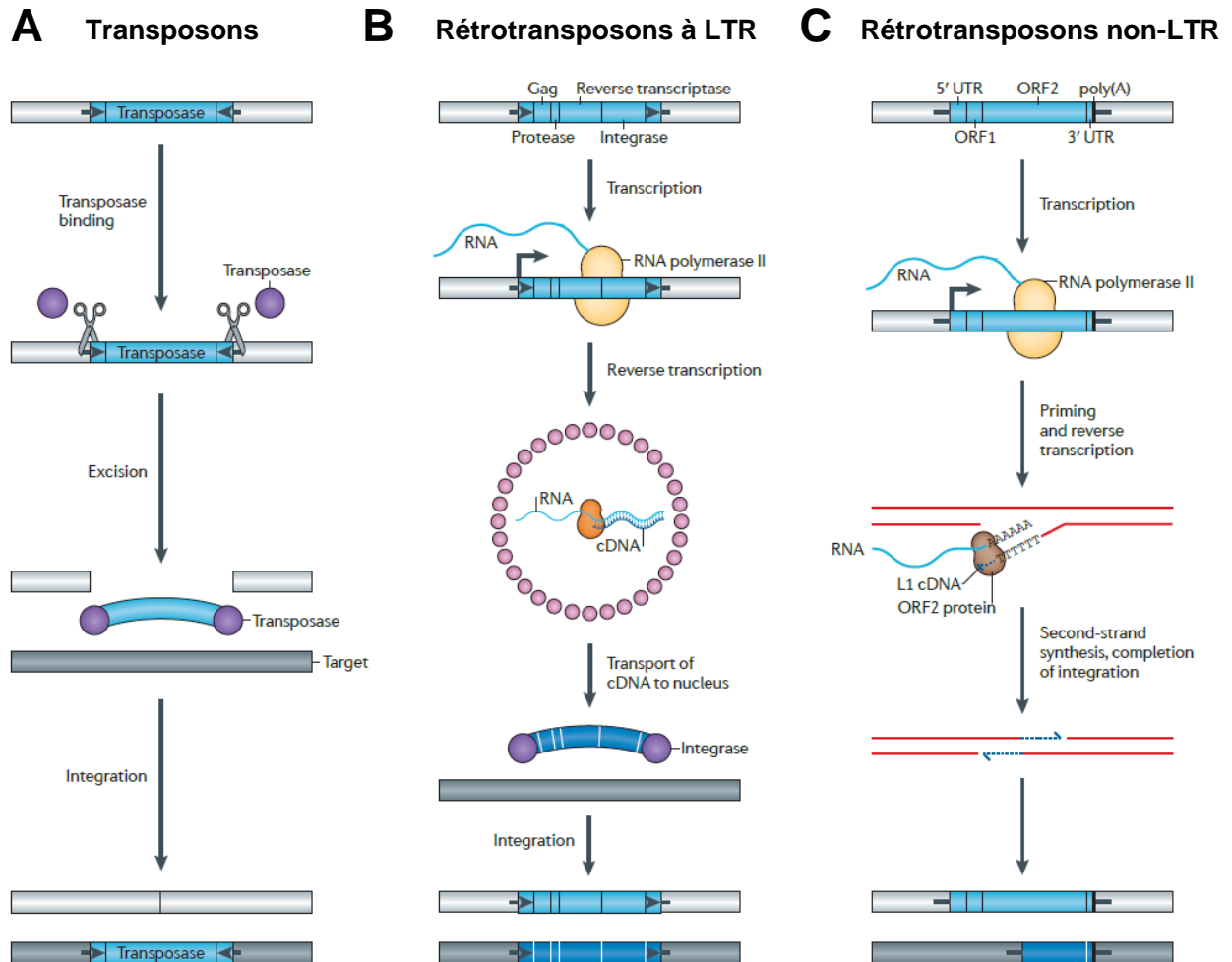


**Figure 2 : Classification et structures des éléments transposables.** Les ET sont divisés en deux grandes classes : les ET de classe I (ou rétrotransposons) et les ET de classe II (ou transposons). LTR: Long Terminal Repeat, HERV: Human Endogenous RetroVirus, LINE: Long Interspersed Nuclear Elements, SINE: Short Interspersed Nuclear Elements, SVA: SINE-R/VNTR/Alu-like. Le pourcentage indiqué entre parenthèse correspond à la proportion de chaque élément au sein du génome humain (Guffanti *et al*, 2014).

Les éléments transposables de classe II, aussi appelés ‘transposons à ADN’, possèdent majoritairement un mécanisme de transposition de type « couper-coller ». Les transposons codent une protéine, la transposase, et contiennent des séquences inversées répétées TIR (Terminal Inverted Repeat) (Figure 2). La transposase permet l’excision de l’élément par

reconnaissance des séquences TIR puis son insertion à une autre position au sein du génome (Figure 3A) (Finnegan, 1992).

Les éléments transposables de classe I, aussi appelés rétrotransposons, sont subdivisés en deux classes : les rétrotransposons à LTR (Long Terminal Repeat) qui possèdent de longues séquences répétées non inversées à leurs extrémités et les rétrotransposons non-LTR qui n'en possèdent pas (Figure 2). Les rétrotransposons à LTR ressemblent structurellement et fonctionnellement aux génomes de rétrovirus endogènes, ils possèdent des cadres ouverts de lecture (ou ORF pour Open Reading Frame) codant les gènes *gag*, *pol* et *env* (Boeke & Chapman, 1991). *pol* possède des domaines permettant l'activité reverse transcriptase et intégrase, nécessaires à la transposition (Figure 3B). Les rétrotransposons non-LTR sont composés de trois grandes familles : les LINE (Long Interspersed Nuclear Elements), les SINE (Short Interspersed Nuclear Elements) et SVA (SINE-R/VNTR/Alu-like). Les ET de type LINE sont constitués d'une région 5'UTR contenant un promoteur interne ainsi que de deux ORFs (Figure 2). L'ORF2 code une endonucléase et une reverse transcriptase (Moran *et al*, 1996). Les SINE sont des ET dérivés des LINE non-autonomes. En effet, contrairement aux LINE, les SINE ont perdu la capacité de se déplacer par eux-mêmes au sein du génome car ils ne possèdent plus les séquences qui le leur permettent mais ceux-ci peuvent transposer grâce à la machinerie des LINE. Enfin, les SVA sont des SINE composites, eux aussi non-autonomes, contenant des séquences Alu-like et une région répétée riche en C-T. Les rétrotransposons possèdent un mécanisme de transposition de type « Copier-coller ». En effet, leur mobilité nécessite la production d'un ARN intermédiaire. Celui-ci est transcrit de façon classique à partir de la séquence d'ET puis reverse-transcrit en ADN complémentaire (ADNc), grâce à l'activité d'une transcriptase inverse codée par l'ET (Figure 3B et 3C) (Finnegan, 1992). Ce mécanisme de transposition permet l'augmentation du nombre de copies d'ET au sein du génome. Il a effectivement été montré que l'accroissement de la taille du génome humain est essentiellement dû à des événements de rétrotransposition (Liu *et al*, 2003).



**Figure 3 : Les différents mécanismes de transposition.** (A) Mécanisme de transposition d'un ET de classe I. La transposase est synthétisée et reconnaît les séquences TIR aux extrémités 5' et 3' de la séquence. Le transposon est alors excisé puis inséré de manière aléatoire dans le génome. (B) Mécanisme de transposition d'un rétrotransposon à LTR. L'élément possède des ORF codant *gag*, *pol* et *env*. La séquence LTR en 5' contient un promoteur reconnu par l'ARN polymérase II, l'élément est transcrit en ARNm. Des protéines Gag (codées par *gag*) s'assemblent pour former une structure proche de celle des particules virales contenant l'ARNm, la reverse transcriptase et l'intégrase. L'ARNm est reverse transcrit en ADNc. L'intégrase permet alors l'insertion de l'ADNc dans le génome. (C) Mécanisme de transposition d'un rétrotransposon non-LTR. L'élément code deux ORFs, avec une activité reverse transcriptase et endonucléase, et est transcrit par l'ARN polymérase II. Ici, la mobilisation se fait par un mécanisme appelé « target-site-primed reverse transcription » (TPRT). L'endonucléase génère une cassure simple-brin au niveau de l'ADN génomique, libérant ainsi une extrémité 3'OH servant d'amorce à la reverse transcription. L'ADNc est intégré en même temps qu'il est synthétisé (Levin & Moran, 2011).

## **2. L'impact des éléments transposables sur le génome**

Compte tenu de la proportion d'ET présents dans les génomes, qu'ils soient autonomes ou non, et de leur diversité, il paraît légitime de se demander : les ET ont-ils un rôle ? Et, si oui, quel est-il ?

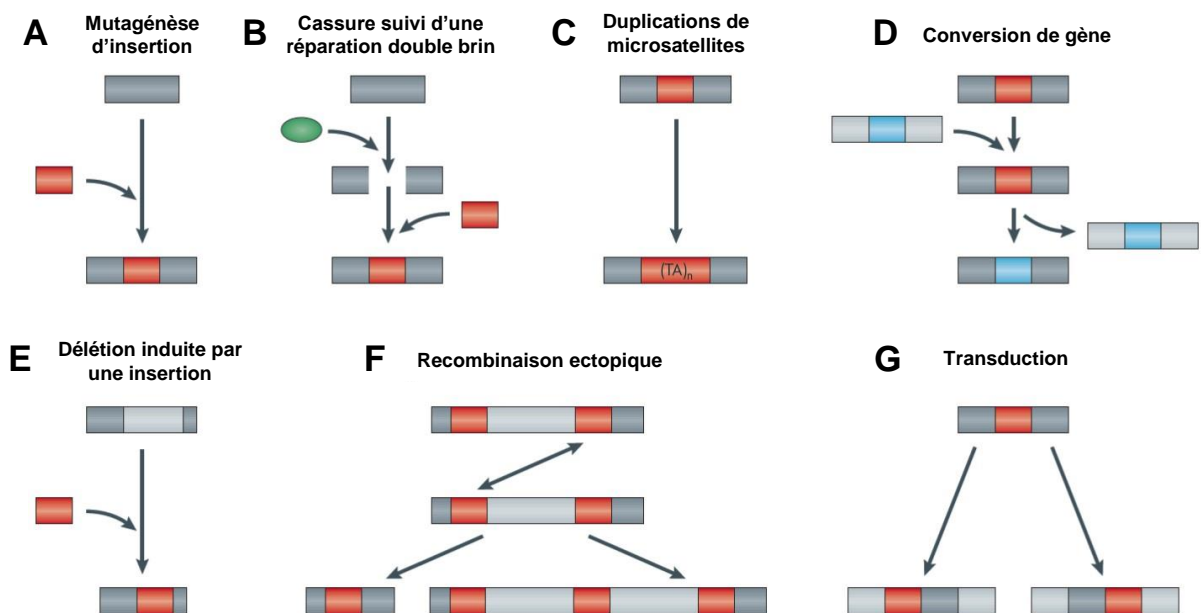
Les ET ont d'abord été classés dans la catégorie de l'ADN « poubelle » (« Junk DNA ») par Susumu Ohno (Ohno, 1972). Il considérait qu'une grande partie du génome était constituée de séquences sans fonction apparente, que nous connaissons actuellement sous le terme d'ADN non-codant. Par la suite, ils ont été considérés comme des éléments « parasites » et « égoïstes » par les auteurs de deux publications conjointes dans *Nature* en 1980 (Doolittle & Sapienza, 1980; Orgel & Crick, 1980). Les ET se servent du génome hôte pour se multiplier sans que ce dernier ne bénéficie d'avantages sélectifs, ce qui représente une dépense énergétique importante et inutile. Bien qu'il ne soit pas clairement établi, les auteurs n'excluent pas un rôle possible des ET dans l'évolution, notamment par l'acquisition possible de nouvelles fonctions. Pourtant, il apparaît assez facilement que la dynamique des éléments transposables peut avoir des conséquences importantes sur la structure et l'évolution des génomes. Des travaux montrent que ces conséquences peuvent être néfastes et générer des maladies tandis que d'autres semblent indiquer que les ET sont aussi responsables d'innovations génétiques pouvant participer à l'adaptation des espèces.

### **2.1. Conséquences structurales et fonctionnelles**

L'activité des éléments transposables peut modifier la structure des génomes de diverses manières. Des bilans de ces différents mécanismes ont été proposés pour les transposons et pour les rétrotransposons (Cordaux & Batzer, 2009; Feschotte & Pritham, 2007).

L'intégration des ET peut se faire dans des régions codantes ou dans des régions régulatrices (Figure 4A). Pour permettre leur insertion, les rétrotransposons non-LTR codent pour une endonucléase. Seulement, l'activité des endonucléases peut générer des cassures doubles brins conduisant à l'activation des acteurs moléculaires impliqués dans la réparation de l'ADN (Figure 4B). La réparation peut se faire par jonction d'extrémités non homologues ou par recombinaison homologue, générant des différences par rapport à la séquence d'origine dans le premier cas mais pas dans le second (Symington & Gautier, 2011). L'insertion d'un

rétrotransposon dans une séquence peut aussi générer des délétions plus ou moins grandes de régions adjacentes, causées par des mécanismes endonucléase-dépendants (Figure 4E). Des séquences peuvent également être ajoutées lors de l'insertion par des phénomènes de transduction (Figure 4G). Ces phénomènes ont lieu lorsque les régions flanquantes en 5' ou en 3' du rétrotransposon sont rétrotranscrites avant d'être excisées puis insérées en même temps que l'élément. Les rétrotransposons non-LTR contiennent de nombreuses séquences microsatellites notamment des séquences poly(A) à leur extrémité 3' et, dans le cas des SINE, au milieu de l'élément (Voir figure 2). Ainsi, ces courtes séquences répétées peuvent se dupliquer après l'insertion augmentant ainsi le nombre de répétitions (Figure 4C). La présence de ces séquences répétées peut engendrer de la recombinaison homologue, soit entre un rétrotransposon et un gène ou bien une partie d'un gène, créant dans ce cas une conversion du gène impliqué dans la recombinaison (Figure 4D), soit entre deux rétrotransposons pouvant générer des délétions, des duplications et des inversions lorsqu'elle se fait de manière non allélique (Figure 4F).



**Figure 4 : Impact des ET sur la structure des génomes.** (A) Mutagenèse d'insertion : l'ET (en rouge) s'insère au niveau d'une séquence codante ou régulatrice ; (B) Cassure suivi d'une réparation double-brin, provoquée par des endonucléases (en vert) pouvant entraîner des différences par rapport à la séquence initiale ; (C) Duplications de microsatellites, ces courtes séquences répétées sont présentes aux extrémités ou au centre de certains éléments ; (D) Conversion de gène : cette conversion intervient lorsqu'il y a une recombinaison homologue entre le gène (en bleu) et un rétrotransposon (en rouge) ; (E) Délétion induite par une insertion : la délétion est provoquée par des mécanismes dépendants des endonucléases lors

de l'insertion de l'ET ; (F) Recombinaison ectopique, elle intervient entre deux rétrotransposons homologues non-alléliques ; (G) Transduction : les régions flanquantes en 5' et en 3' peuvent être rétrotranscrites, excisées puis insérées avec l'élément (Adaptée de Cordaux & Batzer, 2009).

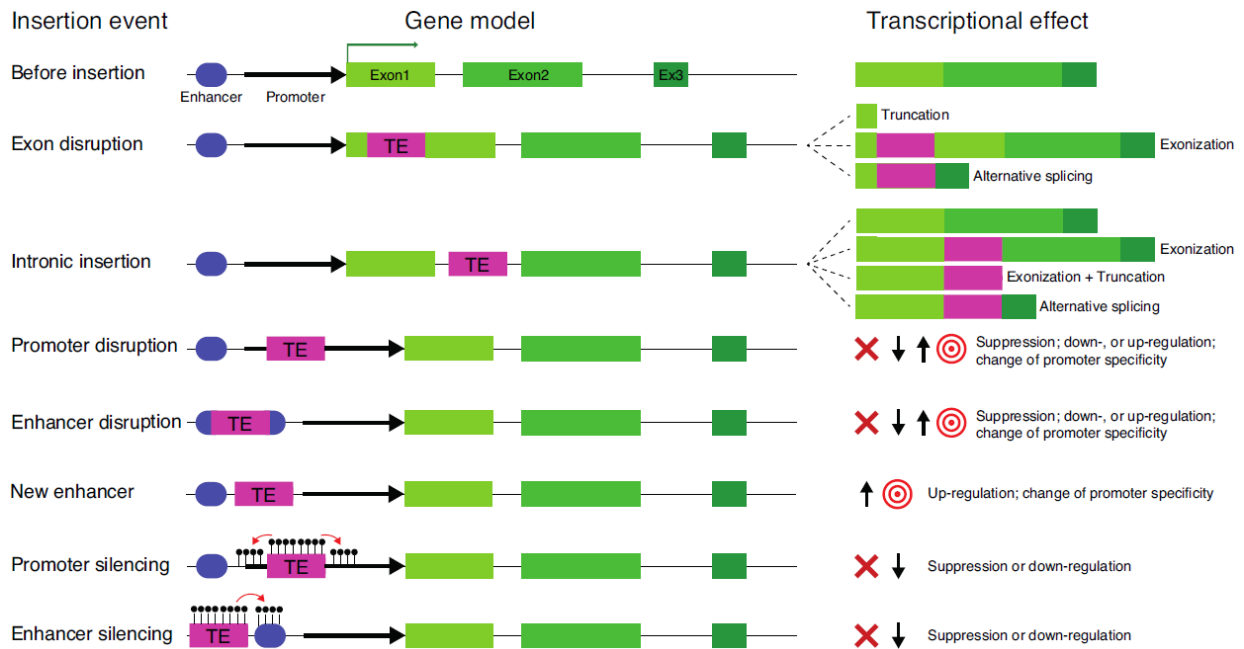
Les ET, au même titre que les gènes codants, possèdent des séquences régulatrices comme des sites d'initiation ou de terminaison de la transcription, des motifs reconnus par des facteurs de transcription, des signaux d'épissage ou encore des insulateurs (Rebollo *et al*, 2012b). Les éléments régulateurs pouvant varier d'un type d'ET à l'autre (Figure 5), les conséquences fonctionnelles au niveau des séquences adjacentes ne seront pas les mêmes d'un ET à l'autre.

	POL II	POL III	Enhancer	Transcription factor	Environment response elements	Insulators	Heterochromatin	Poly(A) sites	Termination site (POL III)	Splice sites
LTR	✓		✓	✓	✓	✓	✓	✓		✓
LINE	✓		✓	✓				✓	✓	✓
SINE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DNA	✓		✓	✓	✓			✓		✓

**Figure 5 : Séquences régulatrices des éléments transposables.** Eléments régulateurs dans différentes familles d'ET de classe I (LTR, LINE et SINE) et de classe II (DNA) (Adapté de (Rebollo *et al*, 2012b).

Il s'avère que ces conséquences fonctionnelles dépendent beaucoup du point d'insertion. Une insertion dans les séquences codantes ne produit pas les mêmes effets qu'une insertion dans les séquences régulatrices (Figure 6). Lorsque les ET transposent dans un exon ou dans un intron, il en résulte souvent des modifications au niveau de la structure des transcrits. Ceux-ci peuvent être tronqués et devenir non fonctionnels, notamment à cause de la présence de terminateurs chez certains ET. La création de nouveaux exons (exonisation) et la présence de sites d'épissage peuvent conduire à la formation de transcrits alternatifs. Lorsque les ET s'insèrent dans les séquences régulatrices, il en résulte une modification de l'expression du gène (répression ou surexpression). De plus, la présence de promoteurs dans les ET peut induire la transcription de séquences non codantes, en sens ou en antisens (Cordaux & Batzer, 2009; Dubin *et al*, 2018; Rebollo *et al*, 2012b). Chez la souris, il a été montré que les ET sont exprimés différemment au cours des stades du développement. L'insertion de ces ET dans la séquence codante de nombreux gènes ou dans les séquences régulatrices entraîne la formation de transcrits chimères pouvant être détectés par RT-PCR. Ces résultats suggèrent que les promoteurs des ET constituent des promoteurs alternatifs pour certains gènes (Peaston

*et al*, 2004). Dans des lignées cellulaires humaines, des miRNA provenant de la région *C19CM* sont transcrits grâce à l'ARN polymérase III à partir des promoteurs des copies d'*ET alu*, insérés en amont (Borchert *et al*, 2006).



**Figure 6 : Effets de l'insertion des ET sur l'expression des gènes.** Les conséquences fonctionnelles de la transposition dans une région codante dépendent du point d'insertion exact et de la structure du locus. L'insertion dans un exon entraîne la production de transcrits tronqués ou aberrants dans la plupart des cas ; si l'insertion se fait en phase avec la séquence codante elle peut conduire à une exonisation et à la formation de nouveaux transcrits par épissage alternatif. L'insertion dans un intron a des conséquences similaires à une insertion exonique. Au niveau des régions non transcrites, les insertions perturbent le fonctionnement des éléments régulateurs (promoteurs ou enhanceurs) ce qui aboutit à des modifications de la transcription. La présence de marques épigénétiques (méthylation de l'ADN) peut s'étendre sur les séquences adjacentes (promoteurs ou enhanceurs) et entraîner une répression de la transcription (Dubin *et al*, 2018).

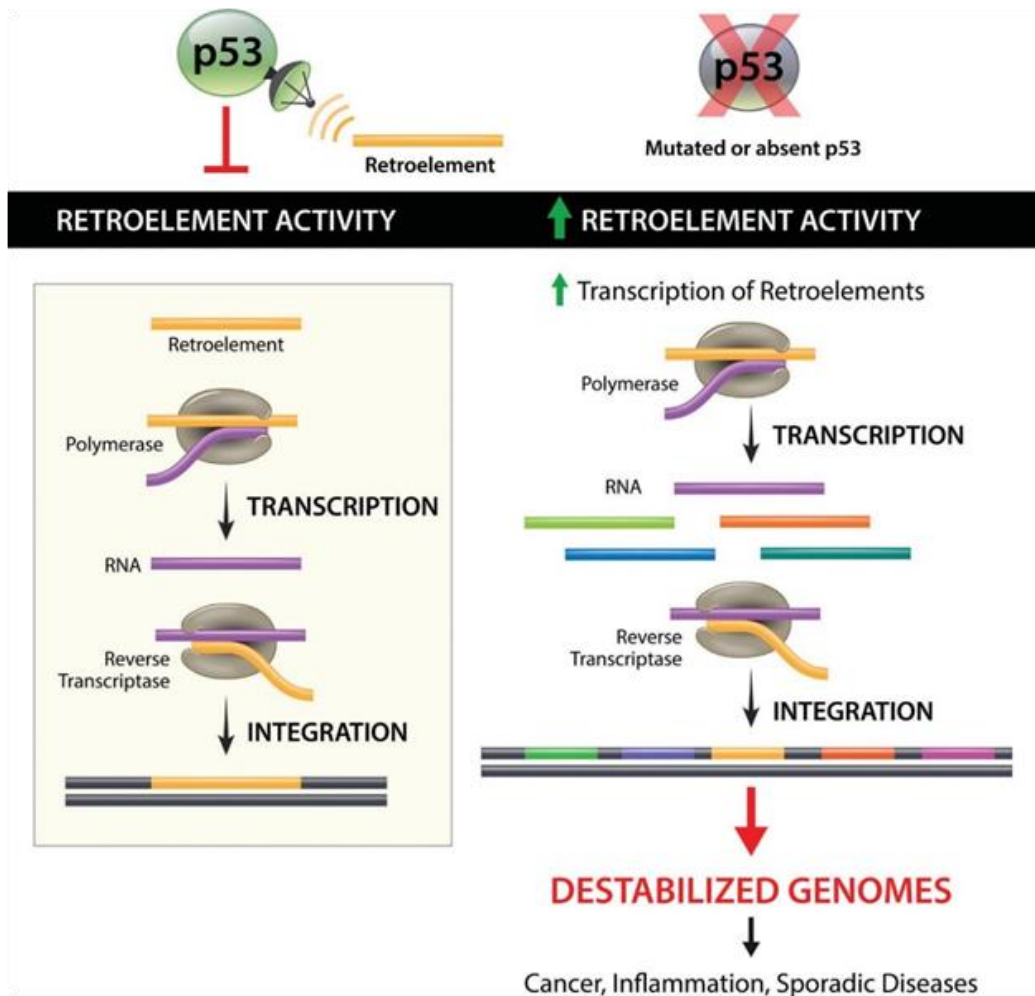
Le taux de mutation dû à des insertions d'ET n'est pas constant, il varie d'une espèce à l'autre (Biémont & Vieira, 2006) et le taux de mutation imputé spécifiquement aux rétrotransposons est plus important que celui des transposons notamment parce que la reverse transcriptase n'a pas d'activité de relecture (McDonald, 1993). De nombreux exemples de gènes fonctionnellement affectés par les éléments régulateurs des ET existent. Beaucoup sont répertoriés sur un catalogue en ligne appelé C-GATE pour Catalogue of Genes Affected by TE (Disponible sur : <https://ritarebollo.wixsite.com/cgate>, décrit dans (Rebollo *et al*, 2012a).

## 2.2. Les éléments transposables, un nid à problèmes !

L'activité des ET modifie structurellement et fonctionnellement les génomes. De plus, le caractère aléatoire de ces modifications suggère que la mobilité des ET est délétère pour l'intégrité du génome. En effet, que dire d'un élément génétique qui peut bloquer l'expression de gènes essentiels ou générer des cassures d'ADN de manière anarchique, si ce n'est qu'il représente un danger ? Et pour cause, l'activité des ET a été identifiée comme étant responsable de beaucoup de maladies dont voici divers exemples. Le lien entre ET et dystrophies musculaires a été établi où des insertions d'ET dans des gènes impliqués dans le développement musculaire, ont été détectées, soit au niveau des séquences codantes (Narita *et al*, 1993), soit au niveau des régions régulatrices (Kondo-Iida *et al*, 1999). Des défauts d'épissage liés à des insertions d'ET dans des introns sont observés dans différents syndromes, par exemple pour le syndrome d'Alport qui entraîne une insuffisance rénale (Knebelmann *et al*, 1995) ou encore le syndrome de lymphoprolifération autoimmune (Tighe *et al*, 2002). L'hémophilie aussi est associée à des insertions d'ET (Ganguly *et al*, 2003), mais dans le cas de cette maladie, différents types de mutation ont été reportés (Mukherjee *et al*, 2004). L'activité des ET semblent aussi impliquée dans des cas de maladies psychiatriques – dépression, bipolarité ou encore schizophrénie – pour lesquelles des surexpressions d'ET ou l'expression aberrante de certains gènes sous contrôles des ET ont été identifiées (Guffanti *et al*, 2014). La protéine Tau, dont l'activité a été caractérisée dans le développement de la maladie d'Alzheimer, semble à l'origine de l'activation de l'expression de différentes classes d'ET, conduisant à une instabilité génomique (Guo *et al*, 2018). Enfin, l'insertion et l'activité des ET ont aussi été identifiées dans de nombreux cas de cancers. L'analyse de tumeurs de patients atteints de cancer du colon montre que l'insertion d'un LINE dans le gène suppresseur de tumeur *APC* conduit à son inactivation (Miki *et al*, 1992). L'analyse de données de RNA-seq issues de différents types de tumeurs a permis d'identifier des transcrits chimères constitués d'ET et d'oncogènes, suggérant l'insertion d'ET dans les séquences régulatrices de nombreux oncogènes. Ces insertions sont globalement associées à la surexpression des oncogènes à proximité desquels les ET sont insérés. Par exemple, l'insertion de l'ET *AluJb* induit l'expression du gène *LIN28B*, dont le rôle a été identifié dans des cancers pulmonaires tandis que la méthylation d'*AluJb* conduit à la diminution de l'expression de *LIN28B* (Jang *et al*, 2019). Enfin, des travaux ont montré chez différents modèles (Drosophile, nématodes, poisson-zèbre et lignées cellulaires humaines) que lorsqu'on mute le gène *p53* l'activité des ET est plus importante (Wylie *et al*, 2016b). *p53* est un gène



très étudié en cancérologie, sa perte de fonction est connue pour être impliquée dans certains cancers. Les auteurs proposent un rôle de *p53* dans le contrôle des ET. Ainsi, les cas de cancers associés à une mutation de *p53* seraient provoqués par la dérégulation des ET, entraînant une instabilité génomique (Figure 7). De manière générale, le large spectre des maladies causées par l'activité des ET, et dont certains exemples ont été donnés dans le paragraphe précédent, pourraient être regroupés dans la catégorie des 'transposopathies' (Wylie *et al*, 2016a).



**Figure 7 : Rôle de *p53* dans le contrôle des rétrotransposons.** Selon ce modèle, le gène *p53* est capable de réguler l'activité des rétrotransposons et ainsi permet de limiter l'instabilité génomique. Lorsque *p53* est muté l'activité des rétrotransposons augmente ce qui génère une instabilité génomique, entraînant différents types de maladies qualifiées de 'transposopathies' (Wylie *et al*, 2016a).

Etant donné leurs effets délétères il peut sembler étonnant que les ET n'aient pas été contre-sélectionnés au cours de l'évolution, pourtant les analyses génomiques révèlent qu'il n'existe pas de génome dépourvu de séquences dérivées d'ET (Pritham, 2009). Ainsi, les ET semblent capables de se maintenir efficacement dans les génomes. Il apparaît d'ailleurs que les ET participe à l'évolution des génomes, notamment parce qu'ils y apportent de la diversité génétique qui peut conduire à l'apparition de nouvelles fonctions. Leur présence pourrait même constituer un avantage sélectif dans des conditions particulières.

### **2.3. Les éléments transposables : un rôle dans l'évolution ?**

Les ET constituent pour les génomes une source de séquences à exploiter ; la variété des modifications structurales et fonctionnelles qu'ils peuvent engendrer semble indiquer qu'ils sont des générateurs de diversité génétique.

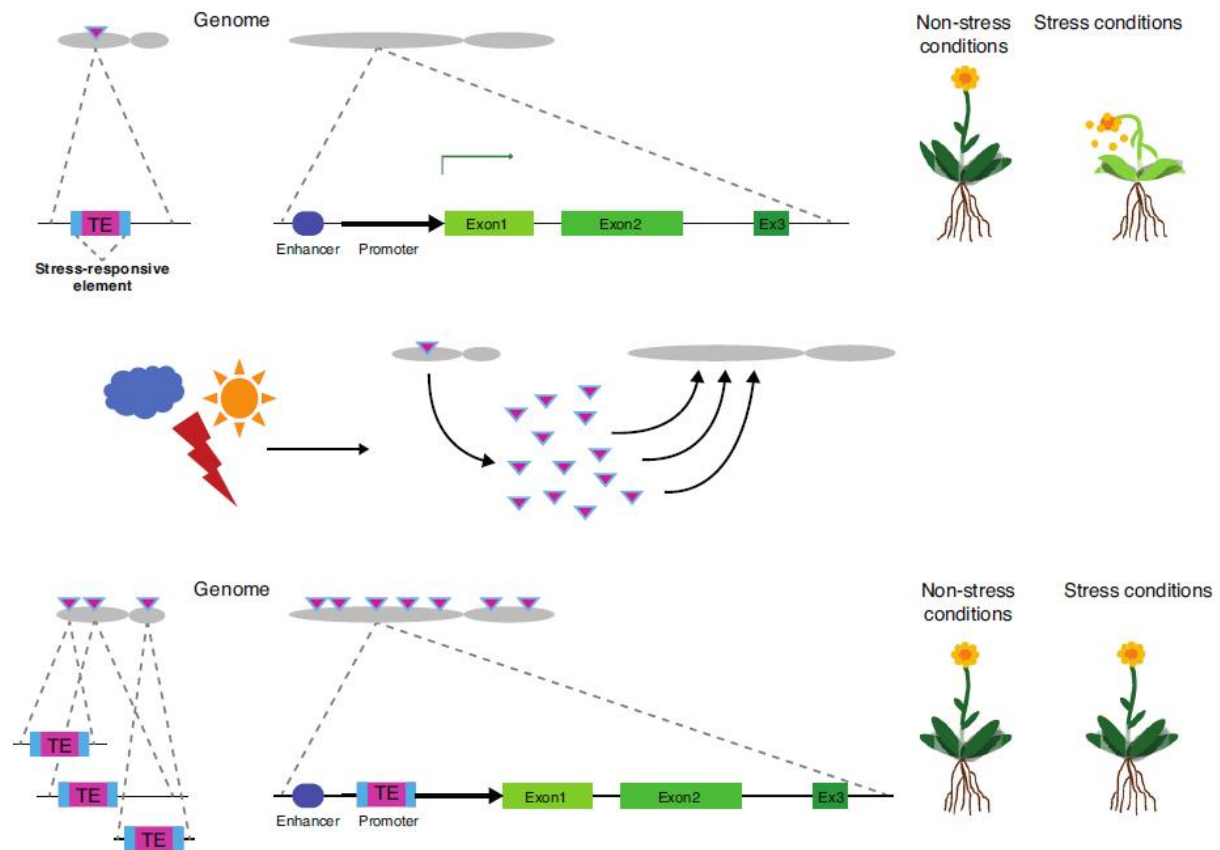
Il existe de nombreux exemples de domestication des ET illustrant qu'ils sont responsables d'innovations génétiques. Ceux-ci ont contribué à l'expansion des certains sites reconnus par des facteurs de transcription (Bourque *et al*, 2008; Schmidt *et al*, 2012) mais également à l'apparition de nouvelles protéines, notamment grâce à la mise en place de cadres ouverts de lecture (Britten, 2006; Kitano *et al*, 2018). Ces innovations modifient les réseaux de gènes et amènent de nouvelles fonctions. De ce fait, les ET sont impliqués dans des processus importants tels que l'immunité innée (Chuong *et al*, 2016; Zhang *et al*, 2019), la reprogrammation cellulaire (Wang *et al*, 2014a) ou les réarrangements génomiques programmés chez les ciliés (Baudry *et al*, 2009; Cheng *et al*, 2010).

L'idée d'envisager que les ET jouent un rôle majeur dans l'évolution n'est pas nouvelle, cependant de plus en plus de résultats renforcent cette hypothèse. Des travaux ont montré que la fréquence des mutations spontanées, seule, ne permet pas d'expliquer l'évolution anatomique rapide et l'augmentation du nombre de chromosomes chez les mammifères (Wilson *et al*, 1974). Les ET sont une source majeure de mutation et sont à l'origine de phénomènes de réarrangements génétiques, suggérant qu'ils pourraient être de bon candidats moléculaires pour justifier la vitesse de l'évolution (McDonald, 1993). Ainsi, des changements dans la régulation des gènes – dus à la présence de nombreuses séquences régulatrices dans les ET (Voir figure 4) – auraient plus d'impact sur l'évolution des génomes qu'une accumulation de mutations spontanées au cours des générations (McDonald, 1995).

De cette manière, l'apparition puis la multiplication rapide des ET au sein d'un génome génèrent des réarrangements génomiques assez importantes pour être à l'origine de la création de nouvelles espèces, voire de nouveaux phylum (Belyayev, 2014; Rebollo *et al*, 2010). De ces processus de spéciation résulteraient l'apparition de phénomènes d'isolements reproductifs associés à trop de divergences génétiques (Serrato-Capuchina & Matute, 2018). Des hypothèses suggèrent même que les ET sont en cause dans l'évolution des eucaryotes : leur abondance et leur activité expliqueraient la taille et la complexité de leurs génomes, par rapport aux procaryotes (Fedoroff, 2012).

En somme, la capacité des ET à induire des mutations, permettant de créer de nouveaux produits (transcrits, protéines...) et ainsi de nouvelles fonctions, peut représenter un atout dans des conditions de stress. Les ET possèdent des séquences de réponse au stress (« Environment Response Element » figure 5) et nombreux sont les stress, biotiques (Infections virales ou bactériennes) ou abiotiques (Température, radiation, polluants...), capables d'augmenter la mobilisation des ET (Esnault *et al*, 2019; García Guerreiro, 2012; Kofler *et al*, 2018; Miousse *et al*, 2015). Des modifications durables de l'environnement contraignent les populations à s'adapter rapidement. Or, l'augmentation du taux de transposition conduit à un accroissement de la variabilité génétique, menant à une probabilité de survie plus importante et à une accélération du processus adaptatif (Piacentini *et al*, 2014). Face à un stress les ET peuvent modifier l'expression des gènes à proximité desquels ils sont insérés. C'est le cas notamment de l'insertion du transposon *bari* près des gènes *Jheh1*, *Jheh2* et *Jheh3* chez *Drosophila melanogaster*. En l'absence de stress, ces gènes sont réprimés et on détecte un enrichissement de la marque d'histone H3K9me3 au niveau de la région. Mais en condition de stress oxydatif, on observe un enrichissement en marques activatrices H3K4me3 et en marques répressives H3K27me3 associé à la modification de l'expression des trois gènes. En l'absence de *bari*, l'ensemble des modifications chromatinienne de la région ne sont pas observées, illustrant le rôle de l'ET *bari* dans la régulation des gènes *Jheh* (Guio *et al*, 2018). Chez la levure, il a également été montré qu'une exposition au chlorure de cobalt (CoCl<sub>2</sub>) entraîne une augmentation de la transcription et de la transposition des ET. Il en résulte que certaines insertions entraînent l'émergence de colonies résistantes au stress. Parmi ces individus, on observe un enrichissement d'insertions d'ET dans les promoteurs de gènes de réponse au stress (Esnault *et al*, 2019). Tous ces résultats indiquent que les ET constituent une réserve de diversité génétique nécessaire pour palier à des stress éventuels. Ainsi les génomes contenant

des ET sont plus susceptibles de répondre aux contraintes environnementales que les génomes sans ET : leur présence apparaît donc comme un avantage sélectif (Figure 8).



**Figure 8: Rôle des ET dans la réponse au stress environnemental.** Les ET contiennent des éléments de réponse au stress et l'insertion d'un ET dans un gène peut modifier son expression (Voir figure 6). Des conditions de stress induisent une augmentation du taux de transposition conduisant à de nombreuses insertions au sein du génome. Si l'insertion se fait dans ou à proximité d'un gène, que l'ET – avec ses éléments de réponse au stress – est maintenu à travers les générations et que le produit du gène représente un avantage pour la tolérance au stress, alors l'organisme peut devenir résistant au stress qui est à l'origine de ces changements (Dubin *et al*, 2018).

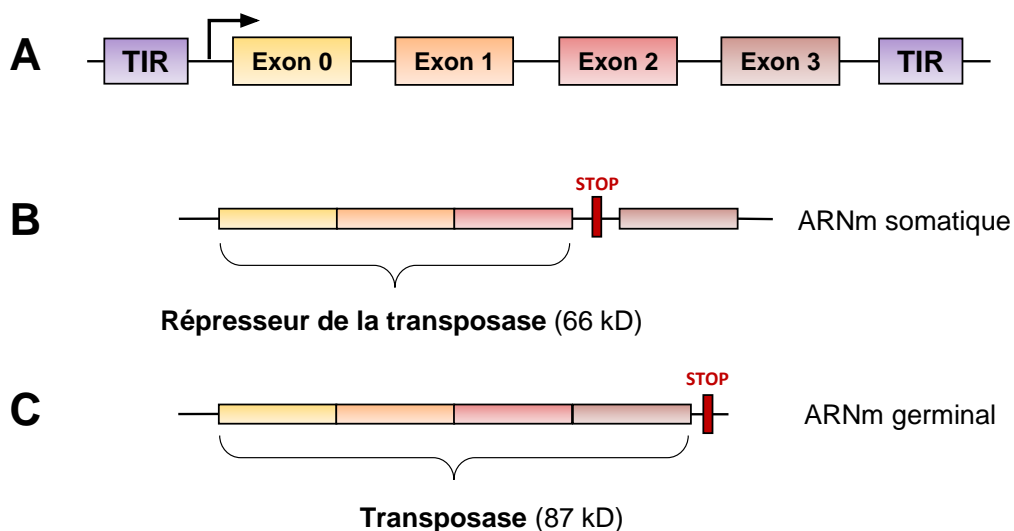
Bien que certaines mutations dues aux ET confèrent un avantage évolutif, il apparaît qu'une partie d'entre elles ont des effets néfastes – comme nous avons pu en discuter dans la partie précédente. Seulement, les mutations délétères ne perdurent pas au cours des générations car elles tendent à être contre-sélectionnées. A l'inverse les mutations qui apportent un avantage sont susceptibles d'être fixées dans les populations (McDonald, 1995). Ainsi, nous pourrions estimer que les effets délétères des ET sont des effets à court terme alors que les effets bénéfiques ne se mesurent que sur le long terme.

### 3. Le cas de l'élément *P*

Les éléments transposables sont ubiquitaires, diversifiés et peuvent avoir des effets délétères autant que bénéfiques. L'élément *P* est un élément transposable particulièrement étudié chez *Drosophila melanogaster*. Son étude est à l'origine de nombreuses découvertes dans les domaines de la génétique et de la biologie moléculaire. Certaines de ces découvertes ont été faites dans notre laboratoire et il nous sert encore actuellement, notamment pour des constructions génétiques, à répondre à de nouvelles questions biologiques.

#### 3.1. Présentation de l'élément *P*

L'élément *P* est un élément transposable de type II d'une taille de 2.9 kb. Il possède des séquences inversées répétées (TIR : Terminal Inverted Repeat) au niveau de ses extrémités et code pour le gène de la transposase qui est constitué de quatre exons et trois introns (Figure 9A) (O'Hare & Rubin, 1983). Dans les cellules somatiques la transposase n'est pas produite, dû à un épissage alternatif (Laski *et al*, 1986). En effet, l'intron situé entre les exons 2 et 3 n'est pas excisé et contient un codon stop. L'ARNm alors obtenu ne permet pas la synthèse d'une protéine fonctionnelle, ainsi l'élément *P* ne peut pas se déplacer (Figure 9B). A l'inverse, dans les cellules germinales, la transposase est synthétisée et fonctionnelle, ce qui permet la mobilité de l'élément *P* (Figure 9C).



**Figure 9** : Schéma représentant la structure de l'élément *P*. (A) Des séquences inversées répétées aux extrémités (TIR) flanquent le gène de la transposase. Ce gène est composé de

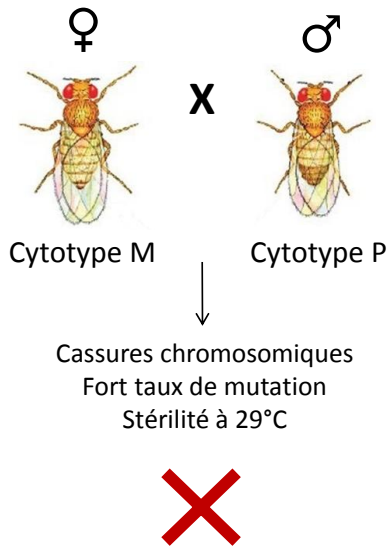
quatre exons et trois introns. (B) Dans les cellules somatiques, la transposase n'est pas synthétisée du fait d'un épissage alternatif. (C) Dans les cellules germinales, la transposase est synthétisée et fonctionnelle.

### 3.2. La dysgénésie des hybrides

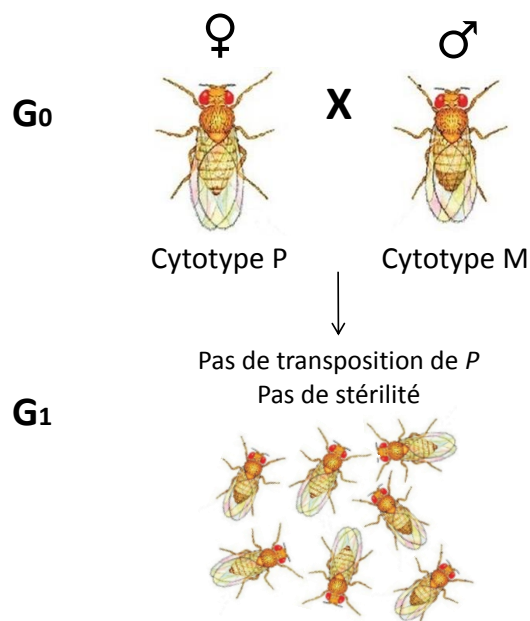
Au cours du siècle dernier, l'élément *P* est apparu au sein du génome de *Drosophila melanogaster* par un transfert horizontal provenant de *Drosophila willistoni* (Daniels *et al*, 1990). En moins de 20 ans, l'élément *P* a rapidement envahi le génome de toutes les populations de *Drosophila melanogaster*. Cette invasion a été découverte à la suite de l'observation d'un phénomène génétique étonnant. En effet, des femelles de vieilles souches de laboratoire, collectées avant l'invasion de l'élément *P* (et dont le génome est vide de l'élément *P*) croisées par des mâles collectés plus récemment (dont le génome contient des copies de l'élément *P*), produisent une descendance qui présente des anomalies génétiques. Cependant, ce phénomène n'est pas observé dans la descendance du croisement réciproque (Femelles avec *P* x Mâles sans *P*). Ce syndrome d'anomalies a été appelé « dysgénésie des hybrides » (Kidwell *et al*, 1977). Les femelles possédant des éléments *P* possèdent un état cellulaire transmis maternellement appelé cytotype « P » qui empêche la dysgénésie tandis que les femelles dépourvues d'éléments *P* possèdent un cytotype « M » qui n'empêche pas la dysgénésie (Figure 10).

Ces résultats révèlent que dans le cas du croisement dysgénésique la transposition de l'élément *P* n'est pas contrôlée, ce qui génère des cassures chromosomiques et un taux de mutation important qui affectent le développement des gonades, mais aussi un phénotype de stérilité à 29°C (Engels & Preston, 1979). Dans le cas du croisement réciproque la transposition de l'élément *P* est régulée dans la lignée germinale et le génome est protégé, ce qui permet le développement d'une descendance fertile.

### Croisement dysgénésique



### Croisement réciproque



**Figure 10: Phénomène de dysgénésie des hybrides.** Les individus de cytotype « P » possèdent des copies de l'élément *P* alors que les individus de cytotype « M » ne possèdent pas de copies de l'élément *P*. Le croisement dysgénésique (Femelles sans *P* x Mâles avec *P*) génère des anomalies génétiques alors que le croisement réciproque (Femelles avec *P* x Mâles sans *P*) n'engendre pas ces anomalies.

### 3.3. Régulation de l'élément *P*

Ainsi, il existe une régulation de l'élément *P* chez la femelle puisque seul le croisement non-dysgénésique témoigne de l'absence d'anomalies génétiques. Il a été montré que cette régulation est sensible à la température. En effet, on observe que les capacités régulatrices augmentent dans la descendance d'un croisement dysgénésique lorsque la température est plus élevée (26.5°C) : le cytotype « M » devient un cytotype « P ». En revanche, à une température plus basse (18°C), on n'observe pas l'émergence de capacités répressives (Ronsseray *et al*, 1984). Des travaux plus récents ont mis en évidence que la vitesse d'invasion du génome de *Drosophila simulans* par l'élément *P* dépend de la température. L'invasion est relativement rapide (moins de 20 générations) à des températures élevées alors qu'elle est beaucoup plus lente (plus de 40 générations) à des températures plus basses. Par ailleurs, cette invasion du génome s'accompagne de l'apparition progressive d'un système de régulation de la transposition (Kofler *et al*, 2018).

On trouve en moyenne entre 30 et 50 copies d'élément *P* par génome chez *Drosophila melanogaster* (Bingham *et al*, 1982). Cependant, l'analyse de la répartition des copies de l'élément *P* montre qu'il existe des points chauds d'insertion (Ronsseray *et al*, 1989). On a identifié des copies de l'élément *P* insérées dans les séquences subtélomériques, appelées *Telomeric Associated Sequence (TAS)*, qui permettent de réprimer toutes les autres copies d'élément *P* du génome (Ronsseray *et al*, 1996). En d'autres termes, il existe des régions particulières au sein du génome agissant comme des « pièges à séquences mobiles ». Lorsqu'un ET s'insère dans l'une de ces régions, toutes les copies génomiques de cet élément sont réprimées. Ces séquences constituent ainsi un catalogue des séquences d'ADN mobiles à réguler. Selon cette propriété, il a été mis en évidence dans la lignée germinale qu'un transgène *P-lacZ* - dérivé de l'élément *P* dans lequel le gène codant la transposase est remplacé par le gène rapporteur *lacZ* - inséré au niveau d'un ces sites régulateurs des régions subtélomériques, est capable de réprimer en *trans* une autre copie de transgène *P-lacZ* dans le génome. Ce phénomène de *Trans-Silencing* implique une régulation de l'élément *P* par homologie de séquences (Roche & Rio, 1998).

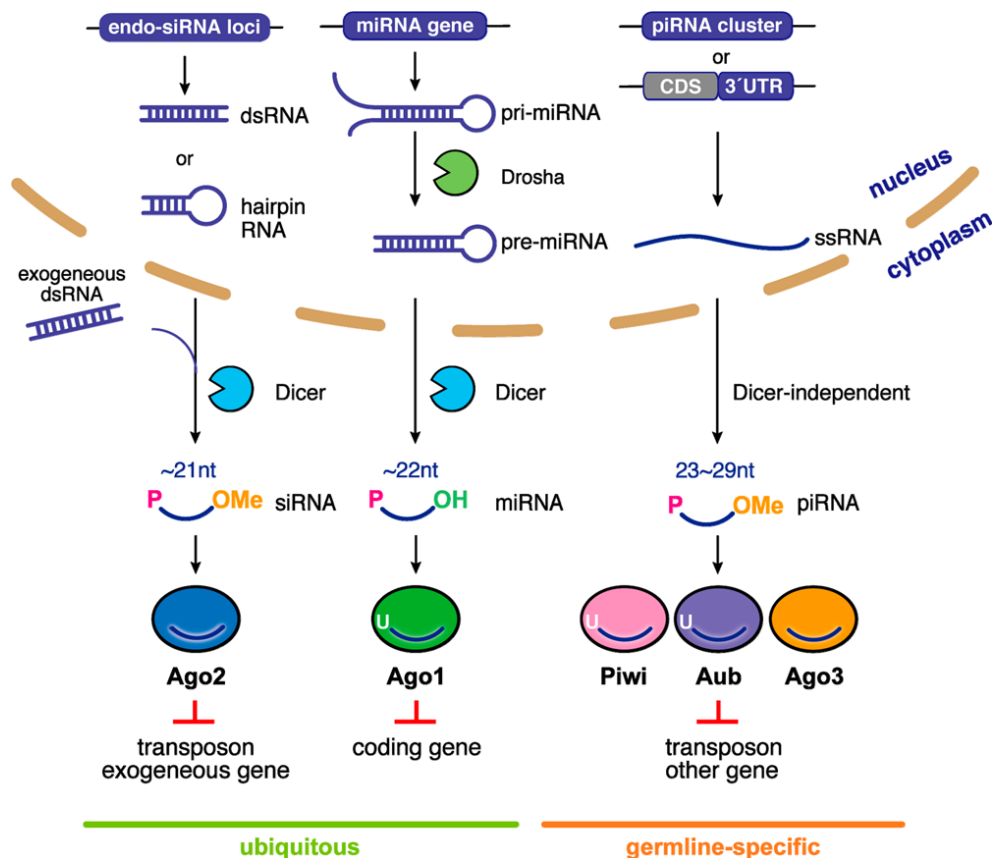
Ce n'est que bien plus tard, grâce au développement des techniques de séquençage, que les bases moléculaires de la régulation de l'élément *P*, et plus généralement la régulation des ET, ont été découvertes. Cette régulation se fait par l'intermédiaire de différentes familles de petits ARN non-codants. Dans les tissus somatiques, la régulation des ET s'effectue grâce à la famille des Endogenous Small Interfering RNA (endo-siRNA) (Czech *et al*, 2008; Ghildiyal *et al*, 2008) tandis que dans les cellules germinales, elle dépend de l'activité des PIWI-Interacting RNA (piRNA) qui sont produits par des séquences spécifiques appelées piRNA clusters (Brennecke *et al*, 2007; Gunawardane *et al*, 2007).





## II. Les PIWI-Interacting RNA

En 2006, une nouvelle famille de petits ARN non-codants a été identifiée. Initialement baptisés Repeat-Associated small interfering RNA (rasiRNA) - avant d'être définitivement appelés PIWI-Interacting RNA (piRNA) - ces petits ARN sont principalement impliqués dans la régulation des ET dans les cellules germinales. Cette voie est distincte des voies d'interférence à l'ARN alors connues : la voie des microRNA (miRNA) et la voie des Short Interfering RNA (siRNA). En effet, alors que les siRNA et les miRNA sont chargés par des protéines Argonautes de la famille AGO, les piRNA interagissent avec des protéines Argonautes de la famille PIWI (Brennecke *et al*, 2007; Gunawardane *et al*, 2007). De plus, leur biogénèse ne fait pas intervenir les protéines Dicer, impliquées dans le clivage d'ARN double-brin (Saito *et al*, 2006; Vagin *et al*, 2006) (Figure 11).



**Figure 11 :** Principales familles de petits ARN non-codants chez *Drosophila melanogaster*. Les siRNA (à gauche) sont produits à partir d'ARN précurseurs double-brin clivés par Dicer-2 pour former des petits ARN de 21 nucléotides. Ces ARN sont pris en charge par Ago2 pour former un complexe RISC. Les ARN précurseurs peuvent être

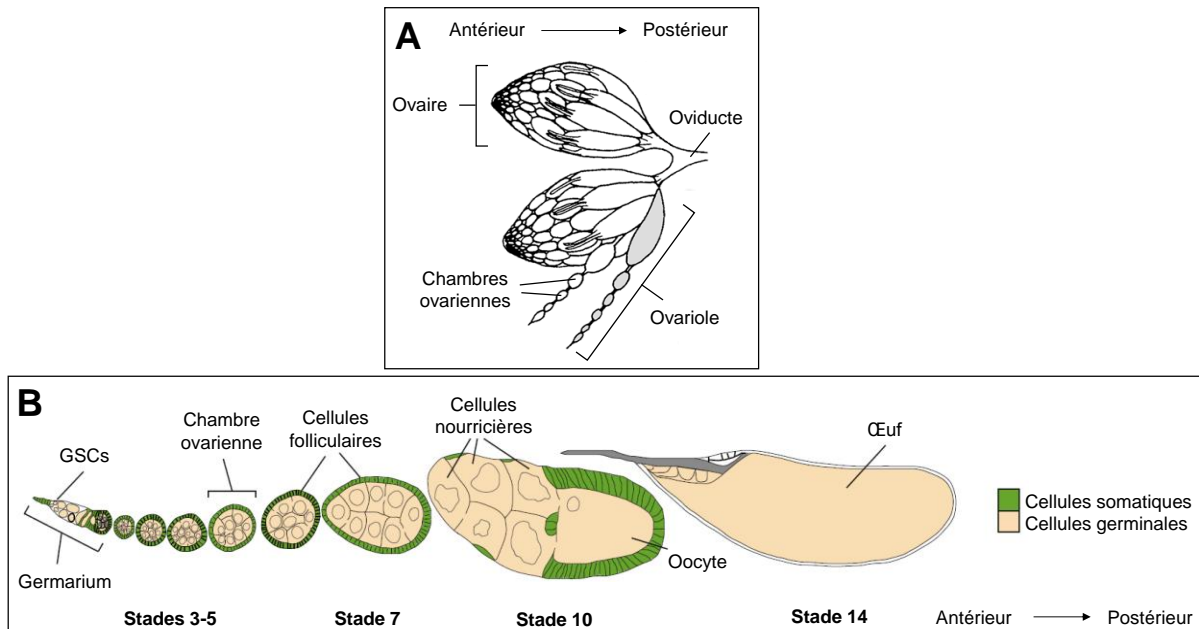
d'origine exogène ou endogène – ces derniers se forment grâce à des locus particuliers et permettent de produire des endo-siRNA. Les siRNA ciblent principalement des séquences exogènes et reconnaissent les ET dans les cellules somatiques. Les miRNA sont produits à partir d'ARN précurseurs double-brin formant une structure tige-boucle et transcrits à partir de gènes non-codants (« miRNA gene »). Ces ARN précurseurs sont clivés successivement par Droscha puis Dicer-1 pour générer des petits ARN de 22 nucléotides. Ces petits ARN interagissent alors avec Ago1 pour former un complexe mi-RISC qui cible les ARN messagers des gènes codants. Les piRNA (à droite) sont produits à partir de précurseurs ARN simple-brin principalement transcrits par des séquences appelées piRNA clusters ou, dans une minorité de cas, par les séquences 3'UTR de certains gènes codant des protéines. Ces ARN précurseurs sont clivés et génèrent des petits ARN de 23 à 29 nucléotides qui sont ensuite chargés par des protéines PIWI (Piwi, Aub et Ago3) pour former des complexes piRISC. Ces complexes ciblent essentiellement des ET mais aussi certains gènes codants. Les siRNA et les miRNA sont présents dans les différents types cellulaires alors que les piRNA sont spécifiques de la lignée germinale. Les miRNA portent un groupement OH en 3' tandis que les siRNA et les piRNA portent un groupement 2'O-méthyl (Yamashiro & Siomi, 2018).

## **1. Régulation des ET par les piRNA**

### **1.1. La lignée germinale de la drosophile**

La drosophile femelle possède une paire d'ovaires contenant entre 16 et 18 ovarioles chacun, eux-mêmes composés de plusieurs chambres ovariennes (Figure 12A). Dans un même ovariole, on trouve tous les stades de développement (14 stades) d'une chambre ovarienne, de la cellule souche aux stades les plus tardifs (14<sup>ème</sup> stade). Le germarium est une structure située au pôle antérieur de l'ovariole et qui contient deux types de cellules souches : deux ou trois cellules souches germinales (aussi appelées GSCs pour germline stem cells) et des cellules souches somatiques. Les GSCs se divisent asymétriquement pour donner une autre GSC – et ainsi permettre le maintien de ce type de cellule souche – et une cellule fille appelée cystoblaste qui pourra se différencier. Cette cellule fille subit quatre mitoses successives avec une cytokinèse incomplète entraînant la formation d'un cyste contenant 16 cellules germinales connectées. Parmi ces cellules, une seule devient le futur oocyte tandis que les 15 autres se différencient en cellules nourricières (Figure 12B, cellules de couleur beige). Ces cellules germinales sont entourées par une couche de cellules somatiques – les cellules folliculaires (Figure 12B, cellules de couleur verte) –, le tout formant une chambre ovarienne (Bastock & St Johnston, 2008; He *et al.*, 2011). L'ovogénèse complète dure approximativement une semaine. Les différents stades de développement des chambres ovariennes sont alignés selon

l'axe antéro-postérieur : les stades précoces se trouvent au pôle antérieur et les stades tardifs au pôle postérieur (Figure 12B). Ainsi, cet alignement s'apparente d'une certaine manière à un axe temporel.



**Figure 12: Lignée germinale de la femelle drosophile.** (A) La femelle possède une paire d'ovaires. Chaque ovaire est constitué d'ovarioles, eux-mêmes constitués de chambres ovariennes. (B) Dans un ovariole on trouve des chambres ovariennes à tous les stades de développement, des GSCs au pôle antérieur (dans le germarium) jusqu'à l'œuf au pôle postérieur. Chaque chambre ovarienne est constituée de cellules germinales (en beige) et de cellules somatiques (en vert). (Adapté de Handler *et al*, 2013).

## 1.2. Les séquences productrices de piRNA

Les piRNA sont produits majoritairement par des séquences particulières appelées 'piRNA clusters', qui constituent un catalogue des séquences mobiles à réguler. Ces séquences, enrichies en marques H3K9me3, sont composées de fragments d'ET et leur taille varie de quelques kilobases jusqu'à plusieurs centaines de kilobases (Aravin *et al*, 2006; Brennecke *et al*, 2007; Girard *et al*, 2006). Les piRNA clusters sont essentiellement localisés dans des régions hétérochromatiques ou à la frontière entre l'euchromatine et l'hétérochromatine. Ainsi, on les trouve au niveau des centromères et régions péri-centromériques, ainsi qu'aux niveaux des télomères et des subtélomères (Brennecke *et al*, 2007). Chez la drosophile, 142 piRNA clusters ont été identifiés et produisent plus de 80% des piRNA totaux (Brennecke *et al*, 2007); de façon similaire, chez la souris on compte une centaine de piRNA clusters produisant plus de 90% des piRNA totaux (Aravin *et al*, 2006; Girard *et al*, 2006). Les

analyses bioinformatiques montrent que la plupart des piRNA sont homologues à plusieurs régions génomiques, indiquant une redondance des séquences présentes dans les piRNA clusters (Brennecke *et al*, 2007).

Chez la drosophile, les piRNA clusters sont divisés en deux grandes catégories : les piRNA clusters de type uni-strand et les piRNA clusters de type dual-strand. Tous les piRNA clusters – qu'ils soient de type uni-strand ou dual-strand - sont transcrits par l'ARN polymérase II pour former de longs ARN précurseurs destinés à la production des piRNA dans le cytoplasme (Goriaux *et al*, 2014; Mohn *et al*, 2014; Zhang *et al*, 2014). Cependant, ces deux types de piRNA clusters se distinguent par leurs modalités de transcription ainsi que par la structure des précurseurs produits.

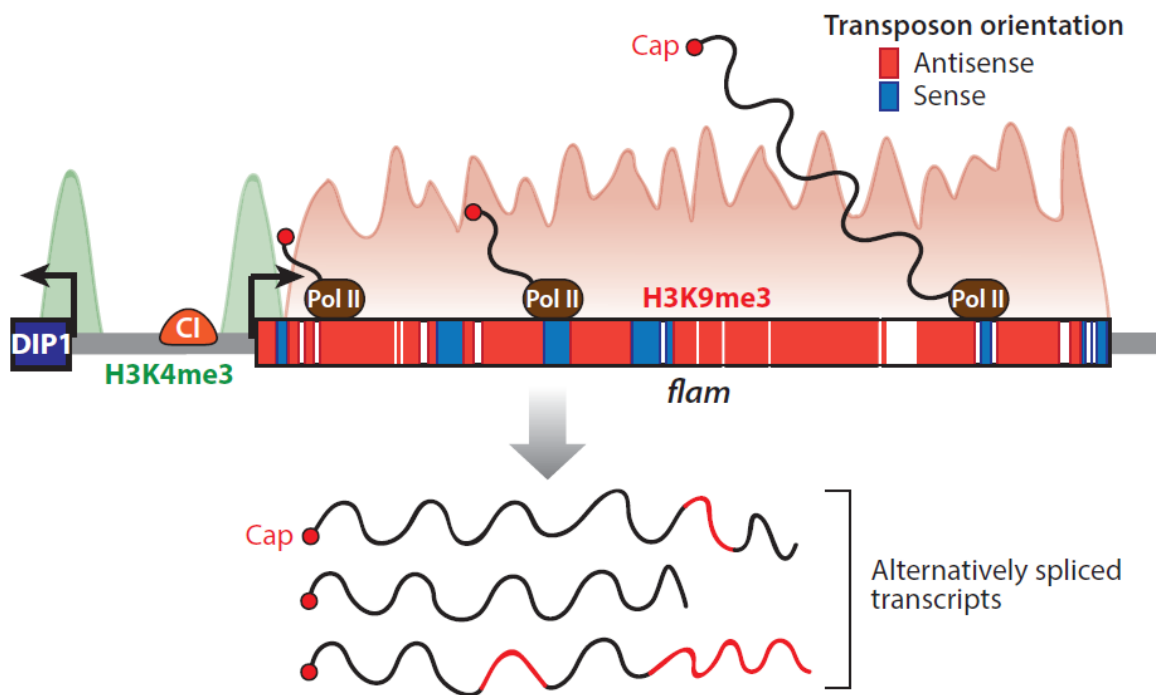
#### - Les piRNA clusters « uni-strand »

Dans les cellules folliculaires ovariennes – qui entourent les cellules germinales (voir Figure 12B) – les piRNA sont produits par des piRNA clusters de type uni-strand. Ceux-ci se caractérisent notamment par une transcription des ARN précurseurs à partir d'un seul des deux brins (Brennecke *et al*, 2007). L'étude des deux principaux piRNA clusters uni-strand – le cluster *20A* et *flamenco* – a contribué à la mise en évidence de la biologie de ce type de piRNA cluster.

Le locus *flamenco* a été identifié comme le régulateur majeur des ET *gypsy*, *Idefix* et *ZAM* dans les cellules folliculaires, dans lesquelles il est actif spécifiquement (Desset *et al*, 2003; Pélisson *et al*, 1994). *Flamenco* est localisé dans la région péricentromérique du chromosome X - à la frontière entre l'euchromatine et l'hétérochromatine (site cytologique 20A1-3) - à proximité du promoteur du gène *DIP1*. Ce cluster mesure environ 180 kilobases et est essentiellement constitué de fragment d'ET orientés en antisens (Goriaux *et al*, 2014). L'analyse des séquences consensus d'initiation de la transcription (Motif Inr) a permis d'identifier le promoteur de *flamenco* en amont du gène *DIP1*. En l'absence de cette séquence consensus, la transcription est abolie. *Flamenco* est transcrit par l'ARN polymérase II et nécessite la présence du facteur de transcription *Cubitus interruptus* (Ci), qui n'est pourtant pas spécifique des piRNA clusters. Les transcrits sont épissés alternativement, générant une dizaine d'ARN précurseurs différents mais qui possèdent tous le premier exon (Goriaux *et al*, 2014).

Le cluster *20A* est actif à la fois dans les cellules folliculaires et les cellules germinales. De même que *flamenco*, il est localisé dans la région péri-centromérique du chromosome X. Les précurseurs sont transcrits à partir d'un promoteur sur un seul brin par l'ARN polymérase II ; ils portent une coiffe 5' méthyl-guanosine (ou cap-5') ainsi qu'une queue poly(A) en 3'. Les analyses moléculaires révèlent également que ce promoteur est enrichi par les marques activatrices H3K4me2/3 tandis que le corps du locus est enrichi en marques H3K9me3 (Li *et al*, 2013; Mohn *et al*, 2014).

Ainsi, la transcription des piRNA clusters uni-strand est similaire à celle des gènes codants : elle se fait *via* l'ARN polymérase II, nécessite la présence d'un promoteur enrichi en H3K4me2/3 et génère des ARN précurseurs épissés, portant une coiffe et polyadénylés. Cependant, à la différence des gènes codants, le corps des piRNA cluster uni-strand est recouvert par la marque H3K9me3 (Figure 13).



**Figure 13 : Transcription d'un piRNA cluster uni-strand – exemple du locus *flamenco*.** Les fragments d'ET qui constituent *flamenco* sont majoritairement orientés en antisens. Le promoteur de *flamenco* est enrichi en H3K4me3 et sa transcription est sous le contrôle du facteur de transcription Cubitus interruptus (Ci). L'ARN polymérase II (Pol II) réalise la transcription du locus, permettant la production de multiples transcrits précurseurs épissés, coiffés et polyadénylés (Czech *et al*, 2018).

## - Les piRNA clusters « dual-strand »

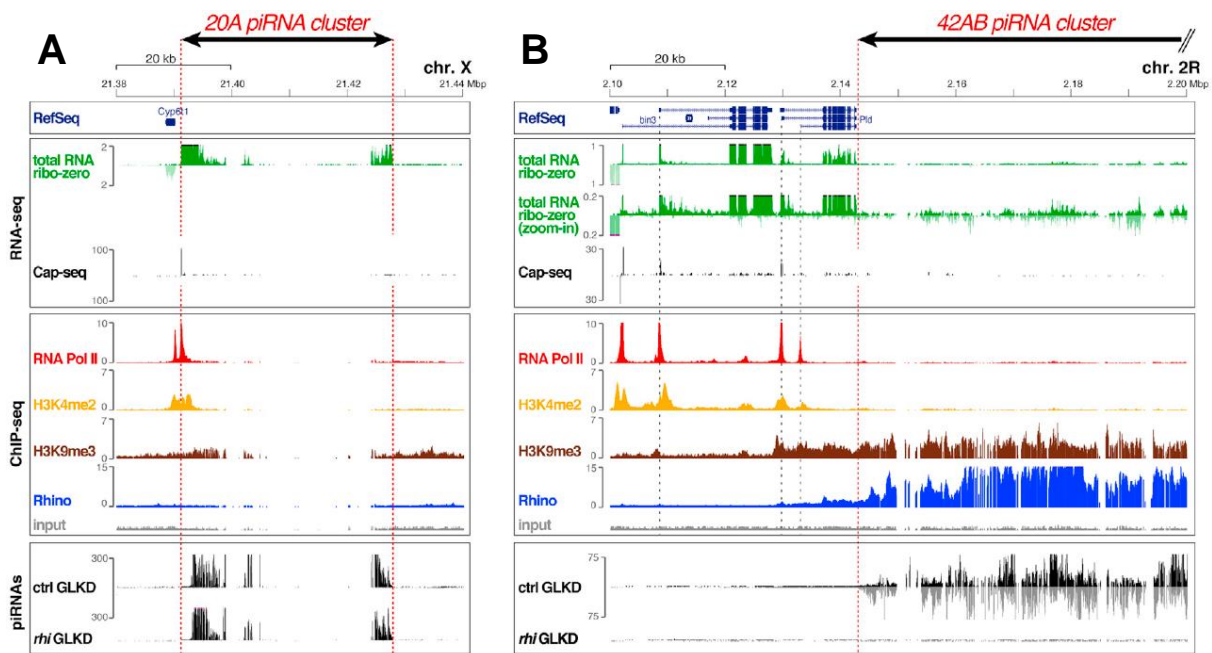
Alors que les piRNA clusters uni-strand sont principalement actifs dans les cellules folliculaires – à l’exception de certains clusters uni-strand comme le cluster *20A* qui est actif dans les cellules folliculaires et dans les cellules germinales – les piRNA cluster dual-strand sont exclusivement transcrits dans les cellules germinales. Ces clusters se caractérisent entre autres par une transcription des deux brins pour la production des ARN précurseurs (Brennecke *et al*, 2007).

Des piRNA clusters dual-strand ont été identifiés dans les séquences subtélomériques – aussi appelées *Telomeric Associated Sequence (TAS)* – des différents chromosomes. Chez *Drosophila melanogaster*, les *TAS* sont divisés en deux grandes familles, en fonction de leur structure : les *TAS-R* et les *TAS-L* (Mason *et al*, 2008). Les *TAS-R* sont localisés sur les bras droits des chromosomes 2 et 3 (appelés *TAS-2R* et *TAS-3R*) ainsi que sur le bras gauche du chromosome X (*TAS-X*). Les *TAS-R* sont constitués de répétitions de l’ET *invader4* ainsi que de sous domaines *TAS* dont les séquences peuvent être redondantes (sous domaines *T1*, *T2*, *T4* et *T5*) ou uniques (sous domaine *T3* présent uniquement dans le *TAS-X*) (Asif-Laidin *et al*, 2017). Les *TAS-L* sont situés sur les bras gauches des chromosomes 2 et 3 (appelés *TAS-2L* et *TAS-3L*) (Karpen & Spradling, 1992) et sont constitués de six répétitions ‘*TLL*’ (*TAS-L like*), un motif particulier de 158 pb qu’on retrouve spécifiquement dans le sous domaine *T3* du *TAS-X* (Asif-Laidin *et al*, 2017). Même si le *TAS-X* peut être perdu dans certaines lignées maintenue dans les laboratoires, on n’a pas identifié de lignées pour lesquelles tous les *TAS* (*TAS-R* et *TAS-L*) sont absents (Asif-Laidin *et al*, 2017). Ces résultats suggèrent que les *TAS* assurent des fonctions biologiques importantes mais qui restent à déterminer.

Le piRNA cluster dual-strand le plus important est le cluster *42AB*. Ce locus, localisé sur le chromosome *2R* au niveau de la région péricentromérique (site cytologique 42AB), mesure environ 240 kilobases et produit plus de 20% des piRNA totaux dans les cellules germinales (Brennecke *et al*, 2007).

Les piRNA clusters dual-strand partagent des similarités avec les piRNA clusters uni-strand. En effet, ils sont transcrits par l’ARN polymérase II et enrichis en marques H3K9me3 (Figure 14A) (Mohn *et al*, 2014; Zhang *et al*, 2014). Cependant, contrairement aux clusters uni-strand, les clusters dual-strand ne semblent pas posséder de promoteur. Pourtant, les fragments d’ET qui composent les piRNA clusters peuvent avoir conservé la séquence de leur promoteur intacte. La mutation de *maelstrom* dans les ovaires entraîne la transcription à partir

des promoteurs d'ET présents à l'intérieur et à proximité des piRNA clusters, sans modification des marques hétérochromatiques. Ainsi la protéine Maelstrom est capable de bloquer les promoteurs potentiellement présents – internes ou externes aux piRNA clusters – ainsi que la transcription canonique, au profit d'une transcription particulière, non canonique, dépendante de facteurs spécifiques (Chang *et al*, 2019). Les marques H3K9me3 sont déposées par l'histone méthyltransférase Egless (Rangan *et al*, 2011) et spécifiquement reconnues par Rhino, une protéine homologue à l'Heterochromatin Protein 1 (HP1) qui reconnaît les mêmes marques (Klattenhoff *et al*, 2009). Ainsi, sur les clusters dual-strand, l'enrichissement en H3K9me3 est associé à un enrichissement en Rhino (Figure 14B). Actuellement, on ne sait pas comment Rhino reconnaît les marques H3K9me3 déposées spécifiquement sur les piRNA clusters par rapport à celles présentes sur le reste du génome.

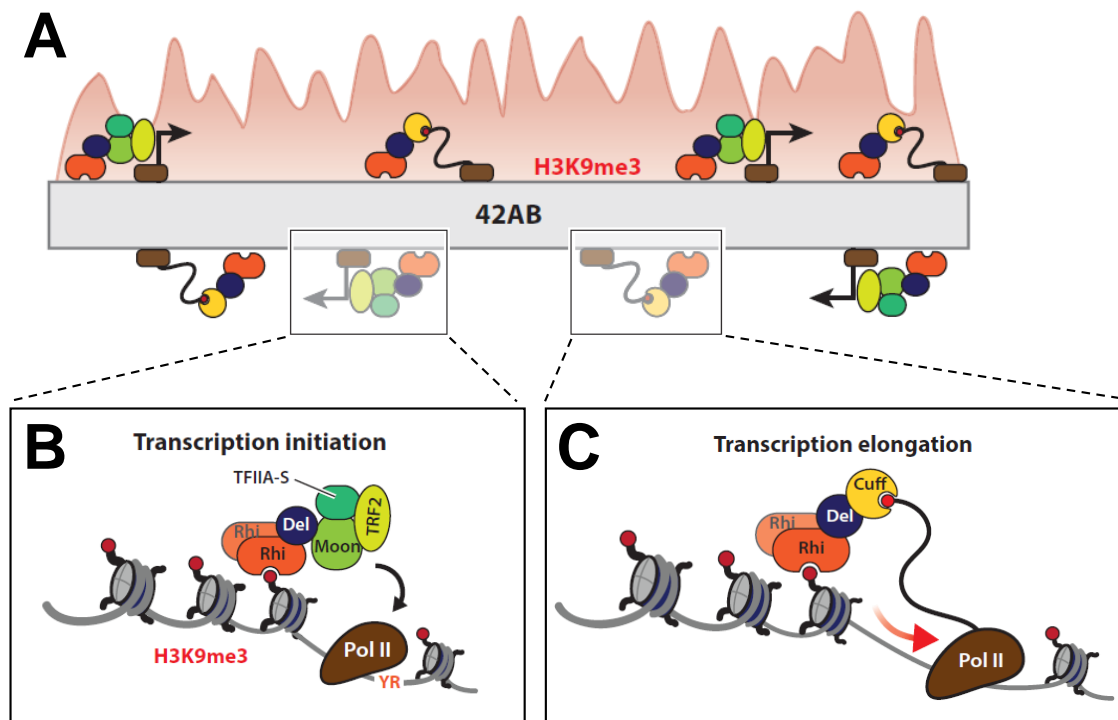


**Figure 14** : Signatures transcriptionnelles des piRNA clusters. (A) Signature transcriptionnelle d'un piRNA cluster uni-strand, le cluster 20A. Les données de RNA-seq et de Cap-seq montrent que le cluster est transcrit et que les ARN précurseurs portent une coiffe méthyl-guanosine à leur extrémité 5'. Le cluster 20A possède un promoteur enrichi en marques H3K4me2 reconnu par l'ARN polymérase II (RNA Pol II). La séquence est enrichie en H3K9me3 mais pas en Rhino. Les piRNA ne sont issus que d'un seul brin et le mutant Rhino n'affecte pas la production de ces piRNA, indiquant que la production des piRNA par les piRNA cluster uni-strand n'est pas dépendante de Rhino (B) Signature transcriptionnelle d'un piRNA cluster dual-strand, le cluster 42AB. Le cluster est transcrit mais les ARN précurseurs ne portent pas de coiffe méthyl-guanosine à leur extrémité 5'. La séquence est enrichie en H3K9me3 et en Rhino mais on ne distingue pas de promoteur spécifique enrichi en H3K4me2, tandis que les promoteurs des gènes adjacents (*Pld* et *bin3*) sont enrichis en



marques H3K4me2 et permettent la production d'ARN portant une coiffe méthyl-guanosine en 5'. Chez le mutant Rhino, le cluster ne produit pas de piRNA indiquant que la production de piRNA issus des deux brins est dépendante de Rhino. Les frontières des piRNA clusters *20A* et *42AB* sont représentées par des pointillés (Mohn *et al*, 2014).

Rhino est requis pour la transcription non-canonique des piRNA clusters dual-strand. Pour ce faire, Rhino recrute les co-facteurs Deadlock et Cutoff pour former le complexe RDC ainsi que Moonshiner (Andersen *et al*, 2017; Mohn *et al*, 2014). Deadlock est une protéine plateforme qui participe à l'interaction indirecte entre Rhino et Cutoff ; son domaine N-terminal interagit avec Rhino tandis que son domaine C-terminal se lie à Cutoff (Mohn *et al*, 2014). Moonshiner est un paralogue de la sous-unité TFIIA-L du complexe d'initiation de la transcription (TFIIA) qui se lie à Deadlock pour recruter les facteurs d'initiation de la transcription TATA-box-binding protein-related factor 2 (TRF2) et transcription factor IIA-S (TFIIA-S). Le complexe ainsi formé reconnaît des motifs de type 'YR' (dinucléotides pyrimidine-purine) pour initier la transcription *via* l'ARN polymérase II (Andersen *et al*, 2017) (Figure 15B). L'existence de tels motifs suggèrent que l'initiation de la transcription peut s'effectuer à de nombreuses positions du piRNA cluster. L'analyse structurale de Cutoff indique une forte homologie avec Rai1/Dxo1 qui, chez la levure, possède une activité permettant le retrait de la coiffe 5'. Chez Cutoff, le domaine catalytique responsable du retrait de la coiffe 5' est muté, rendant l'enzyme inactive. Ainsi Cutoff lie la coiffe à l'extrémité 5' à la place du CAP-binding complex (CBC) et protège les ARN précurseurs de la dégradation. Lors de la transcription, Cutoff bloque les signaux de régulation de la transcription. En effet Cutoff empêche l'épissage – processus auquel l'hélicase UAP56 semble participer puisque sa mutation entraîne un épissage de ces transcrits précurseurs – ainsi que la terminaison de la transcription en bloquant le complexe CPSF (Cleavage-Polyadenylation Specificity Factor) dans la reconnaissance des séquences du signal de polyadénylation (PAS) (Chen *et al*, 2016; Mohn *et al*, 2014; Zhang *et al*, 2014) (Figure 15C).



**Figure 15 :** Transcription d'un piRNA cluster dual-strand – exemple de *42AB*. (A) La séquence de *42AB* est enrichie en marques H3K9me3. (B) L'initiation de la transcription ne dépend pas d'un promoteur. Rhino (Rhi) reconnaît les marques H3K9me3 et interagit avec Deadlock (Del) permettant le recrutement de Moonshiner (Moon). Moonshiner recrute à son tour les facteurs d'initiation de la transcription TFIIA-S et TRF2 qui reconnaissent des motifs de type 'YR'. L'ARN polymérase II est alors recrutée et démarre la transcription à partir de ces motifs. (C) Au cours de la transcription, Cutoff (Cuff) se lie à la coiffe 5' et bloque les signaux de régulation de la transcription (épissage et terminaison). (Adapté de Czech *et al*, 2018).

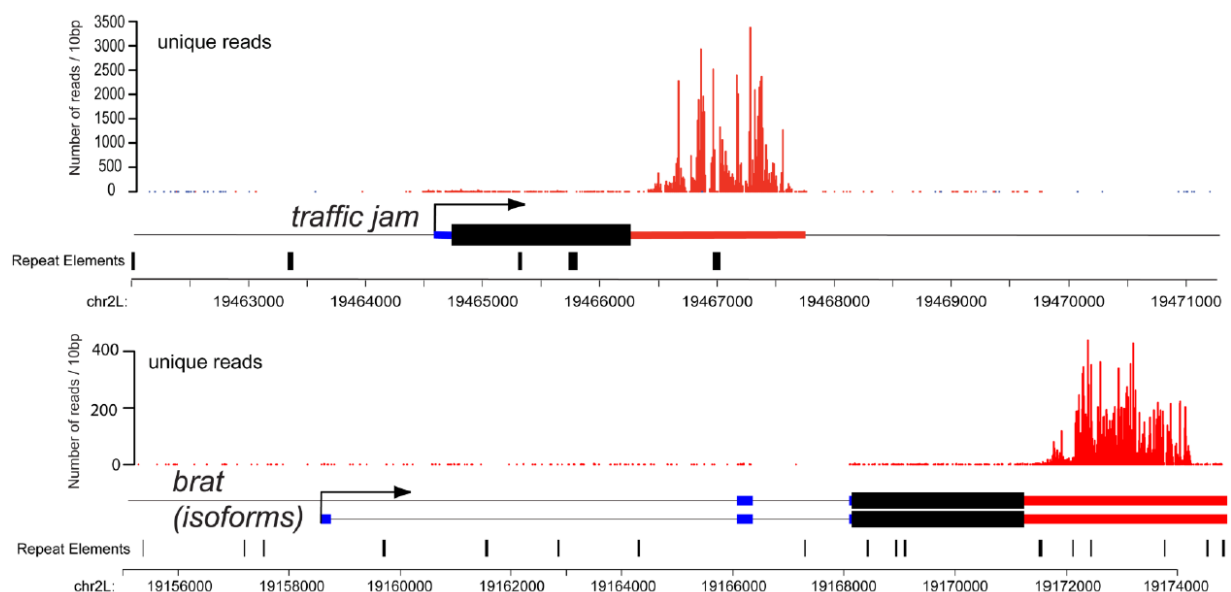
La transcription des clusters dual-strand grâce au complexe RDC-Moonshiner semble se faire de manière aléatoire et générer des transcrits précurseurs de différentes tailles (Andersen *et al*, 2017; Czech *et al*, 2018).

#### - Autres sources de piRNA

Des travaux menés par deux équipes sur les cellules OSS (Ovarian Stem Sheet, lignée cellulaire dérivée des cellules folliculaires) de drosophile ont montré qu'il existe d'autres sources pour la production de piRNA (Robine *et al*, 2009; Saito *et al*, 2009). Des piRNA dérivés de 3'UTR de gènes codants des protéines (aussi appelés piRNA géniques) ont été détectés dans les cellules OSS (voir figure 12B) mais on ne les retrouve pas dans les cellules

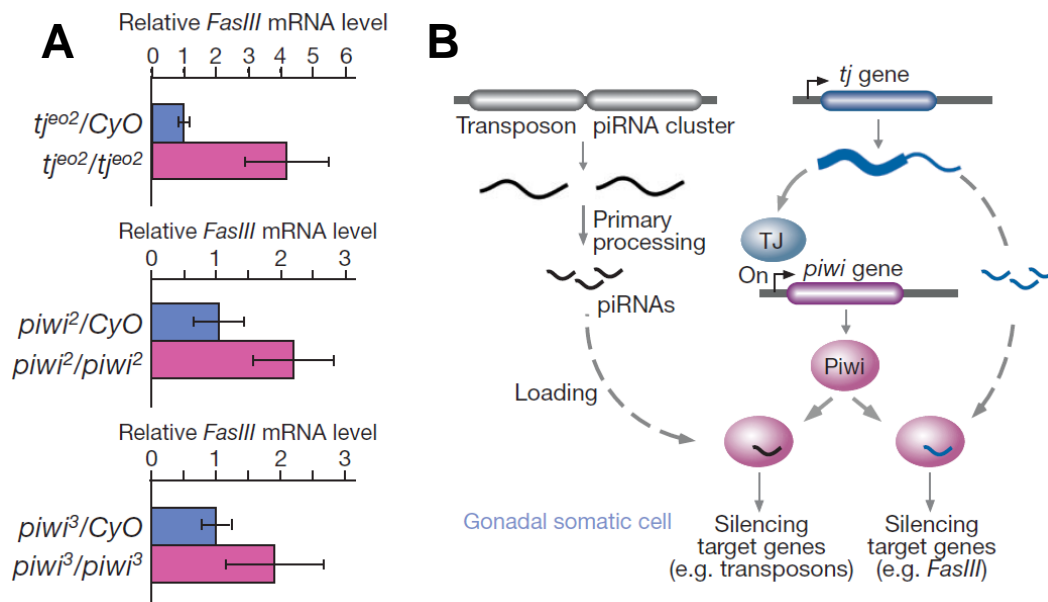
S2 – une lignée cellulaire dérivée de cellules d’embryons de drosophile. Les 3’UTR de 180 gènes ont été identifiés comme source de production de piRNA. Parmi ces régions, les deux loci qui produisent le plus de piRNA géniques sont les 3’UTR des gènes *traffic jam (tj)* et *brain tumor (brat)*, qui sont respectivement un facteur de transcription et un régulateur de la traduction (Robine *et al*, 2009) (Figure 16).

La quantité de piRNA dérivés de 3’UTR n’est pas négligeable puisque les analyses de séquençage ont permis de les estimer à plus d’un million de lectures pour une banque de près de 13,8 millions de lectures (soit 9% des lectures). Mais la quantité de piRNA dérivés d’un 3’UTR n’est pas corrélée à la quantité d’ARNm transcrits (Robine *et al*, 2009). La production des piRNA géniques, puisqu’elle est originaire de séquences non traduites, suggère que les ARNm concernés peuvent à la fois être traduit en protéines pour lesquelles ils codent mais aussi entrer dans différentes voies de biogénèses de petits ARN non-codants – incluant notamment siRNA et piRNA. Il apparait que ces piRNA dérivés de 3’UTR possèdent une fonction biologique.



**Figure 16** : Régions produisant le plus de piRNA dérivés de leur 3’UTR. Exemples de *traffic jam* (en haut) et de *brain tumor (brat)* (en bas). L’alignement des lectures montre un enrichissement des piRNA au niveau des extrémités 3’UTR de ces deux gènes par rapport au reste de la séquence (5’UTR et séquence codante). Cet enrichissement n’est pas corrélé à la présence de séquences répétées, indiquées en dessous des coordonnées par ‘repeat elements’. On peut aussi remarquer que ces piRNA ne sont originaires que du brin sens (Robine *et al*, 2009).

En contexte mutant pour *piwi* ou *tj*, le gène *Fasciclin III* (*FasIII*), qui intervient dans l'adhésion cellulaire, est surexprimé (Saito *et al*, 2009). Ces résultats indiquent que la mutation de *tj* reproduit l'effet de la mutation de *piwi* (Figure 17A). Un modèle propose que la séquence codante de *tj* soit traduite en protéine tandis que la partie 3'UTR produit des piRNA géniques. La protéine produite par *tj* active l'expression de *piwi* et les piRNA dérivés du 3'UTR de *tj* sont chargés par la protéine Piwi, au même titre que les piRNA dérivés de piRNA clusters. Les piRNA dérivés du 3'UTR de *tj* chargés par Piwi sont impliqués dans la répression de *FasIII* (Figure 17B). Ces expériences indiquent que les piRNA ne participent pas seulement au contrôle de l'activité des ET mais aussi à la régulation de certains gènes.



**Figure 17 : Rôle des piRNA dérivés du 3'UTR de *tj* dans la régulation de *FasIII*.** (A) RT-qPCR montrant l'expression relative de *FasIII* en contexte mutant pour *tj* ou *piwi*. Pour ces mutants homozygotes, le gène *FasIII* est plus exprimé par rapport au contrôle mutant hétérozygote. (B) Modèle de la régulation de *FasIII* par les piRNA issus de *tj*. La séquence codante de *tj* est traduite en protéine et l'extrémité 3'UTR produit des piRNA. La protéine TJ active la transcription de *piwi* et les piRNA dérivés du 3'UTR de *tj* sont chargés par la protéine Piwi. Les piRNA dérivés de piRNA clusters chargés par Piwi régulent les séquences d'ET tandis que les piRNA dérivés de *tj* régulent l'expression de *FasIII* (Saito *et al*, 2009).

### 1.3. Biogénèse des piRNA chez *Drosophila melanogaster*

En 2013, trois cribles génétiques basés sur la technique d'ARN interférence ont été entrepris afin d'identifier l'ensemble des facteurs impliqués dans la biogénèse des piRNA dans les cellules folliculaires et germinales de la drosophile (Czech *et al*, 2013; Handler *et al*, 2013; Muerdter *et al*, 2013). Parmi les acteurs moléculaires mis en évidence, on trouve des facteurs connus pour être impliqués dans le transport d'ARN. Selon l'origine des ARN précurseurs – piRNA cluster uni-strand, piRNA cluster dual-strand ou 3'UTR de gène – leur transport ainsi que la biogénèse des piRNA ne requièrent pas exactement les mêmes facteurs, bien que les mécanismes moléculaires soient relativement proches. En effet, alors que les piRNA clusters dual-strand sont actifs spécifiquement dans les cellules germinales, les piRNA clusters uni-strand le sont essentiellement, mais pas exclusivement, dans les cellules folliculaires (Voir figure 12B). Dans le cas particulier des piRNA dérivés de 3'UTR, leur biogénèse nécessite les mêmes facteurs que les piRNA issus de clusters uni-strand notamment parce qu'ils sont produits dans le même type cellulaire.

#### - Export des ARN précurseurs

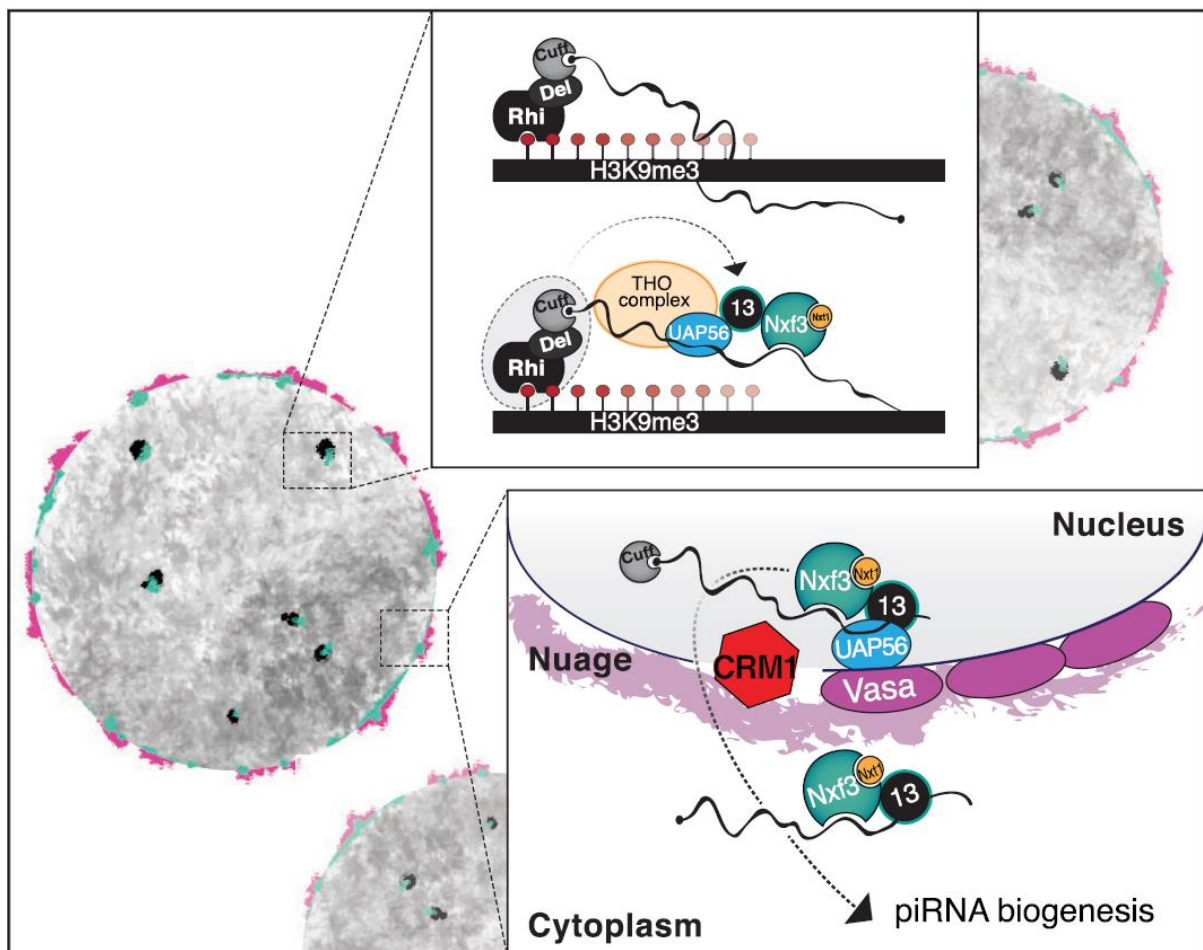
Dans les cellules folliculaires, les ARN précurseurs issus des clusters uni-strand sont exportés vers le cytoplasme par des facteurs d'export généraux, les exportines Nxf1 et Nxt1 ont été identifiés parmi les facteurs requis dans la voie des piRNA (Czech *et al*, 2013; Dennis *et al*, 2016; Handler *et al*, 2013; Muerdter *et al*, 2013). L'export de ces ARN précurseurs nécessite également des facteurs constituant l'Exon Junction Complex (EJC) – incluant Mago, Tsunagi, RnpS1 et Acinus – et l'hélicase UAP56. La mutation de l'un de ces facteurs entraîne une dérégulation des ET ainsi qu'une accumulation des transcrits précurseurs au niveau de leur site de transcription (Czech *et al*, 2013; Dennis *et al*, 2016; Handler *et al*, 2013; Muerdter *et al*, 2013). Grâce à ces facteurs, les ARN précurseurs sont accumulés au niveau d'une structure nucléaire appelée Dot/COM et localisée à proximité de l'enveloppe nucléaire (Dennis *et al*, 2013). Une autre équipe a identifié un site d'accumulation cytoplasmique proche de l'enveloppe nucléaire, appelé flam body (Murota *et al*, 2014). Les deux équipes ont pu mettre en évidence par RNA-FISH que ces deux structures d'accumulation des ARN précurseurs – Dot/COM et flam body – sont proches d'une autre structure cytoplasmique nommée Yb-body (Dennis *et al*, 2013; Murota *et al*, 2014). Les ARN précurseurs se lient à la protéine Yb (aussi

appelé fs(1)Yb) pour être transférés vers les Yb-Bodies. Les Yb-bodies, essentiels à la biogénèse des piRNA, sont constitués de nombreux facteurs protéiques - incluant Vreteno (Zamparini *et al*, 2011), Sister of Yb, Brother of Yb (Handler *et al*, 2011), Shutdown, Hsp83 (Olivieri *et al*, 2012; Preall *et al*, 2012), Armitage et Yb (Saito *et al*, 2010). A l'exception d'Armitage, Yb et Hsp83 (aussi appelé Hsp90), le rôle de ces facteurs protéiques dans la biogénèse des piRNA n'est pas encore connu. Yb et Armi co-localisent au niveau des Yb-bodies et il a été montré qu'Armi est capable de charger des ARN de manière non spécifique en l'absence d'Yb ce qui entraîne la production de piRNA qui ne correspondent pas à des ET. Ainsi, Yb régule la reconnaissance des précurseurs de piRNA par Armi empêchant ainsi la production de piRNA aberrants. De surcroît, Armi interagit également avec la protéine Zucchini (Zuc) impliquée dans le clivage des précurseurs de piRNA en piRNA (Ishizu *et al*, 2019).

Dans les cellules germinales, les ARN précurseurs issus des clusters dual-strand sont exportés vers le cytoplasme grâce à UAP56, qui participe aussi au blocage de l'épissage, aux facteurs d'exports Nxf3 et Nxt1, au co-facteur Bootlegger ainsi qu'au complexe THO/TREX (Transcription/Export) – un complexe conservé chez les animaux et participant à l'export des ARN messagers (ElMaghraby *et al*, 2019; Hur *et al*, 2016; Kneuss *et al*, 2019; Zhang *et al*, 2018). Un crible double hybride chez la levure a permis d'identifier que Deadlock interagit avec Bootlegger et des expériences d'immunofluorescences indiquent que Bootlegger, Nxf3 et UAP56 co-localisent avec Rhino ce qui suggère que le complexe RDC recrute Bootlegger *via* Deadlock et que Bootlegger recrute ensuite Nxf3 et UAP56 (ElMaghraby *et al*, 2019; Kneuss *et al*, 2019; Zhang *et al*, 2012). Des expériences de co-immunoprécipitation suivi de spectrométrie de masse confirment que Nxf3 interagit aussi avec Bootlegger, Nxt1 et UAP56, tandis que des expériences d'immunoprécipitation d'ARN permettent de détecter les ARN précurseurs associés à Nxf3. Ces résultats indiquent que Nxf3 se lie aux ARN précurseurs et recrute les facteurs Nxt1 et UAP56 (ElMaghraby *et al*, 2019; Kneuss *et al*, 2019). Dans le noyau, UAP56 interagit avec le complexe THO, composé de cinq sous-unités protéiques, et co-colocalise avec Rhino à proximité de l'hélicase Vasa, équivalent cytoplasmique d'UAP56 (Zhang *et al*, 2012, 2018). Par ailleurs, on détecte Nxf3, Nxt1 et Bootlegger dans le noyau mais également dans le cytoplasme, contrairement au complexe THO et à UAP56 qui sont exclusivement nucléaires (ElMaghraby *et al*, 2019; Kneuss *et al*, 2019). Lors de la transcription des clusters dual-strand, UAP56 se lie aux transcrits précurseurs et les transfère à

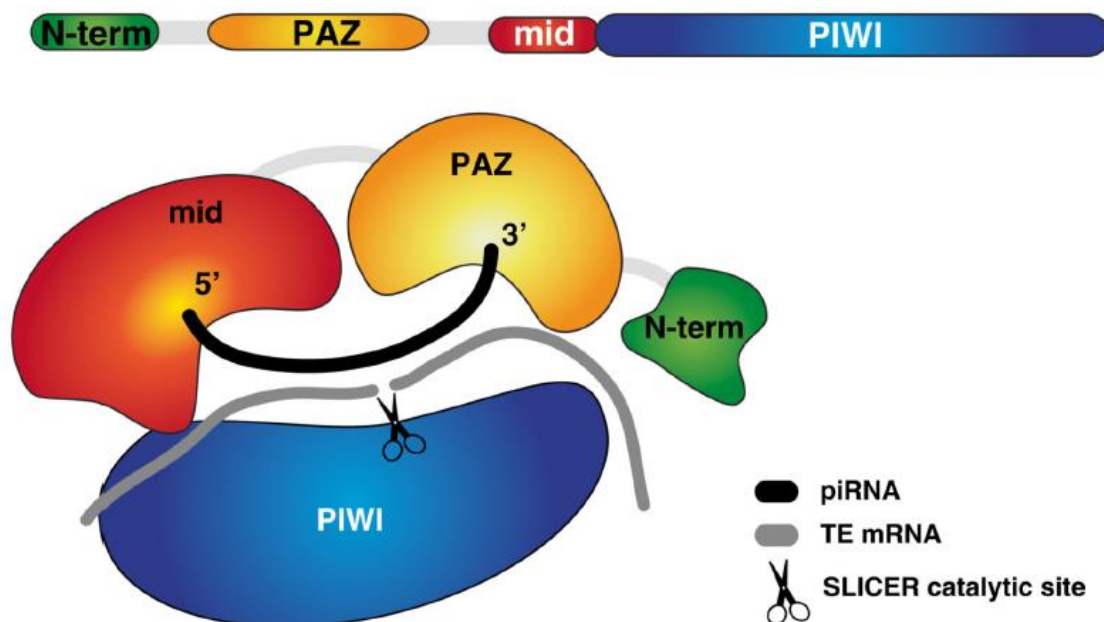
Vasa via l'exportine CRM1 (ElMaghraby *et al*, 2019; Kneuss *et al*, 2019; Zhang *et al*, 2012) (Figure 18).

Vasa est localisé dans une structure périnucléaire appelée 'nuage' qui constitue une alternative aux Yb-bodies des cellules folliculaires (Lim & Kai, 2007). En effet, Yb - facteur essentiel à la formation des Yb-bodies dans les cellules folliculaires - n'est pas exprimé dans les cellules germinales, c'est pourquoi on n'y trouve pas de Yb-bodies (Handler *et al*, 2011). Le nuage regroupe un ensemble de facteurs protéiques parmi lesquels on trouve Shutdown, Hsp83, Vasa, Maelstrom, Krimper, Armitage, Aubergine et Spindle-E (Lim & Kai, 2007; Olivieri *et al*, 2012; Preall *et al*, 2012).



**Figure 18 : Export des ARN précurseurs vers le cytoplasme.** Le complexe RDC, *via* Deadlock (Del) recrute Bootlegger (Boot, '13' sur la figure) qui recrute à son tour Nxf3, Nxt1 et UAP56. UAP56 interagit quant à lui avec le complexe THO. Nxf3 et UAP56 se lient à l'ARN précurseur en cours de transcription puis qu'UAP56 permet le transfert de l'ARN précurseur vers le cytoplasme *via* l'exportine Crm1 (Kneuss *et al*, 2019).

Les Yb-bodies et le nuage sont associés aux mitochondries pour la maturation des piRNA. Lors de leur maturation, les piRNA sont chargés par les protéines PIWI – Piwi, Aubergine (Aub) et Ago3. Ces protéines possèdent une structure commune avec trois domaines : PAZ, MID et PIWI (Figure 19). Les domaines PAZ et MID lient les extrémités des piRNA tandis que le domaine PIWI contient le site catalytique. L'extrémité N-terminal est riche en arginines. Ces résidus peuvent être méthylés par l'arginine méthyl-transférase Capsuleen et son co-facteur Valois. La diméthylation symétrique des arginines du domaine N-terminal permet la liaison avec des protéines possédant des domaines Tudor (Nishida *et al*, 2009; Vagin *et al*, 2009).

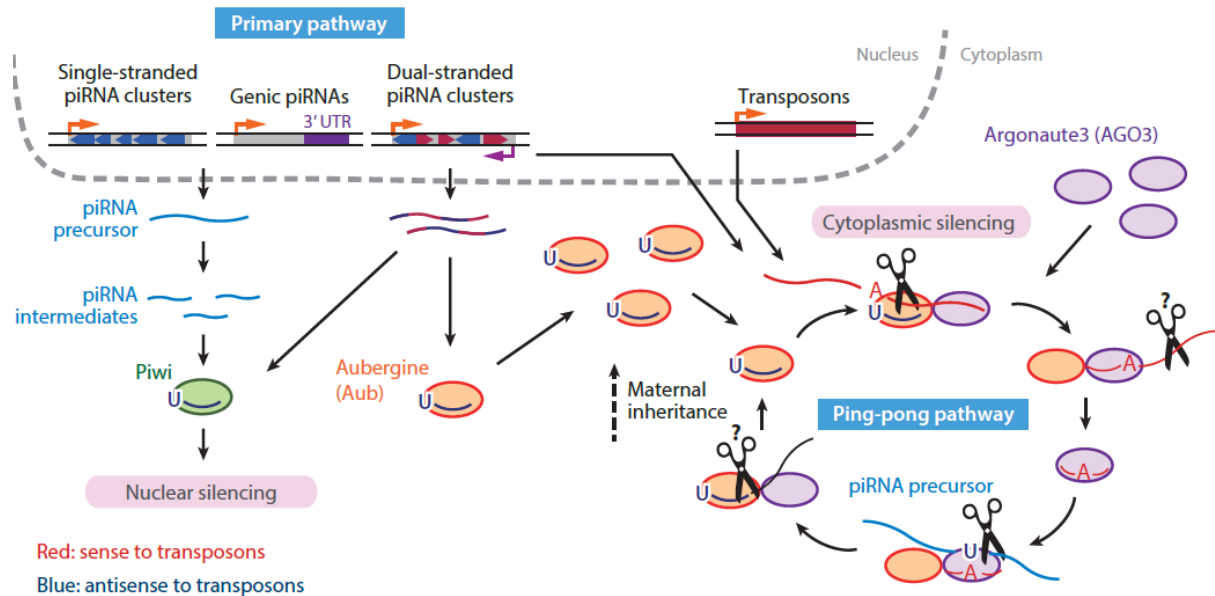


**Figure 19 : Structure des protéines PIWI.** L'extrémité N-terminale est un domaine riche en arginines pouvant être méthylées, il contient aussi un signal de localisation nucléaire pour la protéine Piwi. Les protéines PIWI possèdent trois domaines : PAZ, MID et PIWI. Les domaines MID et PAZ se lient respectivement aux extrémités 5' et 3' des piRNA. Le domaine PIWI contient le site catalytique capable de cliver les transcrits d'ET complémentaires requis pour le PTGS mais pas pour le TGS (Adapté de Tóth *et al*, 2016).

Les piRNA maturés à partir des ARN précurseurs constituent les piRNA primaires et sont produits aussi bien dans les cellules folliculaires que dans les cellules germinales. Ces piRNA primaires sont chargés par la protéine Piwi pour réaliser la régulation transcriptionnelle des ET (Figure 20). Dans les cellules germinales, des piRNA primaires sont chargés par la protéine Aub et permettent la production de piRNA dits secondaires à partir de transcrits



d'ET, *via* un mécanisme appelé cycle ou amplification 'Ping-Pong'. Ces piRNA secondaires sont chargés par les protéines Aub et Ago3 pour accomplir la régulation post-transcriptionnelle des ET (Figure 20). Les gènes *aub* et *ago3* ne sont pas exprimés dans les cellules folliculaires, ce qui explique l'absence de piRNA secondaires.



**Figure 20 : Biogénèse des piRNA dans les ovaires de drosophile.** Les ARN précurseurs issus des piRNA clusters ou dérivés de 3'UTR sont transportés vers le noyaux et maturés en piRNA primaires. Les piRNA primaires sont chargés par Piwi pour réaliser la régulation transcriptionnelle des ET. Dans le cas particuliers des piRNA clusters dual-strand, une partie des piRNA primaires est chargée par la protéine Aub et permet de générer des piRNA secondaires par le mécanisme d'amplification 'Ping-Pong'. Les piRNA secondaires sont chargés par les protéines Aub et Ago3 pour exécuter la régulation post-transcriptionnelle des ET (Iwasaki *et al*, 2015).

#### - piRNA primaires et régulation transcriptionnelle

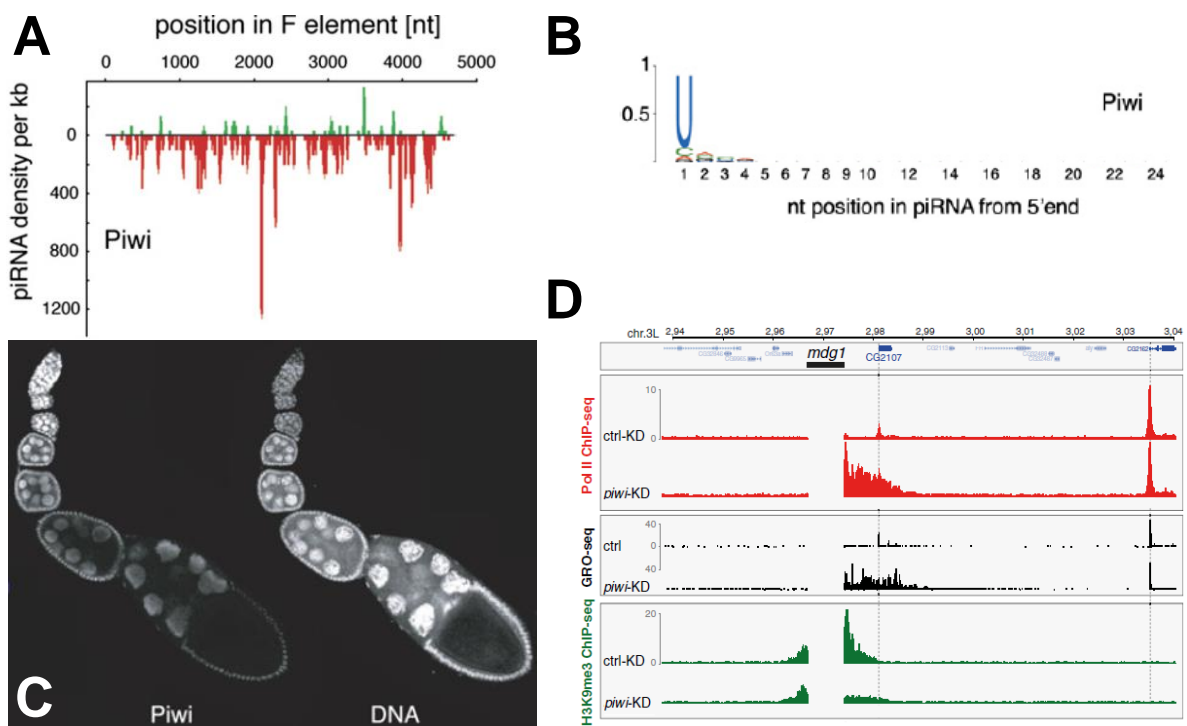
Dans les cellules folliculaires et germinales, les ARN précurseurs sont clivés par l'endonucléase Zucchini, qui possède un signal de localisation mitochondrial en N-terminal lui permettant d'être localisée sur la membrane externes des mitochondries (Pane *et al*, 2007; Saito *et al*, 2010). Zuc clive spécifiquement les ARN simple-brins et nécessite la présence d'autres co-facteurs puisque son activité in-vitro est relativement faible (Ipsaro *et al*, 2012; Nishimasu *et al*, 2012). Dans les cellules folliculaires, les co-facteurs Minotaur et Gasz ont été

identifiés (Czech *et al*, 2013; Muerdter *et al*, 2013; Vagin *et al*, 2013) tandis que dans les cellules germinales la protéine Squash a été mise en évidence (Pane *et al*, 2007) mais les fonctions de ces protéines dans la biogénèse des piRNA ne sont pas encore clairement identifiées. Deadalus (Dead), un facteur commun entre les cellules folliculaires et germinales présent au niveau de la membrane externe des mitochondries, participe à la localisation correcte d'Armi, Zuc et Piwi (Murano *et al*, 2019).

Les fragments d'ARN précurseurs clivés par Zuc sont chargés par Piwi et Aub dans les cellules germinales ou par Piwi seule dans les cellules folliculaires et constituent les piRNA primaires. Shutdown et Hsp83 (aussi appelée Hsp90) sont requis pour la prise en charge des piRNA - primaires et secondaires - par les protéines PIWI (Olivieri *et al*, 2012; Preall *et al*, 2012). Les piRNA primaires sont généralement antisens (Figure 21A) et possèdent majoritairement une uridine en position 1, à l'extrémité 5' – qualifié de biais 1U – suggérant que Zuc clive préférentiellement les ARN précurseurs au niveau des uridines (Brennecke *et al*, 2007; Gainetdinov *et al*, 2018) (Figure 21B). La taille des extrémités 3' des piRNA primaires est modulée par trois facteurs : Zuc, l'exonucléase 5' à 3' Nibbler et l'ARN méthyltransférase Hen1. La mutation de Zuc ou de Nibbler génère des piRNA primaires plus longs mais n'affecte pas la longueur des piRNA secondaires indiquant que ces deux facteurs sont responsables du clivage de l'extrémité 3' des piRNA primaires (Hayashi *et al*, 2016; Wang *et al*, 2016). Hen1 catalyse la 2'O méthylation de l'extrémité 3' des piRNA et sa mutation entraîne la production de piRNA plus courts (Horwich *et al*, 2007; Saito *et al*, 2007; Wang *et al*, 2016). Ainsi, l'ajout d'un groupement 2'O-methyl à l'extrémité 3' semble protéger les piRNA du clivage par Zuc et Nibbler.

Seuls les complexes Piwi-piRISC (piRNA-induced silencing complex) – piRNA guide associé à la protéine Piwi – sont impliqués dans la régulation transcriptionnelle (ou TGS pour Transcriptional Gene Silencing) des ET. La protéine Piwi est localisée dans le noyau, contrairement aux protéines Aub et Ago3 qui sont présentes dans le cytoplasme (Brennecke *et al*, 2007) (Figure 21C). Dans le noyau, les complexes Piwi-piRISC scannent l'ensemble des ARN messagers en cours de synthèse, mais ne ciblent pas les séquences ADN, et répriment les séquences complémentaires aux piRNA chargés par Piwi. Par conséquent, seuls les ET actifs dans le génome peuvent être ciblés par les complexes Piwi-piRISC (Brower-Toland *et al*, 2007; Klenov *et al*, 2014, 2011; Le Thomas *et al*, 2014). Le domaine catalytique de Piwi ne semble pas participer au TGS puisqu'en contexte mutant pour *piwi* l'expression d'une protéine Piwi possédant un site catalytique inactif est capable de restaurer la répression des

ET. En revanche, une délétion de la partie N-terminale de Piwi (mutant *piwi<sup>Nt</sup>*), qui contient un signal de localisation nucléaire (NLS), entraîne sa rétention dans le cytoplasme ainsi que la mobilisation des ET (Klenov *et al*, 2011; Saito *et al*, 2010; Sienski *et al*, 2012). L'association entre une protéine Piwi et un piRNA guide est indispensable pour la répression des séquences cibles : le chargement d'un piRNA primaire stabilise Piwi et entraîne sa translocation vers le noyau tandis que la protéine Piwi seule est retenue dans le cytoplasme (Saito *et al*, 2010). En outre, le recrutement artificiel de Piwi sans piRNA guide ne permet pas la régulation des séquences cibles (Le Thomas *et al*, 2013; Sienski *et al*, 2015). Ces résultats indiquent que l'association entre un piRNA primaire et Piwi est requise à la fois pour la localisation nucléaire et pour le TGS.



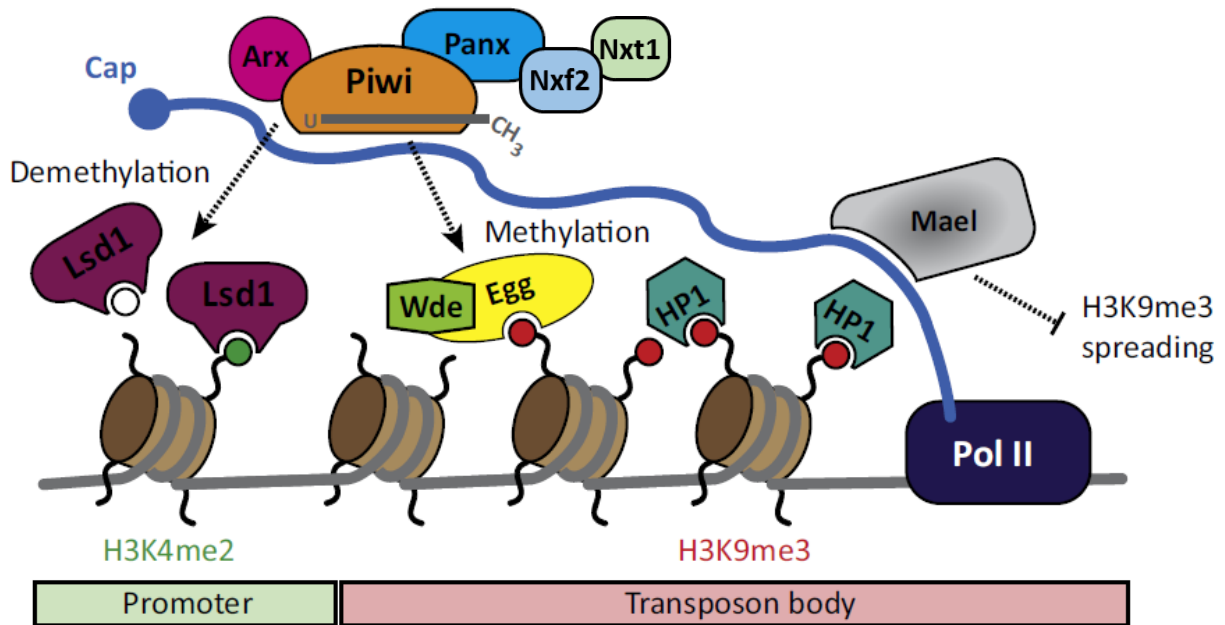
**Figure 21 : Protéine Piwi et piRNA associés à Piwi.** (A) La protéine Piwi charge majoritairement des piRNA antisens (en rouge), exemple de piRNA provenant de l'élément *F*. (B) Les piRNA chargés par Piwi possèdent en général une uridine en position 1, Zuc semble cliver préférentiellement les ARN précurseurs au niveau de ce nucléotide. Cette particularité est appelée biais 1U (C) Immunofluorescence montrant la localisation de Piwi dans les noyaux des cellules d'une ovariole de drosophile (D) Phénotype moléculaire du mutant *piwi<sup>Nt</sup>*, exemple d'une région du chromosome 3 (site cytologique 63E) contenant une insertion de l'ET *mdg1* (rectangle noir). Des expériences de ChIP-seq montrent que l'ARN polymérase II occupe le promoteur et que les marques H3K9me3 ne sont plus présentes chez le mutant *piwi<sup>Nt</sup>*. Des expériences de GRO-seq indiquent également que les ARN naissants sont plus

nombreux chez le mutant *piwi<sup>Nt</sup>* comparé au contrôle (Adapté de Brennecke *et al*, 2007; Sienski *et al*, 2012).

Chez un mutant *piwi<sup>Nt</sup>*, l'absence de Piwi dans le noyau s'accompagne de la diminution de la présence des marques H3K9me3 (Figure 21D) et de la protéine HP1a au niveau des séquences de certains ET, ce qui indique que la présence de Piwi est nécessaire pour l'apposition de ces marques hétérochromatiques (Klenov *et al*, 2014, 2011; Le Thomas *et al*, 2013). Des expériences de co-immunoprécipitation montrent qu'Asterix (Arx) et Panoramix (Panx) interagissent physiquement avec Piwi (Dönertas *et al*, 2013; Sienski *et al*, 2015). De même, des expériences de co-immunoprécipitation et de spectrométrie de masse indiquent que Panx interagit physiquement avec deux facteurs impliqués dans le transport des ARN, Nxf2 et Nxt1, pour former un complexe protéique appelé SFiNX (pour silencing factor interacting nuclear export variant), PICTS (pour Panx-induced co-transcriptional silencing) ou encore Pandas (pour Panoramix-dNxf2-dependent TAP/p15 silencing) en fonction des équipes qui l'on décrit parallèlement (Batki *et al*, 2019; Fabry *et al*, 2019; Murano *et al*, 2019; Zhao *et al*, 2019). La mutation de l'un de ces co-facteurs reproduit le phénotype moléculaire d'un mutant *piwi* entraînant une dérépression des ET, une perte des marques H3K9me3 et une augmentation du nombre d'ARN naissant mais n'affectent pas la quantité de piRNA produits (Batki *et al*, 2019; Dönertas *et al*, 2013; Fabry *et al*, 2019; Murano *et al*, 2019; Ohtani *et al*, 2013; Sienski *et al*, 2015; Yu *et al*, 2015; Zhao *et al*, 2019). Ainsi, ces mutants affectent la régulation transcriptionnelle des ET mais ne sont pas requis dans la biogénèse des piRNA. L'activité de tous ces facteurs : Arx, Panx, Nxf2 et Nxt1 mais également de SUMO, de la SUMO ligase Su(var)2-10 et de son co-facteur Ubc9 sont requises pour la mise en place du TGS. Su(var)2-10 interagit avec Piwi, le piRNA guide et Panx tandis que SUMO intervient dans la sumoylation de la chromatine (Ninova *et al*, 2020). Su(var)2-10 *via* SUMO recrute la méthyltransférase Eggless (Egg, aussi appelé SetDB1) qui reconnaît la sumoylation au niveau de la chromatine et réalise le dépôt des marques H3K9me3 à l'aide de l'activité du co-facteur Windei (Koch *et al*, 2009; Ninova *et al*, 2020; Yoon *et al*, 2008). Parallèlement, les facteurs Lsd1 et CoRest sont impliqués dans la déméthylation des marques H3K4me2/3 tandis les facteurs Mi-2 et Rpd3 participent respectivement au remodelage de la chromatine et à la déacétylation des résidus H3K9 (Mugat *et al*, 2020; Yu *et al*, 2015) (Figure 22). En l'absence des marques H3K9me3, la protéine HP1a qui interagit avec le domaine N-terminal de Piwi ne peut pas se fixer et les ET ne sont pas réprimés (Brower-Toland *et al*, 2007; Klenov *et al*, 2011).

Bien que l'absence de Piwi dans le noyau soit associée à une diminution du dépôt de H3K9me3 et de HP1a, l'analyse par ChIP-seq du mutant *piwi<sup>Nt</sup>* révèle également que cette diminution n'est pas systématique. Chez un mutant *piwi*, l'occupation de l'ARN polymérase II au niveau des promoteurs des ET est plus importante (Klenov *et al*, 2014; Sienski *et al*, 2012) (Figure 21D). Des expériences d'immunofluorescence indiquent également que Piwi co-localise avec l'ARN polymérase II (Le Thomas *et al*, 2013). Ces résultats suggèrent que la régulation de certains ET ne dépend pas nécessairement d'une hétérochromatinisation locale et qu'il existe un mécanisme alternatif dépendant de l'ARN polymérase II. Grâce à un crible génétique basé sur la répression d'un transgène rapporteur, Grzegorz Sienski et ses collègues ont mis en évidence que Maelstrom (Mael) participe à la mise en place de ce mécanisme alternatif (Sienski *et al*, 2012). En effet, le mutant *mael* ressemble à un mutant *piwi*, il génère une dérégulation des ET, une augmentation de l'occupation de l'ARN polymérase II ainsi que du nombre d'ARN naissant. En revanche, la production de piRNA et le dépôt des marques H3K9me3 ne sont pas affectés en contexte mutant pour *mael*. Toutefois, l'absence de Mael génère une extension des marques H3K9me3 sur les gènes adjacents (Sienski *et al*, 2012). Piwi et Mael semblent bloquer le recrutement de l'ARN polymérase II, empêchant la transcription des ET, mais les mécanismes moléculaires impliqués ne sont pas clairement identifiés (Figure 22).

Ainsi, deux mécanismes complémentaires interviennent lors de la mise en place du TGS des ET : l'hétérochromatinisation locale et le blocage de l'ARN polymérase II (Figure 22). Par ailleurs, des facteurs antagonistes au TGS ont été identifiés, notamment le complexe PAF1 qui s'associe à l'ARN polymérase II. En l'absence PAF1, les transcrits d'ET ne sont plus maturés et la voie des piRNA est favorisée (Clark *et al*, 2017). L'ensemble de ces résultats suggèrent qu'il peut y avoir des fuites transcriptionnelles et que le TGS seul n'est pas suffisant pour la régulation des ET.

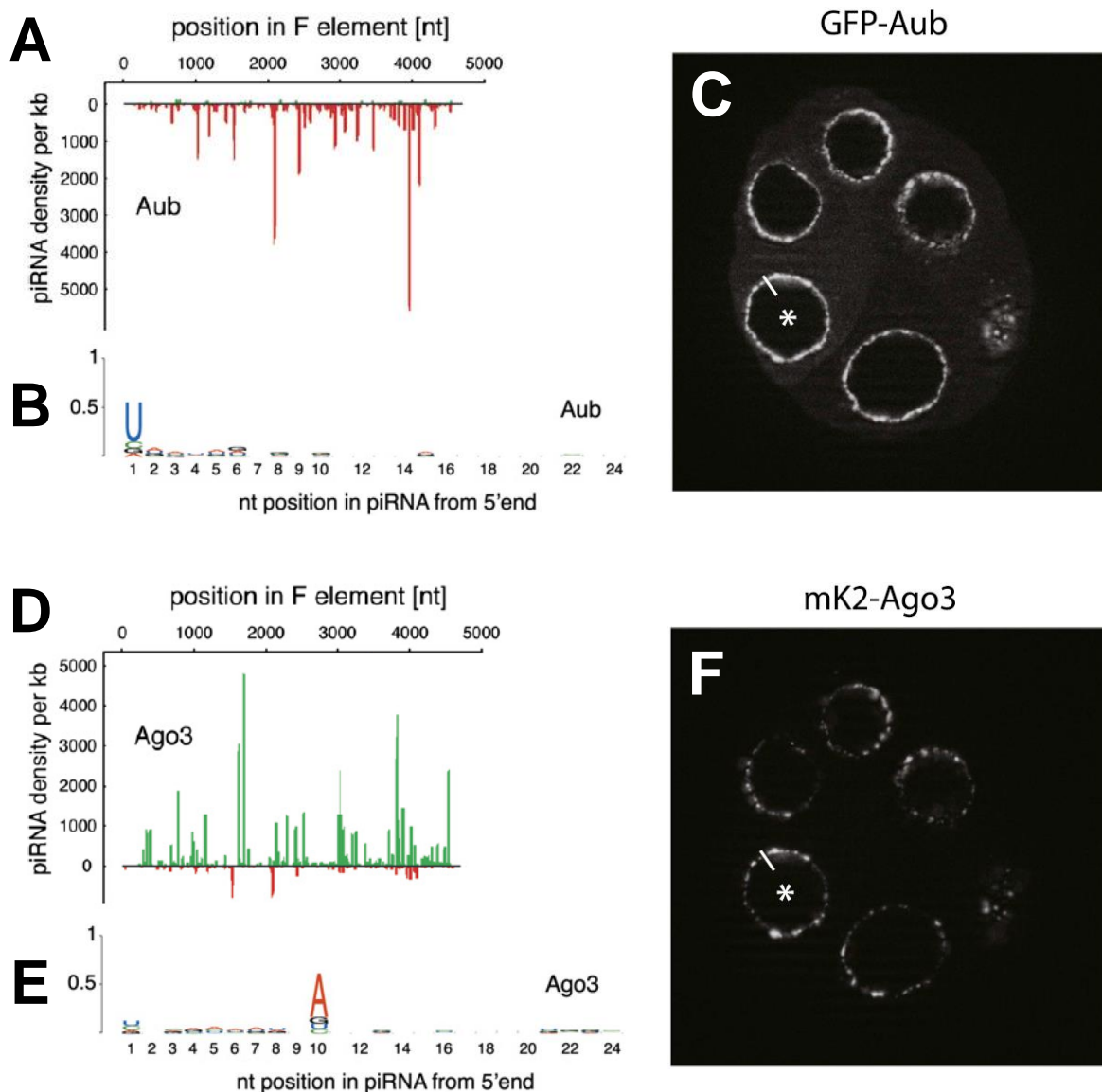


**Figure 22 : Régulation transcriptionnelle des ET.** Les ET sont transcrits par l'ARN polymérase II (Pol II). Le complexe Piwi-piRISC reconnaît les transcrits en cours de synthèse tandis que les co-facteurs Asterix (Arx), Panoramax (Panx), Nxf2 et Nxt1 recrutent la méthyltransférase Eggless (Egg). Egg et son co-facteur Windei (Wde) déposent les marques H3K9me3 tandis que Lsd1 déméthyle les marques H3K4me3. Parallèlement, Piwi et Mael bloquent le recrutement de l'ARN polymérase II. Les marques H3K9me3 sont reconnues par HP1a qui est impliquée dans l'hétérochromatinisation locale de la séquence (Adaptée de (Czech & Hannon, 2016)).

#### - piRNA secondaires et régulation post-transcriptionnelle

Dans les cellules germinales, une partie des piRNA clivés par Zuc est prise en charge par Aub pour la formation de complexes Aub-piRISC. Comme pour Piwi, ces piRNA associés à Aub sont majoritairement antisens et possèdent un biais 1U (Brennecke *et al*, 2007) (Figures 23A et 23B). Aub est localisé dans le cytoplasme, essentiellement au niveau du nuage (Brennecke *et al*, 2007; Webster *et al*, 2015) (Figure 23C). Ces données suggèrent que les complexes Aub-piRISC reconnaissent et clivent les ARN messagers – sens – des ET. Les fragments d'ARN messagers clivés par Aub-piRISC ne sont pas dégradés, ils génèrent de nouveaux piRNA appelés piRNA secondaires et chargés par Ago3, formant des complexes Ago3-piRISC. C'est vraisemblablement parce qu'ils sont dérivés de transcrits d'ET que les piRNA secondaires chargés par Ago3 sont majoritairement sens (Figure 23D) et portent généralement une adénine en position 10 – appelé biais 10A (Figure 23E). Ainsi, les complexes Ago3-

piRISC localisés dans le nuage (Figure 23F) peuvent reconnaître et, de même que Zuc, cliver les ARN précurseurs antisens transcrits à partir des piRNA clusters dual-strand. Ce clivage génère d'autres piRNA secondaires pris en charge par Aub, conduisant à la formation de nouveaux complexes Aub-piRISC capables de reconnaître et de cliver les transcrits d'ET (Figure 24A). Ce mécanisme, appelé cycle ou amplification 'Ping-Pong', assure la régulation post-transcriptionnelle des ET et permet également l'amplification du nombre de complexes Aub-piRISC et Ago3-piRISC. La taille des extrémités 5' des piRNA secondaires est modulée par Aub et Ago3. En revanche, la taille des extrémités 3' dépend de l'activité de Nibbler et de Hen1. Les piRNA secondaires, comme les piRNA primaires, portent un groupement 2'O-méthyl à l'extrémité 3' (Hayashi *et al*, 2016; Wang *et al*, 2016).

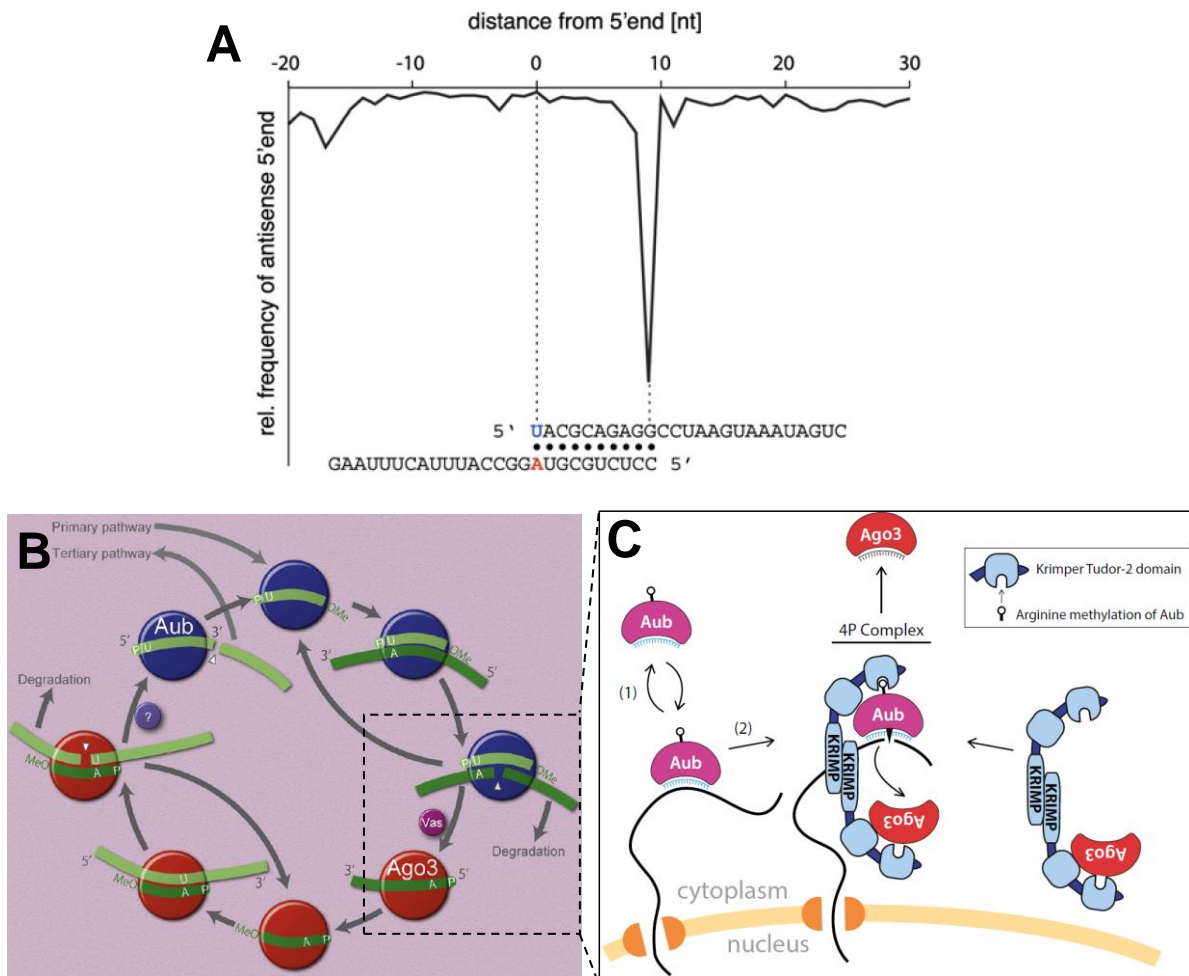


**Figure 23 : Protéines Aub et Ago3 et piRNA associés à Aub et Ago3.** (A) La protéine Aub charge essentiellement des piRNA antisens, exemple de piRNA provenant de l'élément *F* (B) Les piRNA chargés par Aub possèdent en général une uridine en position 1 (biais 1U) (C) Expérience de FRAP (Fluorescence Recovery After Photobleaching) montrant la localisation d'Aub tagguée par la GFP. Aub est présent dans le cytoplasme, essentiellement au niveau du nuage (D) La protéine Ago3 charge majoritairement des piRNA sens, exemple de piRNA provenant de l'élément *F* (E) Les piRNA chargés par Ago3 possèdent généralement une adénine en position 10 (biais 10A) (F) Expérience de FRAP (Fluorescence Recovery After Photobleaching) montrant la localisation d'Ago3 tagguée par mKate2. Ago3 est enrichi au niveau du nuage (Adapté de Brennecke *et al*, 2007; Webster *et al*, 2015).

Une fraction des piRNA associés à Aub et à Ago3 sont complémentaires sur les dix premiers nucléotides de leur extrémité 5'. Lors de l'appariement entre un complexe Aub-piRISC et un transcrit d'ET, l'uridine en position 1 du piRNA guide est en face d'une adénine. Le transcrit d'ET est clivé exactement entre les 10<sup>ème</sup> et 11<sup>ème</sup> nucléotides complémentaires au piRNA guide (Brennecke *et al*, 2007; Gunawardane *et al*, 2007). Le biais 10A des piRNA chargés par Ago3 dépend d'Aub qui semble se lier préférentiellement à des ARN contenant une adénine en phase avec l'extrémité 5' de son piRNA guide (1U) (Wang *et al*, 2014b). Cette caractéristique propre aux piRNA secondaires est appelée signature ping-pong (Figure 24A). Le Ping-Pong requiert de nombreux facteurs protéiques qui composent le nuage. Chez *Bombyx mori*, il a été montré que Vasa interagit avec Siwi – l'orthologue d'Aub – et qu'en contexte mutant pour *vasa* les ARN clivés restent associés aux complexes piRISC. Vasa participerait à la libération des ARN clivés par les complexes piRISC (Nishida *et al*, 2015; Xiol *et al*, 2014). Shutdown et Hsp83 participent à la prise en charge des piRNA secondaires par les protéines PIWI (Aub et Ago3) dans le nuage (Olivieri *et al*, 2012; Preall *et al*, 2012). Les mécanismes permettant le recrutement d'Aub et d'Ago3 vers le nuage sont différents. En effet, en l'absence d'un piRNA guide, Aub est délocalisé du nuage alors que la localisation d'Ago3 n'est pas affectée. En revanche, l'absence de Krimper, une protéine à domaine Tudor, affecte la localisation d'Ago3 mais pas celle d'Aub. Ces résultats indiquent qu'Aub est localisé dans le nuage grâce au piRNA guide alors que la localisation d'Ago3 dépend de Krimper (Webster *et al*, 2015). Des expériences de co-immunoprécipitation montrent que Krimper peut former des dimères *via* son domaine N-terminal, que son domaine Tudor situé en C-terminal interagit avec Ago3 et peut aussi reconnaître les résidus arginines méthylés d'Aub (Webster *et al*, 2015) (Figure 24C). Par ailleurs, la mutation de *krimper* n'affecte pas seulement la localisation d'Ago3, elle entraîne également la prise en charge des piRNA primaires par Ago3 et l'initiation d'un cycle Ping-Pong entre deux Aub (Sato *et al*, 2015). Ces données indiquent que Krimper est nécessaire à l'amplification Ping-Pong car il est impliqué



dans la formation d'un complexe entre Aub et Ago3, qui permet aux ARN clivés par l'un d'être chargés par l'autre (Figure 24B-C). Qin, une autre protéine à domaine Tudor, est requise pour empêcher le chargement par Piwi, au lieu d'Ago3, des ARN clivés par Aub (Wang *et al*, 2015). Enfin, d'autres protéines incluant Spindle-E (Spn-E), Tejas (Tej) et Tapas (Tap) colocalisent avec les composants du nuage : Aub, Ago3, Vasa et Krimper. La mutation de *spn-E* ainsi que la double mutation *tej-tap* entraînent la disparition du Ping-Pong, mais leurs rôles restent encore à déterminer (Andress *et al*, 2016; Patil *et al*, 2014; Patil & Kai, 2010).



**Figure 24 : Mécanisme d'amplification Ping-Pong.** (A) Exemple de représentation de la signature Ping-Pong. La distance par rapport à l'extrémité 5' est représentée en abscisse et la fréquence relative de complémentarité. Le pic en position 9 correspond à la position pour laquelle la probabilité de trouver une piRNA complémentaire à l'extrémité 5' est maximale. Un exemple de complémentarité entre un piRNA associé à Aub (au dessus) et un piRNA associé à Ago3 (en dessous) est représenté. (B) Les complexes Aub-piRISC, qui portent un piRNA guide antisens, clivent des transcrits d'ET sens et génèrent des piRNA secondaires – sens - chargés par Ago3. Les complexes Ago3-piRISC ainsi formés reconnaissent et clivent des transcrits précurseurs antisens provenant de piRNA clusters dual-strand, permettant la formation de nouveau piRNA secondaires. Ces piRNA sont alors chargés par Aub. (C) Les

complexes Aub-piRISC sont recrutés dans le nuage et reconnaissent les transcrits d'ET. Le domaine Tudor de Krimper reconnaît la méthylation des arginines d'Aub et interagit avec Ago3. Krimper peut former des dimères *via* son domaine N-terminal, permettant le rapprochement entre Aub et Ago3. Ainsi, les ARN clivés par Aub sont chargés directement par Ago3 (Adapté de Brennecke *et al*, 2007; Hirakata & Siomi, 2016; Webster *et al*, 2015).

Les mécanismes de TGS et de PTGS dans les cellules germinales sont complémentaires et permettent une régulation très efficace des ET. Ces mécanismes de répression sont mis en place très tôt, dès l'embryon, et sont maintenus tout au long du développement dans les cellules germinales (Marie *et al*, 2017).

#### **1.4. Les piRNA chez d'autres espèces**

Des analyses de données de séquençage montrent que les protéines PIWI sont conservées chez les animaux : elles ont été identifiées chez les spongiaires, les cnidaires et certains bilatériens – incluant les nématodes, la drosophile et l'Homme. Pour ces espèces, la signature Ping-Pong, qui témoigne de la production de piRNA secondaires, a pu être identifiée indiquant que la voie des piRNA est active (Grimson *et al*, 2008). D'autres études permettent de compléter cette liste puisque la voie des piRNA a été identifiée pour d'autres organismes : la souris (Girard *et al*, 2006), les moustiques (Miesen *et al*, 2015), les primates (Hirano *et al*, 2014) ou encore les mollusques (Jehn *et al*, 2018). Cependant, on ne trouve pas de protéines PIWI chez différents groupes de plantes ou chez certains eucaryotes unicellulaires (Grimson *et al*, 2008). L'ensemble de ces données suggèrent que les piRNA seraient apparus chez l'ancêtre commun des métazoaires.

Les facteurs impliqués dans la biogénèse des piRNA sont conservés chez les animaux mais de nombreux facteurs supplémentaires sont spécifiques à la drosophile, notamment les facteurs du complexe RDC, Moonshiner, Panx, Nibbler ou Krimper (Ozata *et al*, 2019). L'existence de facteurs spécifiques suppose que la biogénèse des piRNA chez la drosophile est différente de celle des autres animaux. Chez la plupart des animaux par exemple, la transcription des piRNA clusters repose exclusivement sur la machinerie de transcription canonique, la présence de promoteurs et l'activité de facteurs de transcription (Andersen *et al*, 2017). Ce mode de transcription est le même que celui des piRNA clusters uni-strand chez la drosophile, en revanche il diffère du mode de transcription des piRNA clusters dual-strand. On peut imaginer que dans les cellules germinales de la drosophile l'absence de promoteur clairement

défini a conduit à la sélection d'un mécanisme de transcription alternatif qui fait intervenir le complexe RDC et Moonshiner – des facteurs absents chez les autres animaux. Par ailleurs, chez les autres animaux il semble exister d'autres sources pour la production de piRNA. En effet, des piRNA dérivés d'ARN de transfert ou de longs ARN non-codants ont été identifiés respectivement chez la souris et chez l'Homme (Ha *et al*, 2014; Hirano *et al*, 2014; Keam *et al*, 2014). Une étude comparative indique que chez la plupart des Arthropodes la voie des piRNA est active dans les cellules somatiques et dans les cellules germinales. Sur vingt espèces d'Arthropodes étudiées, quatre ont perdu les piRNA dans les cellules somatiques au cours de l'évolution. *Drosophila melanogaster* fait partie de ces espèces qui ont perdu les piRNA dans les cellules somatiques (hors cellules folliculaires ovariennes) alors que chez une autre espèce de drosophile, *Drosophila virilis*, les piRNA somatiques ont été conservés. Il s'avère également que chez les espèces qui ont conservé la voie des piRNA dans les cellules somatiques, les piRNA sont capables de cibler les ET mais aussi des ARN messagers (Lewis *et al*, 2018). Ainsi, les piRNA semblent être impliqués dans d'autres fonctions que celle de la régulation des ET.

## 2. D'autres fonctions pour les piRNA

De nombreux travaux montrent que les piRNA ne jouent pas seulement un rôle dans la régulation des ET. Les piRNA sont notamment responsables de la régulation de gènes impliqués dans des fonctions biologiques variées ou encore dans la défense virale chez certaines espèces de moustiques.

### 2.1. Régulation de gènes

Parmi les gènes dont la régulation dépend de l'activité de piRNA chez la drosophile, on trouve le gène *FasIII*. Les piRNA géniques produits à partir du 3'UTR du gène *traffic-jam* sont chargés par les protéines PIWI et permettent la régulation de *FasIII* (Saito *et al*, 2009).

En 2010, il a été montré que le gène *Nanos* (*Nos*), qui participe à la formation de l'axe antéro-postérieur et à la détermination de la lignée germinale chez la drosophile, est régulé par les piRNA au stade embryon (Rouget *et al*, 2010). Cette étude montre que dans l'embryon de drosophile, la taille de la queue poly(A) des ARNm de *Nos* diminue progressivement au cours

du temps et entraîne la dégradation des ARNm. Le déadénylation et la dégradation des ARNm dépendent respectivement de l'activité du complexe CCR4-NOT et de la protéine Smaug (Smg). En contexte mutant pour la voie des piRNA, la déadénylation des ARNm de *Nos* n'a plus lieu, en conséquence les ARNm ne sont plus dégradés. Des expériences de co-immunoprécipitation indiquent que CCR4-NOT, Smg et Aub interagissent et que les ARNm de *Nos* co-précipitent avec la protéine Aub. De plus, l'analyse de la séquence de l'ARNm de *Nos* révèle d'une part que le 3'UTR contient des sites de liaisons à Smg et d'autre part que le 5'UTR possède des séquences complémentaires aux ET *roo* et *412*. Les piRNA complémentaires au 3'UTR de *Nos* co-immunoprécipitent avec Aub et l'absence des séquences complémentaires à *roo* et *412* situées dans le 5'UTR empêche la déadénylation des ARNm de *Nos*. Enfin l'injection d'anti-piRNA de *roo* et *412* ne permet pas le développement normal de l'embryon, montrant qu'un piRNA guide est requis pour la régulation de *Nos*. Ces résultats indiquent que, dans l'embryon de drosophile, les complexes Aub-piRISC ciblent les ARNm de *Nos*, puis recrutent le complexe CCR4-NOT et Smg pour induire la déadénylation puis la dégradation de ARNm de *Nos*. La régulation de *Nos* au stade embryon est essentielle pour le développement normal de l'individu puisque l'expression ectopique de ce gène dans le pôle postérieur, dont les ARNm sont détectés par hybridation in-situ, bloque la traduction des gènes *bicoid* et *hunchback*, ce qui empêche le développement normal de la tête (Rouget *et al*, 2010). Quelques années plus tard, des expériences de iCLIP ont permis de mettre en évidence qu'une centaine d'ARNm interagit avec Aub au stade embryon chez la drosophile. Une part de ces ARNm, qui codent des facteurs impliqués dans la détermination de la lignée germinale, est dégradée par CCR4-NOT et Smg (Barckmann *et al*, 2015).

La régulation par les piRNA de gènes impliqués dans différentes fonctions a aussi été mise en évidence chez d'autres espèces. Chez la souris, par exemple, les piRNA sont impliqués dans la dégradation d'ARNm au cours de la spermatogénèse via l'activité du complexe de déadénylation CCR4-CAF1-NOT (Gou *et al*, 2014). Chez l'anémone *Nematostella vectensis*, l'analyse de données de séquençage montre qu'une partie des piRNA s'aligne sur des gènes, pour lesquels des produits de dégradation sont détectés (Praher *et al*, 2017). Chez l'aplysie, on détecte des piRNA dans le cerveau qui participent à la mise en place de la mémoire à long terme. Les complexes piRISC régulent indirectement le répresseur de transcription *CREB2* en aidant à la méthylation de son promoteur (Rajasethupathy *et al*, 2012). Chez *Bombix mori*, des piRNA particuliers sont produits à partir d'un locus constitué de séquences répétées et sont impliqués dans la détermination du sexe. Ces piRNA répriment les ARNm du gène *Masc*,

requis pour la production d'une isoforme femelle-spécifique du gène *doublesex* (Kiuchi *et al*, 2014).

## 2.2. Défense virale

Toutes les études ne s'accordent pas sur le rôle de la voie des piRNA dans la réponse antivirale chez *Drosophila melanogaster*. Dans une première étude, un crible perte de fonction basé sur la technique d'ARN interférence a montré que certains gènes de la voie des piRNA sont requis dans la réponse antivirale chez *Drosophila melanogaster*. Les mutants pour les gènes *Piwi*, *Aub* ou *armi* survivent moins à l'infection par le Drosophila X Virus (DXV) que les contrôles (Zambon *et al*, 2006). En revanche, dans une étude plus récente, les mutants des gènes *Piwi*, *Aub* et *Ago3* n'apparaissent pas plus sensibles à différentes infections virales (DXV, Sindbis virus et Drosophila C Virus) que les contrôles (Petit *et al*, 2016). Dans les cellules OSS (Ovarian Somatic Sheet) de drosophile, des petits ARN dérivés de virus ont été identifiés après infection par différents virus d'insectes. Ces petits ARN mesurent entre 24 et 30 nucléotides et ont un biais 1U mais pas de biais 10A, suggérant qu'il s'agit de piRNA primaires dérivés de virus, aussi appelés vpiRNA (Virus-derived piRNA) (Wu *et al*, 2010).

Des vpiRNA ont aussi été détectés dans des tissus somatiques de moustiques, *Aedes aegypti* et *Aedes albopictus*. Ces vpiRNA mesurent entre 23 et 30 nucléotides et ont un biais 1U et 10A, ce qui indique d'une part qu'il y a des vpiRNA primaires et secondaires, et d'autre part que leur biogénèse fait intervenir le cycle Ping-Pong. Après infection par le virus du chikungunya (CHIKV), la production des vpiRNA augmente (Morazzani *et al*, 2012) et l'inactivation de la voie des piRNA en lignée cellulaire entraîne l'augmentation de la production de particules virales (Schnettler *et al*, 2013). Les vpiRNA limitent la pathogénicité du CHIKV en contexte mutant pour *dicer-2*, qui abolit la voie des siRNA, aussi impliquée dans la réponse anti-virale. Ainsi, les voies des siRNA et des piRNA sont redondantes dans la défense virale, et l'inactivation de l'une de ces deux voies peut être compensée par l'autre (Morazzani *et al*, 2012).

L'analyse comparative de données de séquençage chez différentes espèces de moustiques a permis d'identifier sept protéines Piwi, contre une seule chez *Drosophila melanogaster* (Campbell *et al*, 2008). La perte de fonction de *Piwi5* et d'*Ago3* empêche la production de vpiRNA secondaires. L'absence des autres protéines Piwi bloque la production des piRNA de

la voie classique mais n'empêche pas celle des vpiRNA. Ces résultats montrent que certaines protéines Piwi sont impliquées dans la biogénèse des piRNA tandis que Piwi5, notamment, participe à la biogénèse des vpiRNA. Piwi5 et Ago3 co-localisent dans le cytoplasme, Piwi5 lie préférentiellement les piRNA antisens avec un biais 1U alors qu'Ago3 lie les piRNA sens avec un biais 10A (Miesen *et al*, 2015). Ces données rappellent l'amplification Ping-Pong dans les cellules germinales chez *Drosophila melanogaster*, où Piwi5 est l'orthologue d'Aub.

### 2.3. Les piRNA, biomarqueurs de maladies ?

Les piRNA et les protéines PIWI sont impliqués dans différents processus biologiques : ils participent à la régulation des ET ainsi que de certains gènes et ils sont requis dans la défense virale. De ce fait, la dérégulation de la voie des piRNA peut avoir des conséquences biologiques importantes.

Depuis quelques années les travaux établissant un lien entre piRNA et cancers se multiplient. Chez l'Homme, il existe quatre protéines PIWI, appelées PIWIL (pour PIWI-Like). La dérégulation de l'expression des protéines PIWIL a été identifiée dans de nombreux types de cancers grâce à des tissus ou sur des lignées cellulaires (Mei *et al*, 2013). Il a été montré que PIWIL2, *via* son domaine PAZ, se lie à la protéine STAT3 (Signal Transducer and Activator of Transcription 3) et induit la formation d'un complexe protéique PIWIL2/STAT3/c-Src. STAT3 est alors phosphorylé par c-Src entraînant la translocation du complexe PIWIL2/STAT3/c-Src vers le noyau. Enfin, le complexe PIWIL2/STAT3/c-Src se lie au promoteur du gène *p53* et empêche sa transcription (Lu *et al*, 2012). Une autre étude montre que PIWIL2 régule l'expression de *c-Myc* en se liant à NME2, un activateur transcriptionnel de *c-Myc*. La mutation de PIWIL2 empêche l'expression de *c-Myc* et permet de réduire la prolifération cellulaire (Yao *et al*, 2014). Ainsi, la protéine PIWIL2 participe à la tumorigénèse, d'une part en réduisant l'apoptose et d'autre part en induisant la prolifération cellulaire (Lu *et al*, 2012; Yao *et al*, 2014).

La dérégulation de nombreux piRNA a été identifiée dans plusieurs types de cancers : des piRNA sont surexprimés tandis que d'autres sont moins exprimés (Mei *et al*, 2013; Ng *et al*, 2016). Certains de ces piRNA semblent avoir des effets inverses puisque une partie d'entre eux est impliquée dans la prolifération cellulaire et d'autres dans l'inhibition de la prolifération cellulaire. La comparaison des transcriptomes de plusieurs organes indique que

le profil d'expression mais aussi la distribution en taille des piRNA sont différents entre des tissus sains et des tumeurs (Martinez *et al*, 2015). D'autres piRNA semblent participer à la migration cellulaire, c'est notamment le cas de piR-Hep1 qui est surexprimé dans des lignées cellulaires de carcinome. En effet, l'inactivation de piR-Hep1 diminue la motilité et le caractère invasif des cellules tumorales alors que sa surexpression entraîne l'augmentation de la migration cellulaire (Law *et al*, 2013). Un autre exemple est celui de piR-823 qui est surexprimé dans des tumeurs de cancers colorectaux (CRC). La surexpression de piR-823 dans des lignées cellulaires non tumorales induit la prolifération cellulaire tandis que l'inhibition de piR-823 dans des lignées cellulaires dérivées de CRC bloque la prolifération cellulaire, entraîne l'apoptose ainsi que la répression de l'expression de gènes codant des protéines heatshock. Des expériences d'immunoprécipitation d'ARN et de western blot montrent que piR-823 se lie et induit la phosphorylation de la sérine 236 (Ser<sup>236</sup>) de HSF1 (Heat Shock Factor 1), un facteur de transcription impliqué dans l'activation de protéines heatshock. La phosphorylation du résidu Ser<sup>236</sup> de HSF1 est nécessaire pour sa translocation vers le noyau puis son activation (Yin *et al*, 2017). Par ailleurs, il a été montré que l'expression de *HSF1* et d'autres protéines heatshock est associée au développement de cancers colorectaux, ce qui suggère un rôle de piR-823 dans la tumorigénèse *via* l'activation de HSF1 (Cen *et al*, 2004).

Les origines de ces piRNA sont variées, puisque certains proviennent des piRNA clusters (entre 10 et 15%) et une grande partie correspond à des séquences codantes ou des transcrits non-codants (près de 70%) (Martinez *et al*, 2015). L'expression aberrante des protéines PIWIL ou la dérégulation de certain piRNA correspond à une signature moléculaire qui varie d'un cancer à l'autre, en fonction du type de tissu touché ou du type de tumeur. Cette signature moléculaire pourrait constituer un marqueur de cancer et aider à l'établissement de pronostic sur l'évolution de la maladie (Mei *et al*, 2013; Ng *et al*, 2016).

Des analyses transcriptomiques d'échantillons provenant de cerveaux humains révèlent qu'environ 150 piRNA sont dérégulés chez des patients atteints de la maladie d'Alzheimer, la plupart étant surexprimés. Par ailleurs, ces piRNA sont prédits pour la régulation des gènes impliqués dans des processus neurophysiologiques ou l'apoptose (Roy *et al*, 2017). Ces données suggèrent que la dérégulation des piRNA pourrait être à l'origine de la maladie d'Alzheimer ou d'autres maladies neuro-dégénératives. Pour le moment, le lien entre la dérégulation des piRNA et d'autres maladies n'a pas encore été établi.

### 3. La biologie des piRNA clusters

#### 3.1. Des outils d'étude de la biologie des piRNA clusters

Notre équipe de recherche s'intéresse, entre autres, à la biologie des piRNA clusters germinaux ; nous cherchons à caractériser l'émergence des piRNA clusters, à identifier leurs propriétés fonctionnelles et moléculaires, ainsi qu'à comprendre leur dynamique. Pour ce faire, nous utilisons un ensemble de transgènes. Parmi ceux-ci, il y a des transgènes rapporteurs, aussi appelés transgènes « cibles », qui permettent de déterminer s'il y a, ou non, production de piRNA ; et des transgènes dits « silencer » parmi lesquels le transgène *P(lArB)* qui est inséré dans un piRNA cluster naturel et le cluster *BX2* qui est constitué de transgènes répétés et mime un piRNA cluster.

##### - Transgènes « cibles »

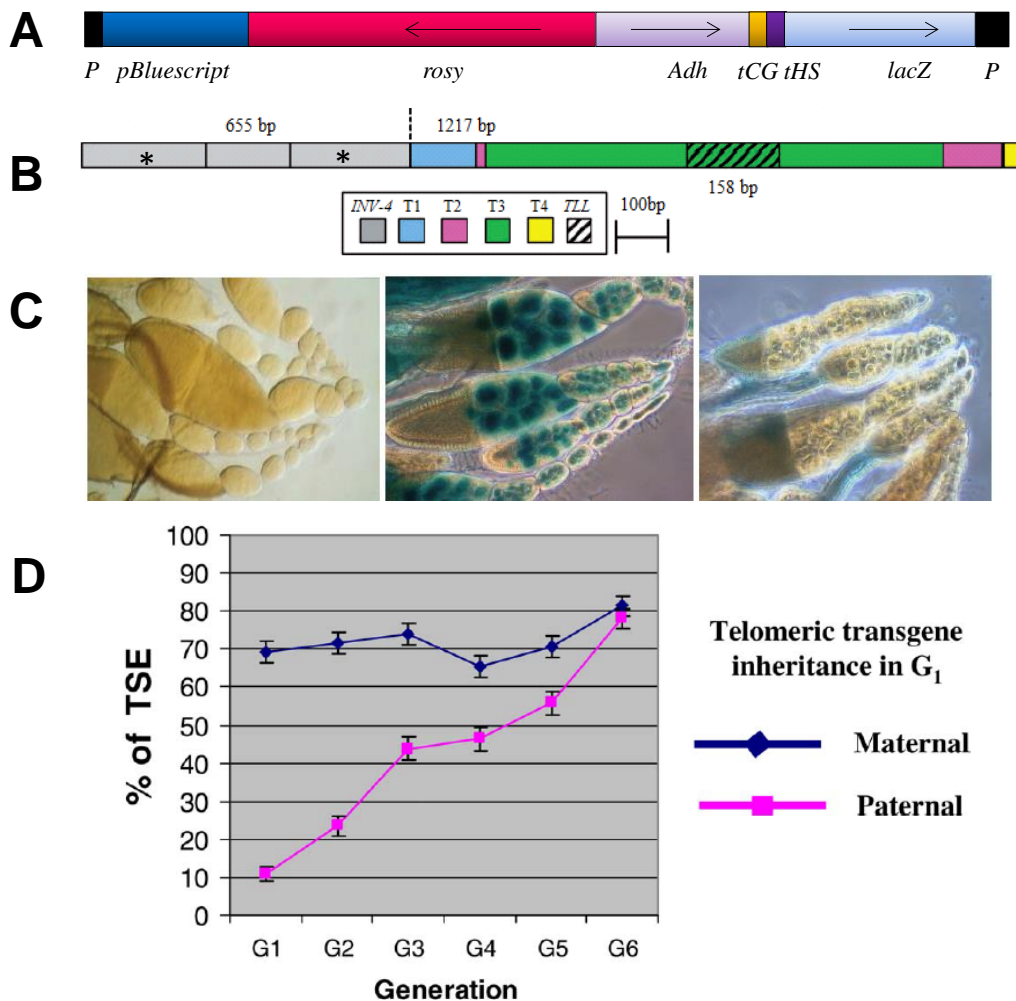
Au laboratoire, nous pouvons utiliser deux techniques différentes pour déterminer si une séquence produit des piRNA ou non, l'une moléculaire et l'autre fonctionnelle. La première technique est le séquençage des petits ARN (small RNA-seq) suivi d'analyses bioinformatiques des banques de données générées. La seconde technique est basée sur les propriétés de *Trans-Silencing* des piRNA clusters. En effet, l'insertion d'un transgène *P-lacZ* au niveau des régions subtélomériques permet de réprimer une autre copie de transgène *P-lacZ* dans le génome (Roche & Rio, 1998). Cette répression, qualifiée de *Trans-Silencing Effect* (TSE), dépend des piRNA ; la perte de fonction des gènes impliqués dans la biogénèse des piRNA affecte le TSE (Josse *et al*, 2007; Todeschini *et al*, 2010). Les transgènes silenciers et les transgènes cibles (= *P(TARGET)*) contiennent la séquence du gène *lacZ*, par conséquent il existe des séquences d'homologies entre ces deux types de transgènes (Voir dans matériel et méthodes « Balanceurs et transgènes »). L'expression de la  $\beta$ -galactosidase, protéine codée par le gène *lacZ*, permet de déterminer si un transgène silencer produit ou non des piRNA. En effet, lorsqu'un transgène silencer ne produit pas de piRNA, le transgène cible s'exprime et la  $\beta$ -galactosidase est produite ; à l'inverse lorsqu'un transgène silencer produit des piRNA, le transgène cible est réprimé et la  $\beta$ -galactosidase n'est pas produite. On peut déterminer si la  $\beta$ -galactosidase est produite ou non par une technique de mesure de son activité enzymatique. Cette technique, appelée 'coloration *lacZ*' ou 'coloration  $\beta$ -galactosidase' est basée sur



l'utilisation de X-Gal (5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside), un substrat chromogénique dont l'hydrolyse par la  $\beta$ -galactosidase entraîne, entre autres, la formation d'un composé bleu. En l'absence de  $\beta$ -galactosidase, le X-Gal n'est pas hydrolysé et la coloration bleue n'apparaît pas.

#### - **Le transgène *P(lArB)***

Le transgène *P(lArB)* est une construction dérivée de l'élément *P* contenant notamment les gènes *rosy*, *adh* et *lacZ* et qui est inséré dans les séquences subtélomériques du chromosome *X* (ou *TAS-X*) (Figure 25A). Le *TAS-X* a été identifié comme étant un piRNA cluster (cluster *IA*) composé des plusieurs séquences de l'ET *invader-4*, de domaines possédant des séquences d'homologies avec les *TAS-R* et d'un domaine unique appelé *T3*, qu'on ne retrouve pas dans les *TAS* des autres chromosomes ni ailleurs dans le génome (Asif-Laidin *et al*, 2017) (Figure 25B). Le transgène *P(lArB)* est capable de réprimer un transgène cible, ce qui indique qu'il peut produire des piRNA (Figure 25C). Lorsque le transgène *P(lArB)* est hérité maternellement pendant plusieurs générations, le pourcentage de TSE (qui est une estimation des capacités répressives du transgène *P(lArB)*) reste constant. En revanche, lorsque le transgène *P(lArB)* est hérité paternellement à la première génération, le pourcentage de TSE augmente au cours des générations (Figure 25D). Ces résultats indiquent que le TSE a un effet maternel et suggèrent que le phénomène de répression dépend de composants cellulaires apportés par la mère. Notons que le terme TSE est peu utilisé actuellement, on parle plutôt de pourcentage de répression.



**Figure 25 : Structure, localisation et propriétés du transgène *P(lArB)*.** (A) Le transgène *P(lArB)* est dérivé de l'élément *P* et mesure 18.3 kb. Il contient les pieds de *P* en 5' et en 3', le plasmide *Bluescript*, ainsi que les gènes *rosy*, *Adh* et *lacZ* – les flèches représentent le sens de transcription de chaque gène. Les terminateurs de transcription de *Adh* et *lacZ* correspondent respectivement aux terminateurs des gènes *CG8231* (*tCG*) et *Hsp70Ab* (*tHS*). (B) Le *TAS-X* est constitué de plusieurs séquences de l'ET *Invader-4* (*Inv-4*) en gris, de domaines qu'on retrouve dans les *TAS-R* (*T1*, *T2* et *T4*, respectivement en bleu, rose et jaune) et d'un domaine spécifique qui est absent des autres *TAS*, le domaine *T3* (en vert). Le domaine *T3* contient une séquence particulière qui est conservée chez d'autres espèces de drosophiles, appelé *TAS-L like* (*TLL*, en hachuré). Deux transgènes *P(lArB)* sont insérés dans le *TAS-X*, leur localisation est représentée par une étoile '\*'. (C) Photos d'ovaires de drosophile après coloration *lacZ*. A gauche, les ovaires contiennent deux copies du transgène *P(lArB)* insérés dans le *TAS-X* ; on voit que la  $\beta$ -galactosidase n'est pas produite car les ovaires sont blancs, ce qui indique que le gène *lacZ* du transgène *P(lArB)* n'est pas capable de produire la  $\beta$ -galactosidase. Au milieu, les ovaires contiennent un transgène cible *PlacZ* seul (transgène *P(A92)* de la lignée BQ16, voir Matériel et Méthodes) mais pas d'insertion du transgène *P(lArB)* dans le *TAS-X* ; on observe une coloration bleue dans les noyaux des cellules germinales ce qui indique que le gène *lacZ* du transgène cible s'exprime. A droite, les ovaires contiennent deux copies du transgène *P(lArB)* insérés dans le *TAS-X* ainsi qu'un transgène cible *PlacZ*; on n'observe pas d'expression du gène *lacZ* par le transgène cible, ce qui indique que le transgène *P(lArB)* réprime l'expression du transgène cible. (D) Pourcentage de TSE jusqu'à la sixième génération lorsque que le transgène *P(lArB)* est hérité maternellement (en bleu) ou

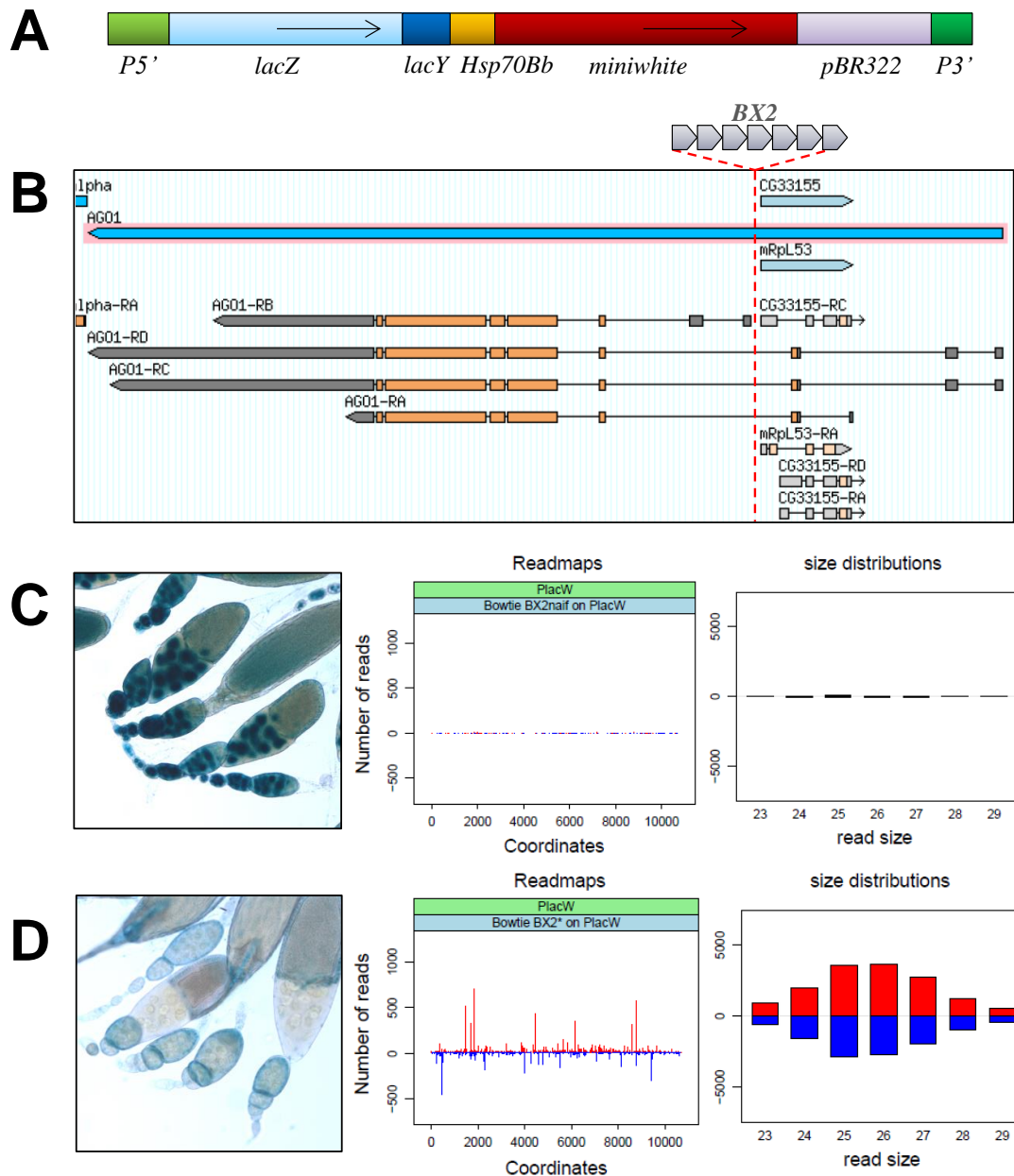
paternellement (en rose). Le pourcentage de TSE est constant lorsque le transgène *P(lArB)* est hérité maternellement alors qu'il augmente lorsque le transgène *P(lArB)* est hérité paternellement. Notons toutefois que le transgène est hérité maternellement aux générations suivantes ; l'héritage maternel ou paternel du transgène ne concerne que la première génération (Adaptée de Asif-Laidin *et al*, 2017; Josse *et al*, 2007).

#### - Le cluster *BX2*

*BX2* est un cluster constitué de sept transgènes *P-lacZ-white* (*P[lacW]*) répétés en tandem et inséré dans le gène *ago1*, sur le chromosome 2 (site cytologique 50C) (Figure 26A et 26B). Le cluster *BX2* a été obtenu grâce à des événements de transposition successifs du transgène *P[lacW]* chez des drosophiles possédant une mutation du gène *white* endogène (Dorer & Henikoff, 1994). Le gène *white* est impliqué dans la pigmentation de l'œil de la drosophile, par conséquent lorsque le gène s'exprime les yeux de la drosophile sont rouges mais lorsqu'il ne s'exprime pas, les yeux de la drosophile sont blancs. Ainsi, l'expression du transgène *P[lacW]* restaure la pigmentation de l'œil en contexte mutant pour le gène *white*. Cependant, la formation d'une structure répétée contenant plusieurs transgènes *P[lacW]* engendre un phénotype de variéation dans l'œil qui induit la répression du gène *white* de manière stochastique ; plus le nombre de copies de *P[lacW]* est élevé, plus la variéation est importante. Ce phénomène appelé RIGS (Repeat-Induced Gene Silencing) s'accompagne de la formation d'hétérochromatine au niveau de la région répétée (Dorer & Henikoff, 1997, 1994).

En exposant une lignée portant le cluster *BX2* à des rayons X, Douglas Dorer et Steven Henikoff ont obtenu un cluster appelé *T-1* (Dorer & Henikoff, 1997). Des expériences de southern blot montrent que le traitement par des rayons X dans ces lignées a induit des réarrangements chromosomiques, on détecte notamment une translocation des chromosomes 2 et 3. Pourtant, ces réarrangements chromosomiques n'affectent ni la structure du transgène, ni sa localisation ; par conséquent, les clusters *BX2* et *T-1* sont génétiquement identiques : ils sont constitués de sept copies du transgène *P[lacW]* et sont insérés au même site cytologique. Pourtant, bien qu'ils soient génétiquement identiques, les clusters *T-1* et *BX2* ne sont pas épigénétiquement identiques. En effet, dans les cellules germinales des ovaires de drosophile, le cluster *T-1* est capable de réprimer l'expression d'un transgène cible alors que *BX2* ne possède pas cette capacité (Ronsseray *et al*, 2001). Des analyses de small RNA-seq montrent que *T-1* produit des piRNA correspondant à la séquence du transgène *P[lacW]* alors que le cluster *BX2* n'en produit pas. Pourtant, lorsqu'on croise des femelles portant le cluster *T-1* par

des mâles portant le cluster *BX2*, on constate que le cluster *BX2* hérité du père dans la descendance possède des capacités répressives et produit des piRNA. Le cluster *BX2* est converti épigénétiquement dans la descendance par les piRNA produits par le cluster *T-1* hérités de la mère (de Vanssay *et al*, 2012). Ce phénomène de conversion épigénétique est appelé paramutation et le cluster *BX2* paramuté – c’est-à-dire possédant des capacités répressives associées à la production de piRNA – est désigné par *BX2\**. De façon cohérente, le phénomène de paramutation se produit lorsqu’on croise des femelles portant le cluster *BX2\** par des mâles portant le cluster *BX2*. Ainsi, le cluster *BX2* peut être paramuté à chaque génération, ce qui permet de maintenir la production de piRNA par le cluster au cours des générations (de Vanssay *et al*, 2012). La production de piRNA par *BX2\** est associée à un enrichissement de la marque H3K9me3 (Le Thomas *et al*, 2014). Ainsi, nous possédons le cluster *BX2* dans deux états épigénétiques différents : un état épigénétique ‘inactif’ où le cluster ne possède pas de capacités répressives, ne produit pas de piRNA et n’est pas enrichi en marques hétérochromatiques, on l’appelle alors *BX2<sup>OFF</sup>* (aussi appelé *BX2<sup>naif</sup>* dans certaines publications du laboratoire) (Figure 26C); et un état épigénétique ‘actif’ où le cluster possède des capacités répressives, produit des piRNA et est enrichi en marques hétérochromatiques, on l’appelle dans ce cas *BX2<sup>ON</sup>* (Figure 26D). Ces deux états épigénétiques peuvent être maintenus de manière extrêmement stable (plus de 150 générations) lorsque la température ne varie pas (25°C). Il est à noter qu’on n’observe pas de différence de quantité de transcrits entre les clusters *BX2<sup>OFF</sup>* et *BX2<sup>ON</sup>* ce qui confirme que la production de piRNA n’est pas corrélée à la quantité de transcrits produits par le cluster (de Vanssay *et al*, 2012). Ces résultats rappellent le cas des piRNA géniques, produits à partir des 3’UTR de gènes, pour lesquels la quantité de transcrits et la production de piRNA ne sont pas corrélées non plus (Robine *et al*, 2009).



**Figure 26 : Structure, localisation et propriétés du cluster *BX2*.** (A) Le transgène *P[lacW]* est dérivé de l'élément *P* et mesure 10691 pb ; il contient les pieds de *P* aux extrémités 5' et 3' ainsi que le gène *lacZ*, une partie du gène *lacY*, le gène *miniwhite* (construction ne contenant que les exons du gène *white*), le terminateur du gène *Hsp70Bb* et du plasmide *pBR322*. Les flèches représentent le sens de transcription des gènes *lacZ* et *miniwhite*. (B) Le cluster *BX2*, constitué de sept copies du transgène *P[lacW]* répétées en tandem, est inséré dans un intron du gène *Ago1* et en amont des gènes *mRpL53* et *CG3155* (C) *BX2<sup>OFF</sup>* ne possède pas de capacité répressive et ne produit pas de piRNA. A gauche, la photo représente des ovaires contenant *BX2<sup>OFF</sup>* et un transgène cible après coloration *lacZ*. Le transgène cible n'est pas réprimé, par conséquent le gène *lacZ* s'exprime et la  $\beta$ -galactosidase est produite. Au milieu et à droite, les diagrammes représentent respectivement l'alignement des petits ARN de 23 à 29 nucléotides sur la séquence du transgène *P[lacW]* et leur distribution par taille : *BX2<sup>OFF</sup>* ne produit pas de piRNA (D) *BX2<sup>ON</sup>* possède des capacités répressives et produit des piRNA. A gauche, la photo représente des ovaires contenant *BX2<sup>ON</sup>* et un transgène cible après coloration *lacZ*. Le transgène cible est réprimé, par conséquent le gène *lacZ* ne s'exprime pas et la  $\beta$ -galactosidase n'est pas produite. Au milieu et à droite, les

diagrammes représentent respectivement l'alignement des petits ARN de 23 à 29 nucléotides sur la séquence du transgène *P[lacW]* et leur distribution par taille : *BX2<sup>ON</sup>* produit des piRNA de 23 à 29 nucléotides qui sont s'alignent sur tous les transgènes *P[lacW]*.

### 3.2. Paramutation et effets transgénérationnels

Chez la drosophile, la paramutation du cluster *BX2* requiert l'héritage de piRNA maternels ainsi que les facteurs impliqués dans la biogénèse des piRNA (de Vanssay *et al*, 2012; Hermant *et al*, 2015). Ce phénomène de paramutation s'accompagne de modifications chromatinienne avec un enrichissement en marques H3K9me3 (Le Thomas *et al*, 2014). Les mécanismes moléculaires de la paramutation par les piRNA hérités maternellement restent encore à éclaircir. En 1956, Alexander Brink définit la paramutation comme une interaction entre deux allèles qui conduit à la modification de l'expression de l'un des allèles par l'autre allèle, sans modification de la séquence d'ADN. Cet état épigénétique est ensuite transmis aux générations suivantes en l'absence de l'allèle à l'origine du changement d'état (Brink, 1956).

La paramutation a d'abord été mise en évidence chez le maïs, notamment grâce à l'étude du locus *b1* qui code un facteur de transcription impliqué dans la synthèse de pigments violets et dont deux épi-allèles ont été identifiés, *B-I* et *B'*, dont les séquences ADN sont identiques. Lorsque la plante est homozygote pour *B-I*, le locus *b1* est fortement exprimé ce qui induit une pigmentation sombre au niveau des tiges. En revanche lorsque la plante est homozygote *B'*, le locus *b1* est peu exprimé ce qui induit une pigmentation claire des tiges. Lorsque ces deux homozygotes sont croisés, la pigmentation de la tige chez les descendants hétérozygotes *B-I/B'* est claire. Ces résultats indiquent que l'épi-allèle *B'* est capable de convertir *B-I*. L'épi-allèle nouvellement converti, désigné par *B''*, est à son tour capable de convertir *B-I* dans les générations suivantes (Coe, 1966). Une région contenant sept séquences répétées en tandem, requise pour le processus de paramutation, a été identifiée environ 100 kb en amont du locus *b1*. Le profil de méthylation de cette région répétée est différent selon l'épi-allèle ; de manière étonnante, la région répétée est plus méthylée dans le cas de l'épi-allèle *B-I* que pour *B'*, alors que *B-I* est associé à une expression forte du locus *b1* (Stam *et al*, 2002). La séquence répétée en amont du locus *b1* est transcrite dans les deux sens. Ces transcrits peuvent former des ARN dual-strand à l'origine de la production de siRNA de 25 nucléotides. En contexte mutant pour *mediator of paramutation 1 (mop1)*, un gène codant un ARN polymérase ARN-dépendant, ces siRNA ne sont plus produits. Pourtant, les siRNA ne sont pas suffisant pour induire la

paramutation puisqu'elle requiert également l'activité de *mop1* (Alleman *et al*, 2006; Arteaga-Vazquez *et al*, 2010). Cependant, le lien entre séquence répétée, siRNA, *mop1* et paramutation de *B-I* reste encore à préciser.

La paramutation a aussi été identifiée chez la souris avec le gène *Kit*, qui code un récepteur tyrosine kinase, entre autres impliqué dans la melanogénèse. Les souris homozygotes sauvages *Kit+* sont viables et le phénotype associé est une coloration blanche sur le bout des pattes et de la queue tandis que des souris homozygotes portant l'allèle *Kit<sup>tm1Alf</sup>*, dans lequel l'insertion d'une cassette *lacZ-neo* conduit à la perte de fonction de *Kit*, ne sont pas viables. Les souris hétérozygotes *Kit+/ Kit<sup>tm1Alf</sup>* sont viables et associées à un phénotype d'extension de la coloration blanche au niveau des pattes et de la queue. Ce phénotype est maintenu à la génération suivante chez des souris homozygotes sauvages issues de parents hétérozygotes, où l'allèle *Kit+* et la descendance sont alors désignés *Kit\**. Des expériences de northern blot chez les souris *Kit+/ Kit<sup>tm1Alf</sup>* et chez les souris *Kit\** indiquent que la quantité d'ARN polyadénylé est divisée de moitié par rapport aux souris sauvages *Kit+/Kit+*. Ce phénotype moléculaire est associé à la surexpression du gène *Kit* chez les *Kit+/ Kit<sup>tm1Alf</sup>* et *Kit\**. De plus, la micro-injection d'ARN provenant de souris *Kit+/ Kit<sup>tm1Alf</sup>* dans des embryons *Kit+/ Kit+* au stade une-cellule reproduit le phénotype associé à celui des souris *Kit+/ Kit<sup>tm1Alf</sup>* et *Kit*, ce qui indique que la paramutation du locus *Kit* est ARN dépendant (Rassoulzadegan *et al*, 2006). Toutefois, mentionnons deux points ; premièrement, contrairement à la paramutation chez la drosophile ou le maïs, dans cette étude la séquence d'ADN des deux allèles n'est pas identique ; deuxièmement ce type de paramutation chez la souris n'est pas stable au cours des générations, contrairement à la paramutation de *BX2* chez la drosophile.

Ces différentes études réalisées chez différents types d'organismes montrent que la paramutation est un processus partagé dans le monde du vivant et dont les mécanismes moléculaires semblent, au moins en partie, dépendants de l'ARN. La paramutation fait partie des mécanismes moléculaires impliqués dans les phénomènes d'épigénétique transgénérationnelle.

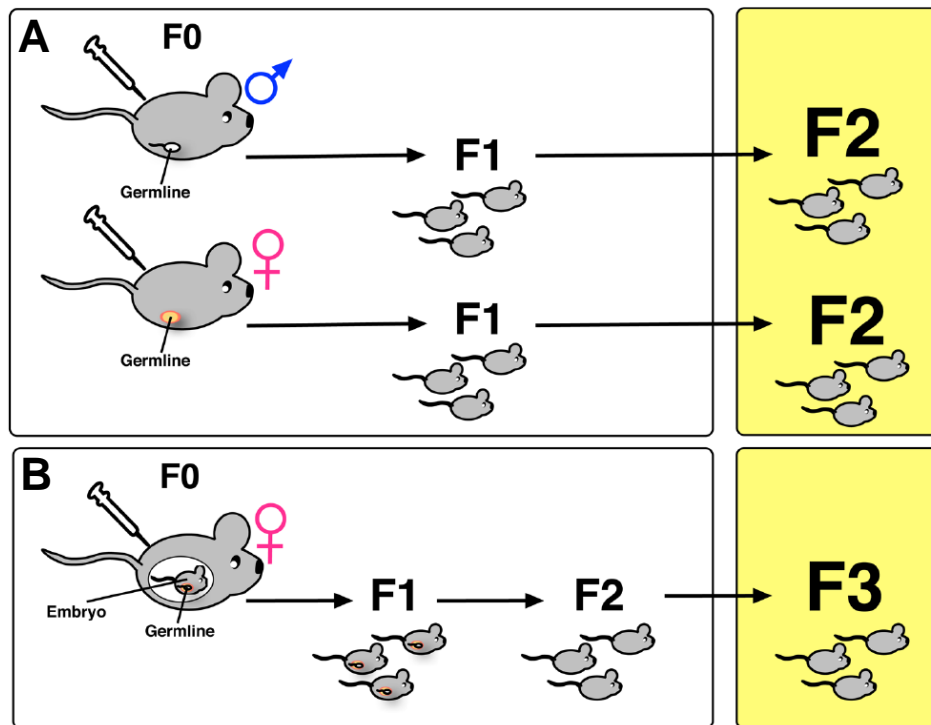
### **III. L'épigénétique transgénérationnelle et le rôle des piRNA**

#### **1. Description de la revue**

Depuis une vingtaine d'années, l'épigénétique transgénérationnelle est un domaine en pleine expansion, notamment grâce au développement des techniques de séquençage à haut-débit. L'épigénétique transgénérationnelle correspond à l'étude des modifications de l'épigénome qui persistent après plusieurs générations, et ce en l'absence de l'effecteur qui a induit ces modifications. De nombreux stress ou facteurs environnementaux ont été rapportés comme étant capables d'altérer durablement l'épigénome, notamment des stress chimiques, osmotiques ou physiques, des facteurs sociaux ou encore différents types de régimes alimentaires. Pour que les effets d'un stress soient maintenus à la génération suivante, ceci implique que le stress affecte les cellules germinales.

Il est important de noter qu'on parle d'effets transgénérationnels seulement si les individus ne sont pas directement exposés au stress qui induit ces effets. En effet, un stress n'affecte pas seulement l'individu - mâle ou femelle - qui y est exposé (G0), il peut également impacter les cellules de la lignée germinale de l'individu, qui constituent la génération suivante (G1) (Skinner, 2008) (Figure 27A). Par extension, chez une femelle gestante, le stress affecte la femelle gestante elle-même (G0), l'embryon (G1) mais encore les cellules germinales de l'embryon (G2) (Skinner, 2008) (Figure 27B). Par conséquent, les effets transgénérationnels ne peuvent être estimés qu'à la seconde génération (G2) ou, dans le cas particulier où le stress est induit chez une femelle gestante, à la troisième génération (G3) (Figure 27).





**Figure 27 : Estimation des effets transgénérationnels, quelle génération ?** (A) Lorsqu'un individu, mâle ou femelle, est soumis à un stress, la G0 et la G1 (lignée germinale de l'individu) sont affectées, ainsi les effets transgénérationnels ne peuvent être quantifiés qu'en G2. (B) Lorsqu'une femelle gestante est soumise à un stress, celui-ci impacte à la fois la G0, la G1 (embryon) mais aussi la G2 (lignée germinale de l'embryon), dans ce cas particulier, les effets transgénérationnels ne peuvent être estimés qu'en G3 (Yohn *et al*, 2015).

Les supports moléculaires qui ont été identifiés comme étant impliqués dans les effets transgénérationnels sont la méthylation de l'ADN, les modifications d'histones et les ARN non-codants, incluant les petits et les longs ARN non-codants (Heard & Martienssen, 2014; Lim & Brunet, 2013). Le rôle des piRNA dans la mémoire épigénétique a été rapporté chez la drosophile et chez le nématode (de Vanssay *et al*, 2012; Shirayama *et al*, 2012). Dans cette revue publiée dans *Cells*, nous abordons le rôle des piRNA dans les phénomènes d'épigénétique transgénérationnelle en réponse à des stress environnementaux, en discutant des mécanismes moléculaires impliqués ainsi que des conséquences possibles dans les domaines de l'évolution et du développement de maladies.

## 2. Review: Environmentally-Induced Transgenerational Epigenetic Inheritance: Implication of PIWI Interacting RNAs

Review

# Environmentally-Induced Transgenerational Epigenetic Inheritance: Implication of PIWI Interacting RNAs

Karine Casier , Antoine Boivin , Clément Carré  and Laure Teyssset \* 

Transgenerational Epigenetics & small RNA Biology, Sorbonne Université, CNRS, Laboratoire Biologie du Développement, Institut de Biologie Paris-Seine, UMR7622, 75005 Paris, France; karine.casier@upmc.fr (K.C.); antoine.boivin@sorbonne-universite.fr (A.B.); clement.carre@sorbonne-universite.fr (C.C.)

\* Correspondence: laure.teyssset@sorbonne-universite.fr

Received: 16 August 2019; Accepted: 17 September 2019; Published: 19 September 2019



**Abstract:** Environmentally-induced transgenerational epigenetic inheritance is an emerging field. The understanding of associated epigenetic mechanisms is currently in progress with open questions still remaining. In this review, we present an overview of the knowledge of environmentally-induced transgenerational inheritance and associated epigenetic mechanisms, mainly in animals. The second part focuses on the role of PIWI-interacting RNAs (piRNAs), a class of small RNAs involved in the maintenance of the germline genome, in epigenetic memory to put into perspective cases of environmentally-induced transgenerational inheritance involving piRNA production. Finally, the last part addresses how genomes are facing production of new piRNAs, and from a broader perspective, how this process might have consequences on evolution and on sporadic disease development.

**Keywords:** piRNA; transgenerational epigenetics; environment; transposable elements

## 1. Introduction

The impact of environmental stress on gene expression has been known since the 1960s. In the presence of lactose, glucose starvation in *Escherichia coli* induces derepression of the lactose operon, allowing cells to metabolize this sugar. These gene regulation changes are maintained until glucose is added to the nutritive medium without impacting subsequent generations [1].

In animals, plants, or fungi, some stress can induce epigenetic changes that can be maintained for at least more than one generation even if the stress is removed. Early studies analyzing human populations bring to light some physiological effects of parental food intake on their children's health, for instance, in Sweden [2], during the Leningrad siege (1941–1944) [3] or the Dutch famine (1944–1945) [4,5]. A restricted number of studies identified epigenetic modifications associated with parental stress [6]. Identification of accurate molecular associated modifications remains difficult to pinpoint, because of (1) the obvious impossibility of establishing a classical design of experiment, such as reproducibility and control experiments during several human generations, (2) past cultural and ecological factor influences [7], and (3) multiple stresses (nutritional, psychological, or social) that go along with wars and famines. Therefore, to better understand the consequences of specific stress at the epigenetic level and their possible transmission, it is critical to develop animal models.

Epigenetic inheritance is characterized by a reversible phenotypic transmission to the next generation in the absence of DNA sequence modification. This definition can be applied in an extended sense, including social learning and communication or in a tight sense, like mother to daughter cell transmission, called 'cellular epigenetic inheritance' [8]. In the latter, it is also important to differentiate intergenerational, multigenerational, and transgenerational inheritance. If the stress condition impacts

the F0 gestating female and the F1 progeny (embryos), it will be referred to as an intergenerational effect. If the F2 progeny (embryonic germ cells) are also affected, it will be referred to as a multigenerational effect. In contrast, transgenerational inheritance corresponds to an epigenetic transmission of stress passing through the germ cells, but without direct exposure. Hence, male and non-gestating female transgenerational phenotypes can be considered in F2, whereas for gestating female, only F3 has to be considered [9].

In the past twenty years, the development of next generation sequencing (NGS) approaches led to the accumulation of convincing evidence for environmentally-induced transgenerational epigenetic inheritance dependent on DNA methylation, specific histone modification, and non-coding RNA biology. However, the precise mechanisms remain unclear, leaving open questions to be tackled. PIWI-interacting RNAs (piRNAs), a class of small non-coding RNAs that tightly repress transposable element (TE) activity in animal gonads, have also been demonstrated to induce stable epigenetic memory. In this review, we present some examples of environmentally induced transgenerational epigenetic inheritance to underscore the role of piRNAs in epigenetic inheritance and more specifically, in transgenerational epigenetic inheritance in response to environmental stress. Finally, the implications of piRNAs in environmentally-induced transgenerational inheritance in evolution and potential implication in human health will be discussed.

## 2. Evidence and Mechanisms of Environmentally-Induced Transgenerational Epigenetic Inheritance

The universal aspect of transgenerational epigenetic inheritance is now well documented in several species from different taxa, including bacteria, protists, fungi, plants, and animals [8]. About one-third of these transgenerational effects are linked to environmental factors, such as nutritional stresses (starvation or supplementation), chemical stresses (arsenic), and physical stresses (temperature, pressure, light and dark exposure).

### 2.1. Multiple Molecular Supports for Environmentally-Induced Epigenetic Inheritance

Four main types of support have been identified including (1) DNA that can be methylated on CG, CH, and CHH motifs, where H corresponds to a nucleotide other than G, (2) histones that can be modified (trimethylation of histone 3 lysine 4 (H3K4me3) associated with transcriptionally active marks and di/trimethylation of histone 3 lysine 9 (H3K9me2/3) and trimethylation of histone 3 lysine 27 (H3K27me3) associated with transcriptionally repressive marks) or replaced in the late-stage spermatids of many animals by the protamines forming a highly compact genome [10]. Defect of histone replacement during gametogenesis, called 'histone retention', could be associated to some infertility in men [11], (3) small non-coding RNAs (microRNAs (miRNAs), short interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs), and transfer RNA-derived small RNAs (tsRNAs) (also known as tRNAs Fragments (tRFs)), and (4) long non-coding RNAs (lncRNAs) [12–14].

Two additional supports for epigenetic inheritance are possible, although evidence of environmentally-induced epigenetic inheritance is still lacking. First, protein structures involving prions or the Hsp90 chaperone protein have both been linked to chromatin remodeling factors in *Drosophila* and could be good candidates for involvement in such mechanisms (for more details, see [15,16]). Second, over the years, more than 100 RNA modifications have been described, some linked to human diseases [17,18]. Among the enzymes involved in these modifications, the m<sup>5</sup>C Dnmt2 RNA methyltransferase has been shown to be required in mice for epigenetic transmission of Kit and Sox9 phenotypic variants [19].

Although evidence is still lacking concerning protein structure and RNA modifications, transgenerational environmentally-induced epigenetic inheritance associated with DNA methylation, histone modifications, or non-coding RNAs has been reported. Some examples are presented in Table 1, which highlights that one stress can induce multiple epigenetic modifications. For instance, in rat sperm, vinclozolin or DDT pesticide exposure modifies DNA methylation, drives histone retention,

and decreases the global number of expressed small non-coding RNAs (miRNA, piRNAs, and tsRNAs) and long non-coding RNAs, affecting expression of genes implicated in metabolism, signaling, and transcription [20,21] and is correlated with the occurrence of several pathologies [22–25]. In *Drosophila*, the elevation of temperature during development causes functional piRNA emergence associated with H3K9me3 enrichment from a particular repeated locus mimicking a piRNA cluster for over 50 generations [26] (Table 1). Such cross-talk between several epigenetic modifications is not restricted to animals. In fission yeast, the RNA interference (RNAi) machinery is necessary to direct H3K9me3 mark deposition, inducing heterochromatin formation and silencing on a centromeric transgene. This new epigenetic state, sensitive to the H3K9 methyltransferase mutant, can then be maintained up to 32 generations [27]. Taken together, H3K9me3 marks and small non-coding RNA could act together to reinforce epigenetic stability through generations thanks to a feedback loop: on the one hand, H3K9me3 leads to heterochromatin formation and recruits the RNAi machinery, and on the other hand, the RNAi machinery targets H3K9 residues and induces their trimethylation [26,27].

**Table 1.** A list of environmentally-induced transgenerational epigenetic inheritance.

Stress	Model/tested tissues	Epigenetic modifications	Reference
<b>Pesticides and pollutants</b>			
Vinclozolin	Rat/Sperm	DNA methylation	[28,29]
		miRNA, piRNA, tsRNA	[30]
	Rat/Prostate cells	DNA methylation, histone retention, miRNA, piRNA, tsRNA, and lncRNA	[20]
		DNA methylation, miRNA, piRNA, tsRNA, and lncRNA	[23]
		miRNA	[31]
		DNA methylation, piRNA	[25]
Dichlorodiphenyltrichloroethane (DDT)	Rat/Sperm, testis, ovaries, kidney, prostate, whole organism	DNA methylation	[32]
	Rat/Sperm	DNA methylation	[33]
		DNA methylation, histone retention, miRNA, piRNA, tsRNA, and lncRNA	[21]
	Rat/Ovaries	DNA methylation, piRNA	[25]
	Rat/ovaries, sperm, prostate, kidney, whole organisms	DNA methylation	[22]
	Rat/Ovaries	DNA methylation	[34]
Pesticide mixture (Permethrin and DEET), plastic mixture (bisphenol A and phthalates), dioxin and jet fuel hydrocarbon	Rat/Sperm, Ovaries	DNA methylation	[35]
	Rat/Sperm, testis, prostate, kidney, ovaries, whole organism	DNA methylation	[36]
	Rat/Sperm, testis, prostate, kidney, ovaries, whole organism	DNA methylation	[24]
Chlordecone Hydrocarbon (jet fuel JP8)	Mouse/Testis	H3K4me3 modification	[37]
	Rat/Sperm, ovaries, kidney, prostate, whole organism	DNA methylation	[38]
Bisphenol A	Nematode/Germinal cells	H3K27me3 and H3K9me3 modifications	[39]
Heavy metals (Cu, Cd, Cr, and Hg)	Rice/Leaf	DNA methylation	[40]
<b>Traumatic stresses</b>			
Maternal separation	Mouse/Sperm, brain	DNA methylation	[41]
Fear conditioning	Mouse/Sperm	DNA methylation	[42]
<b>Diet</b>			
Low protein diet	Rat/Liver	DNA methylation	[43]
Starvation	Nematode/Whole organism	siRNA	[44]
Feeding with bacteria expressing ds-RNA (GFP)	Nematode/Germinal Cells	siRNA	[45]
Avoidance of pathogenic bacteria	Nematode/Whole organism	piRNA	[46]
Alcohol	Rat/Sperm, POMC neurons	DNA methylation	[47]
Nitrogen deficiency	Rice/Leaf	DNA methylation	[48]
Drought	Rice/Seed	DNA methylation	[49]
<b>Osmotic and Thermic stresses</b>			
NaCl	Daphnia/Whole organism	DNA methylation	[50]
NaCl, Heat shock	<i>Drosophila</i> /Eyes	Heterochromatin disruption	[51]
High temperature (25 °C)	Nematode/Whole organism	H3K9me3 modification	[52]
	Nematode/Whole organism	H3K9me3 modification and siRNA	[53]
High temperature (29 °C)	Nematode/Oocytes	siRNA	[54]
	<i>Drosophila</i> /Ovaries	H3K9me3 modification and piRNA	[26]

## 2.2. How Long Does the Imprinting of Transgenerational Epigenetic Inheritance Last?

At each generation, the totipotency of primordial germ cells is ensured by the erasure of epigenetic modifications requiring numerous genetic factors, such as DNA demethylation, histone methyltransferases, and histone demethylases [55–57]. However, analyses of DNA methylation patterns show that some loci escape this resetting, suggesting that antagonist mechanisms provide maintenance of some epigenetic marks [58]. In fact, the DNA methyltransferases *Dnmt3a* and *Dnmt3b*, the PRC2 complex in mice and the PRC2 complex associated with H3K27me3 marks in worms have been shown to be involved in the maintenance of DNA methylation and histone modification pattern on discrete loci [59–61].

Some studies have addressed this question for small non-coding RNAs. In *Caenorhabditis elegans*, the RNAi machinery provides its own regulation by targeting RNAi genes and leads to a progressive decrease of small RNA at each generation, varying from one to four generations [62,63]. Antagonist molecular mechanisms govern the duration of the small RNA transgenerational response. As for other modifications, transgenerational small RNA inheritance relies on many genes coding for RNA-dependent RNA polymerase (RdRp), Argonaute proteins, or histone methyltransferases (Reviewed in [64–66]). The duration of the transgenerational response can be extended due to siRNA, which constitutes a second RNAi trigger that drives tertiary siRNA production [67]. Hence, the duration of epigenetic inheritance depends on the influence of these two opposite processes, namely resetting and stimulation of small RNA production [68,69]. Disruption of the endo-siRNA pathway leads to hypersensitivity to exogenous small RNAs, suggesting competition between these RNAi pathways, probably for molecular factors, such as DCR-1, which is the unique *Dicer* ortholog in *C. elegans* [70]. In contrast, piRNA producing loci transcription is actively maintained at each generation by maternal piRNA inheritance in *Drosophila* [71,72]. Disruption of this maternal deposition is followed by a complete erasure after only one generation [26,71].

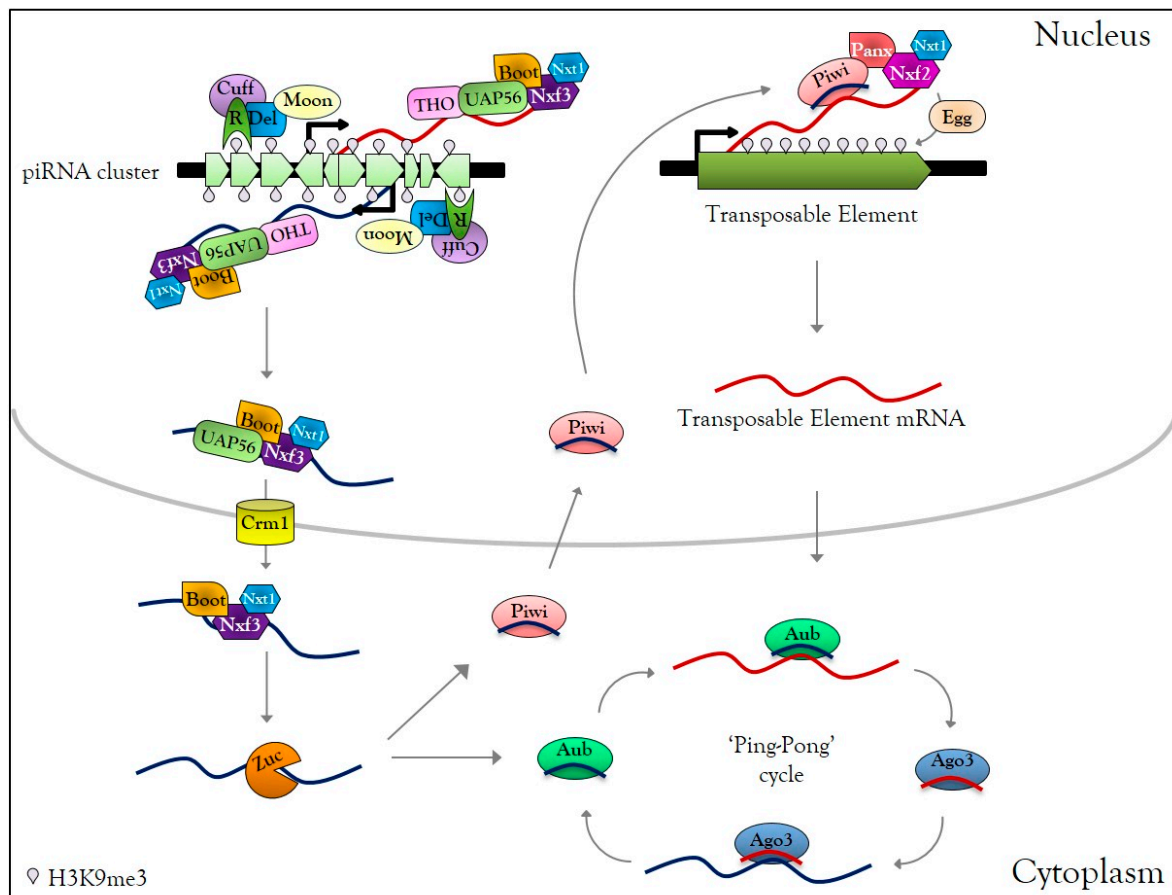
The family of small non-coding RNAs is frequently considered in transgenerational epigenetic inheritance, probably because there are responsible for endogenous gene regulation and involved in important developmental programs (Reviewed in [73]). In the particular case of piRNAs, their function in epigenetic memory associated with their role in response to environmental factors has been recently demonstrated [26,46].

## 3. piRNAs in Environmentally-Induced Transgenerational Epigenetic Inheritance

### 3.1. Overview of piRNA Biogenesis and Functions

The piRNAs correspond to a class of small non-coding RNAs interacting with the PIWI proteins (Piwi, Aub, and Ago3 in *Drosophila*), belonging to the AGO subfamily proteins, mainly active in animal gonads to protect genome integrity against TEs. In *Drosophila* germ cells, piRNAs are produced from specific heterochromatic loci, composed of numerous TE fragments, called piRNA clusters [74,75]. Those clusters are enriched in the H3K9me3 mark and Rhino HP1-like protein [76] (Figure 1). Rhino interacts first with the co-factors Deadlock and Cutoff, then Deadlock recruits the transcriptional initiation machinery via the large subunit of transcription initiation factor II A (TFIIA-L) paralog called Moonshiner, as well as Bootlegger, UAP56, the THO complex, and nuclear export factors Nxt1 and Nxf3 [76–79]. The piRNA precursor transcripts emerging from piRNA clusters are then exported by the Crm1 (chromosomal maintenance 1) exportin to the cytoplasmic piRNA biogenesis site, where they undergo a single strand hydrolysis by the Zucchini RNase to finally give rise to piRNAs [78–83] (Figure 1). In contrast to si- and miRNA, piRNAs are charged directly as single guide RNA onto PIWI proteins to form piRNA-induced silencing complexes (piRISC) [74]. After nuclear translocation, piRISC complexes recognize nascent TE transcripts and perform transcriptional gene silencing (TGS) by local heterochromatinization of TEs. This silencing requires nuclear Piwi-piRNAs that direct Panoramix, Nxf2, its co-factor, Nxt1, and the histone methyltransferase Eggless to the TE loci [84–91] (Figure 1). In the cytoplasm, mRNAs of TEs are cleaved by piRISC complexes, leading to post-transcriptional

gene silencing (PTGS) of mobile elements. Cleaved-transcripts are not degraded; they are charged by PIWI protein to form new piRISC complexes, corresponding to an amplification mechanism called the 'Ping-Pong cycle' [74] (Figure 1). piRNA-dependent silencing occurs at each gonadal developmental stage [92], and disruption of the piRNA pathway leads to TE mobilization [93], which induces DNA damage [94,95].



**Figure 1.** Germline PIWI-interacting RNAs (piRNA) biogenesis. In *Drosophila melanogaster* germ cells, active piRNA clusters are enriched in H3K9me3 and HP1-like protein Rhino (R), which interact with Deadlock (Del), Cutoff (Cuff). Moonshiner (Moon) interacts with Del and recruits transcription initiation factors, allowing transcription of piRNA clusters in the sense and antisense direction. Bootlegger, UAP56, the THO complex, Nxt1, Nxf3 are recruited to the piRNA cluster transcription site. The transcripts are then exported to the cytoplasmic piRNAs biogenesis site via the CRM1 exportin. In the cytoplasm, piRNA precursors are cleaved by Zucchini (Zuc) into primary piRNA and charged by Piwi or Aubergine (Aub) proteins, forming Piwi-piRISC and Aub-piRISC complexes. Piwi-piRISC is translocated into the nucleus and recognizes nascent TE transcripts due to the interaction with Panoramix, the Nxf2, and Nxt1. The association of these factors to nascent TE transcripts causes recruitment of the histone methyltransferase Eggless that transcriptionally silences TEs by local heterochromatinization. In the cytoplasm, Aub-piRISC cleaves TE transcripts into secondary piRNAs, which are loaded onto Ago3, forming Ago3-piRISC complexes. Ago3-piRISC recognizes and cleaves piRNA precursors into new secondary piRNAs, loaded onto Aub. This mechanism, known as the 'Ping-Pong' cycle, allows the amplification of piRISC complexes and post-transcriptional regulation of TEs.

PIWI protein distributions reveal that the piRNA pathway is highly conserved in animals, but they exist neither in plants nor in prokaryotes [96]. Interestingly, recent studies show that piRNAs are also found in somatic tissues of mammalian, arthropods, and mollusks [97–100]. In *Drosophila*

*melanogaster*, however, piRNA expression is restricted to the gonads, including the germline nurse cells and the somatic follicular cells [101,102]. These findings might suggest that the ancestral capacity to synthesize piRNAs in somatic tissues (outside of gonads) has been lost in *D. melanogaster*. However, this is still subject to debate [103–105].

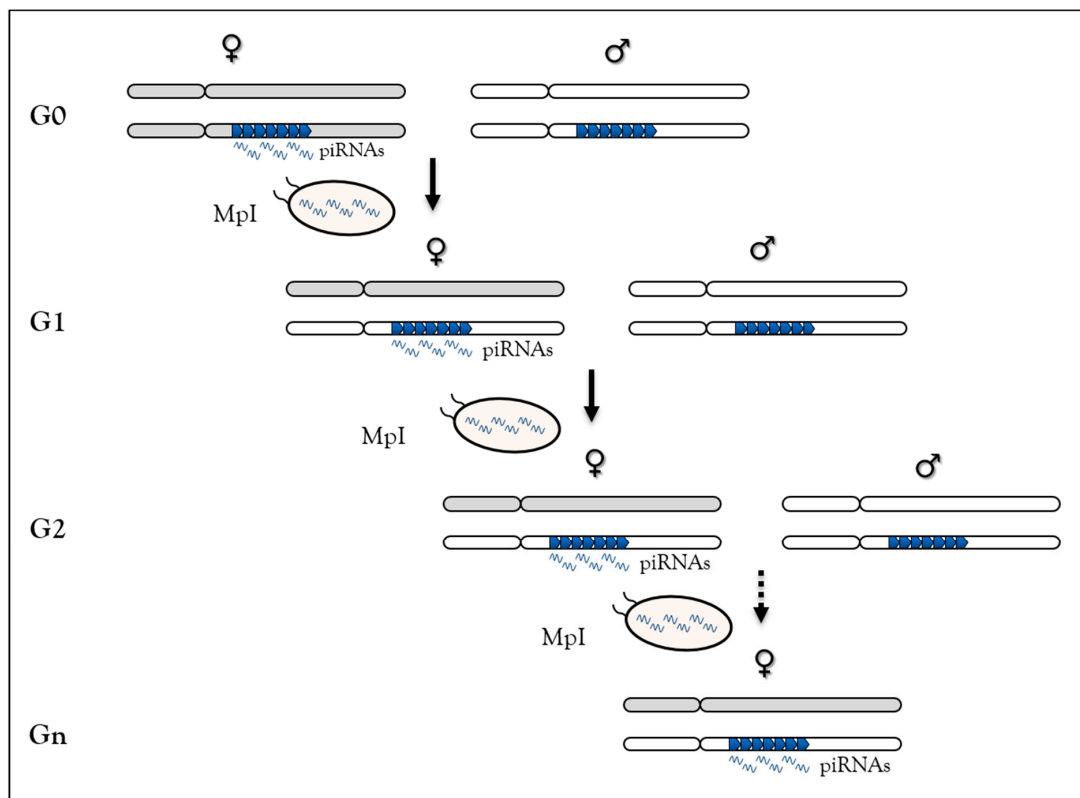
### 3.2. Role of piRNAs Outside of TE Regulation

Genomic annotations of small RNA libraries isolated from various metazoans reveal that piRNAs are not restricted to TE regulation but can also target coding genes. In *Drosophila*, piRNAs derived from the 3'UTR of *traffic-jam* (*tj*) matching *Fasciclin III* (*FasIII*) can be identified in the ovarian somatic cells (OSC) line. This finding suggests that *FasIII* is regulated by piRNAs originated from the *tj* 3'UTR [102,106]. In *Drosophila* early embryos, *nanos* (*nos*) is regulated by piRNAs. Evidence suggests that maternal *nos* mRNAs are targeted by a complex composed of piRNAs, Aub, and CCR4-NOT deadenylation factors in the soma, whereas in the germ plasm, those transcripts are stabilized by the interaction with Aub [107,108]. In *Aplysia*, piRNAs seem to be implicated in *CREB2* promoter methylation, a transcriptional repressor involved in long-term memory [109]. As the last example, sex determination in silkworm is dependent on the piRNA pathway that regulates *Masculinizer*, a locus required for the production of the *doublesex* female-specific isoform [110]. Taken together, these data highlight the idea that piRNAs are required in gene regulation of a broad range of cellular functions.

### 3.3. piRNAs in Epigenetic Memory

In *Drosophila*, the insertion of exogenous *lacZ* sequences into a subtelomeric piRNA cluster can functionally repress a euchromatic *lacZ* transgene in *trans* [111]. This epigenetic silencing mechanism, referred to as the *trans*-silencing effect (TSE), requires maternally but not paternally inherited piRNAs [71,112–116]. These results strongly suggest an active role for piRNAs in TSE and show that a piRNA cluster, without maternal piRNA inheritance, is not sufficient to promote this silencing process [112]. Involvement of piRNAs in epigenetic memory has been associated with epiallele emergence through a paramutation phenomenon in *D. melanogaster* and in *C. elegans* [71,117]. Paramutation is defined as a modification of one allele induced by another allele without DNA sequence modification and corresponds to an epigenetic conversion (for reviews see [118–120]).

In *Drosophila*, using a locus made of repeated *P-lacZ-white* transgenes mimicking a piRNA cluster (called *BX2* cluster), but inactive for piRNA synthesis, we have shown that homologous maternally-inherited piRNAs can stably convert this locus into an active piRNA cluster for over 200 generations [71] (Figure 2). This newly-activated cluster has the same properties as other *Drosophila* piRNA clusters. This piRNA-induced paramutation process requires molecular actors of the piRNA pathway [72] and performs TE-derived sequence control generation after generation. Finally, the reversibility of paramutation occurs when maternal piRNA inheritance is interrupted, underscoring the idea that this is an epigenetic process [26,71]. Taken together, based on sequence complementarity, these data show that maternally-inherited piRNAs are competent to initiate new piRNA production in the next generation and suggest that piRNAs are responsible for their own maintenance across generations. Two non-exclusive molecular mechanisms have been proposed to explain this case of epigenetic memory in which piRNA production requires piRNA inheritance [121]. In the first, inherited-piRNAs would induce heterochromatin formation on the piRNA cluster by leading to H3K9me3 deposition, inducing local Rhino enrichment [26]. The piRNA cluster-specific transcription machinery will then assemble at the locus to promote piRNA synthesis [77]. In the second, inherited piRNAs could trigger their production by Ping-Pong amplification from piRNA cluster precursors and TE transcripts [74,75].



**Figure 2.** The paramutation process in *Drosophila melanogaster* relies on maternal piRNA inheritance. At G0, females (grey chromosomes) carrying the BX2 clusters made of repeated *P-lacZ-white* transgenes (blue arrowheads) producing piRNAs are crossed to males (white chromosomes) carrying the same cluster not producing piRNA. At G1, maternal piRNA deposition in the embryo leads to activation of piRNA production from the paternal cluster allele. The newly activated cluster can activate the paternally-inherited inactive cluster in G2 due to maternal piRNA inheritance (MpI). Then, this process of activation leads to stable piRNAs production across generation.

In *C. elegans*, piRNAs correspond to 21 nucleotide (nt) long small RNAs with a uridine bias at the 5'-end (21U RNA) distinct from the endo-siRNAs that are 22 nt long small RNAs with a guanosine bias at the 5'-end (22G RNA) [122]. The biogenesis of nematode piRNAs depends on the Piwi protein, PRG-1 [66,123,124]. Like in *Drosophila*, transgenes are used to decipher the molecular mechanisms of germline repression. Analysis of the piRNA sensor indicates that piRNAs are required to trigger silencing through imperfect base-pairing [125–127]. This silencing is then transgenerationally propagated by the 22G RNA dependent on numerous factors, such as RdRp, NRDE-2, the Argonaute proteins WAGO-9/HRDE-1 and WAGO-4 and chromatin components, such as the histone methyltransferases SET-25, SET32, and HPL-2, a Heterochromatin 1 Protein (HP1) orthologue [45,128–133]. Indeed, specific H3K9 methyltransferases, SET-25 and SET-32, and endo-siRNAs have been shown to work together to promote silencing inheritance [134]. In addition, neuronal RDE-4 and germline HRDE-1 have been implicated in transgenerational regulation of endo-siRNAs and mRNAs controlling chemotaxis [135]. The authors proposed the existence of an intricate interaction between the neuronal system and the germline to control transgenerational behavior. Finally, precise dissection of the silencing mechanism determines that piRNAs lead to the production of secondary and tertiary siRNAs, necessary for full repression and leading to an epigenetic conversion form of paramutation [67].

Altogether, the piRNA silencing pathway in *Drosophila* involves piRNAs for the initiation and maintenance through generations [26,71,72], while in worms piRNAs are involved in triggering the silencing that is then maintained by siRNAs in the progeny, independent from piRNAs [124,136–138].



### 3.4. Heritable Stress-Induced piRNA Synthesis

In F0 mice, an early maternal separation causes downregulation of piRNA cluster 110 [139], and a Western-like diet leads to differential expression of 190 piRNAs distributed on 63 piRNA clusters [140]. In rat, a high-fat diet induces deregulation of 1092 piRNAs, and only three of them are still differentially expressed at the next generation [141]. In *C. elegans*, high temperature (25 °C) induces motif-dependent piRNA biogenesis downregulation [142]. This downregulation is associated with differential gene expression and fitness reduction, which persist at the next generation. However, in those studies, piRNA expression in subsequent generations, as well as their roles in a phenotypic response and the potential consequences on the epigenome, was not directly addressed. In the case of rat stressed with vinclozolin or DDT, differentially expressed piRNAs were identified in F3, but the mechanism and role of piRNAs were not questioned [20–23,25] (Table 1).

This year, two cases of environmental stresses inducing a genomic response dependent on piRNAs have been published in *C. elegans* and in *D. melanogaster* [26,46].

#### 3.4.1. Behavior in *C. elegans*

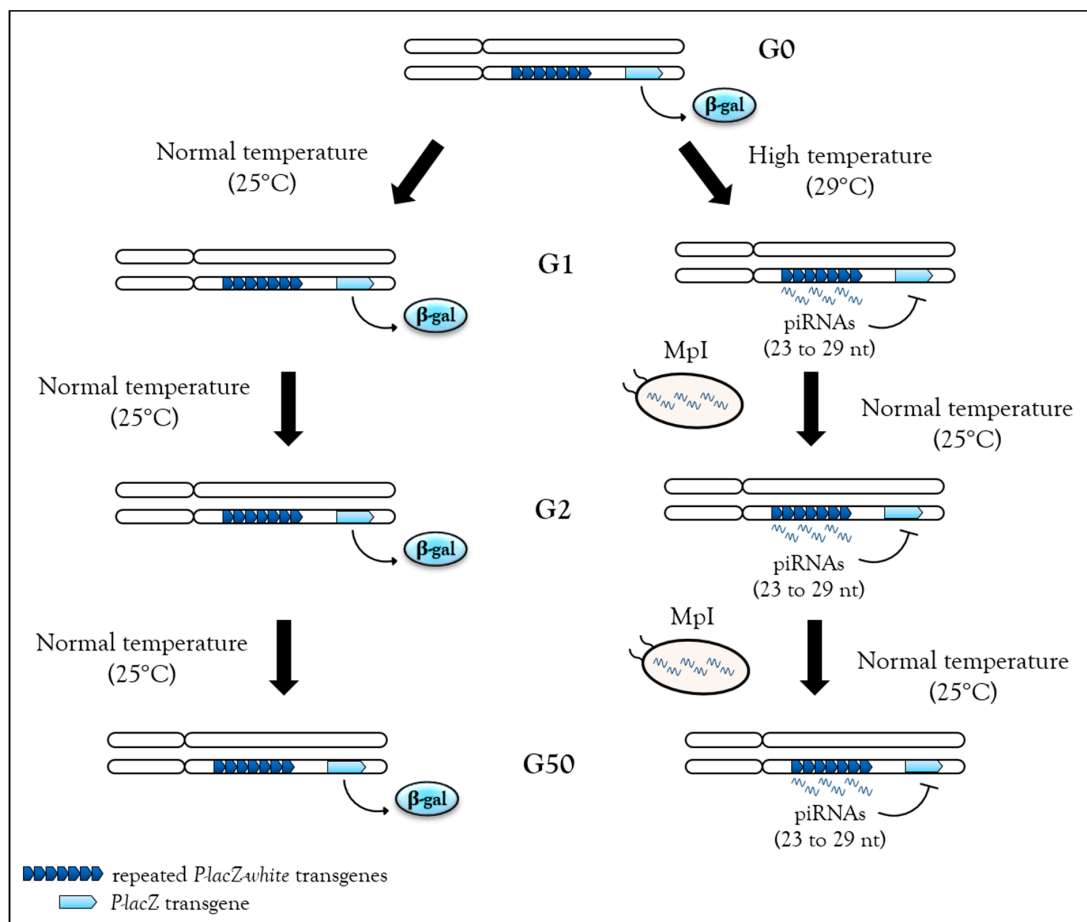
Wild type male or female larvae of *C. elegans* exposed to pathogenic *Pseudomonas aeruginosa* learn to avoid it and switch their feeding preference to nonpathogenic bacteria. This learning to distinguish between pathogen and non-pathogen does not correspond to a general phenomenon but rather is specific for *P. aeruginosa* [46]. Surprisingly, this behavior of avoidance can be transgenerationally inherited to the naïve F1 progeny up to the fourth generation (F4). This peculiar inheritance is completed at the fifth generation, where progeny exhibits a naïve behavior. Differential transcriptome analyses were performed on treated vs. non-treated mothers fed with pathogenic *P. aeruginosa*s and their respective progeny. Moore and colleagues showed that the TGF- $\beta$  ligand DAF-7, previously identified as a neuronal responder of the pathogen [143], was expressed more in specific sensory neurons of treated mothers relative to non-treated animals as well as in the F1, F2, F3, and F4 progeny, thus mirroring transgenerational avoidance [46]. The authors further show that in the *daf-7* loss of function mutant, while the mother can learn to avoid the pathogen, the F1 progeny did not receive the information of avoidance. The same pattern of inheritance was observed for mother defective for the Piwi encoding gene *prg-1*. This mutation is associated with a defect in piRNA expression in the mothers and in *daf-7* induction in the ASI neurons in F1.

Altogether, the studies indicate that transgenerational inheritance involved in the avoidance of pathogenic *P. aeruginosa* depends on the TGF- $\beta$  pathway linked to the small non-coding RNAs as well as piRNA molecular partners. Therefore, piRNAs might participate in the control of specific reversible behaviors, which might have consequences on adaptive survival, providing advantages to progeny. In addition, the loss of inheritance after five generations might be considered an advantage in worms that may colonize a new environment in their natural habitat. The precise molecular basis of this resetting occurring at the fifth generation was not addressed but might be reminiscent with previous studies [63].

#### 3.4.2. Heat Response Induces de novo piRNAs

Recently, high temperature during development (29 °C instead of 25 °C) has been shown to be sufficient to induce de novo piRNA production and heterochromatin associated with H3K9me3 and Rhino enrichment from the BX2 cluster in *Drosophila* [26]. One generation raised at 29 °C is sufficient to activate piRNA production from the BX2 cluster, and once this production is started, it can be maintained over 50 generations even after return to normal temperature (25 °C) (Figure 3). These piRNAs can functionally silence homologous *lacZ* reporter transgenes, a phenomenon referred to previously as TSE. Hence, heat stress is able to induce piRNA production from a locus comprised of repeats. Once established, this production remains stable even after stress removal, i.e., when flies return to a normal temperature of 25 °C. Small RNA sequencing analyses were unable to identify

other genomic loci used for de novo piRNA synthesis as if all loci competent to become piRNA cluster were already activated. Analyses performed at 29 °C to identify molecular factors required for the activation of de novo piRNA production suggest that there is an increase of transcription of the *BX2* cluster flanking region along with the transcription of a homologous euchromatic sequence. When *BX2* is maternally inherited, piRNA production is maintained, but when *BX2* is paternally inherited, piRNA production is abolished. This result suggests that when the stress is removed, piRNA can be self-maintained over generations by the paramutation process, as long as the inheritance is maternal [26] (Table 1).



**Figure 3.** Stable, heritable stress-induced de novo piRNAs. In flies, the *BX2* cluster, a sequence made of repeated *P-lacZ-white* transgenes (dark blue arrowheads), is capable of producing piRNA after one generation at 29 °C, then these piRNAs can functionally repress a homologous *P-lacZ* sequence (light blue arrowheads) in *trans*. This silencing capacity is maintained at the second generation after return to normal temperature (25 °C) and then up to 50 generations. Hence, when stress is removed, piRNAs can be self-maintained by maternal piRNA inheritance (MpI) at each generation.

In *Drosophila*, environmental stresses can generate piRNA emergence, and this new piRNA production is maintained across generations, allowing stable repression of homologous sequences. How can environmental stress impact piRNA production? It remains unclear whether piRNA cluster chromatin containing H3K9me3 marks, piRISC complex formation and/or stability, or piRNA pathway genes are involved in the stress response, individually or simultaneously. To date, it is quite difficult to estimate the global effect of piRNA deregulation over many generations, in particular, if it might result in the alteration of inherited gene expression patterns. Future investigations will help our understanding of environmentally-induced piRNA transgenerational inheritance and allow the estimation of potential consequences.

## 4. Speculative Insights on Transgenerational piRNA Inheritance Consequences

### 4.1. Implications in Evolution

According to Lamarck's theory, the environment and the experiences of the organism mediate changes that could be inherited in subsequent generations. These changes are then subject to natural selection and adaptation [144,145]. In the case of transgenerational inheritance, the environment can impact the epigenome by causing epimutations, through DNA methylation, histone modifications, and non-coding RNA misregulation. These epimutations lead to phenotypic variations and can give the organism a selective advantage. In *C. elegans*, it has been shown that stress conditions (starvation, supplemented diet, salt, or arsenic) enhance longevity and stress resistance in subsequent generations [146–148]. If they are stable enough to be maintained over time, epigenetic modifications and their gradual accumulation in response to various stresses, may have an important role in speciation [149]. For Brevik and colleagues, transgenerational epigenetic inheritance may also explain the rapid emergence of insect population resistance against pesticides [150]. Many studies have shown that pesticides, notably vinclozolin or DDT, induce transgenerational changes in DNA methylation, histone modification, and/or non-coding RNA deregulation in rodent models (Table 1).

Among the epigenetic mechanisms driving adaptation and evolution, piRNAs are playing a fundamental role in TE regulation across generations. For instance, during the 20th century, the *D. melanogaster* genome was invaded by the *P* element after a horizontal transfer [151]. It appears that in the absence of *P* element regulation, progeny display genetic defects, leading to gonad atrophy at high temperature, whereas no anomaly occurs when *P* element regulation is established [152,153]. It was later shown that this specific control of gonadal *P* element occurs through piRNAs [101]. Interestingly, the *Drosophila simulans* genome has been recently invaded by *P* elements [154], and the invasion dynamics are widely influenced by temperature. The transposition rate at elevated temperature is higher compared to the rate at low temperature [155]. At elevated temperature (29 °C), after 20 generations, the *P* element copy number in the genome reaches a plateau. These invasions are correlated with piRNA production against *P* element sequences, strongly suggesting that the regulation is established in response to invasion.

Analyses of *P* element distribution in the *D. melanogaster* genome showed that subtelomeric regions of the X chromosome are hotspots for *P* element insertions [156]. Those subtelomeric regions happened to be one of the active *Drosophila* piRNA clusters [74]. Once *P* element or *P* derived transgenes insert into subtelomeres, new piRNAs targeting those sequences will be synthesized, leading to repression of euchromatic copies [71,101,114,116,157]. Particular sequences in subtelomeric piRNA clusters, called *TAS-L like (TLL)*, found in the X chromosome of *D. melanogaster*, are also present in autosomic subtelomeric piRNA clusters of other members of the *melanogaster* subgroup, including *D. simulans*. The finding that some *P* element insertions were found very close to *TLL* in *D. simulans* suggests that *P* element regulation could be conserved between *D. melanogaster* and *D. simulans* [155,157]. As with all piRNA clusters, the subtelomeric regions are made up of repeated domains that were once activated for piRNA production. However, the type of event that converted these loci remains unclear. Recently, we have shown that high temperature can convert a locus composed of tandemly repeated sequences into a stable piRNA cluster, highlighting the sensitivity of such regions to environment stresses for piRNA production [26]. Once activated, de novo piRNA production can be transmitted through generations by maternal inheritance [71]. Because these loci are hotspots of insertion for TEs, specific piRNAs are synthesized that functionally repress homologous sequences, leading to maintenance of genome integrity over time. Therefore, piRNA clusters constitute an epigenetic memory of TE invasion and are essential to provide genome defense.

TEs contain many regulatory sequences, including insulators, RNA polymerase signals, splice sites, and/or environment response elements [158]. Furthermore, TEs can modulate the expression of neighboring genes [159]. In addition, numerous stresses, including radiation, pollutants, temperature, or viral infection, can induce TE activation [160,161]. Thus, similar to *P* element activity under

elevated temperature conditions, the emergence of piRNA production can occur in response to stress-induced TE activation. We can expect that organisms with TE regulatory mechanisms are likely to be selected as compared to organisms without this defense system. Indeed, TE mobilization induced by environmental stress can affect genome organization by creating mutations, chromosomal breaks, or genomic rearrangements [162,163]. In the long-term, these modifications can be beneficial for organisms by the creation of new molecular functions, but until this occurs, TEs must be regulated in the short-term to restrict genomic damage. In a way, if TEs are defined as molecular actors of evolution, piRNAs could act as molecular regulators of evolution.

Finally, piRNAs and molecular factors of the piRNA pathway have been described to play a role in reproductive isolation by maintaining interspecies sterility [164,165]. In *D. melanogaster* and *D. simulans* hybrids, *D. simulans* Rhino protein co-localizes with *D. melanogaster* Deadlock, but they cannot interact, because of the rapid evolution of their chromo domains. As a consequence, TE mobilization leads to hybrid sterility due to the absence of a functional piRNA pathway [164]. Additionally, the *AT-chX* locus, located close to the pericentromeric region of the X chromosome, is one of the *D. melanogaster* piRNA clusters absent from other *Drosophila* species. This locus contains sequences related to the *vasa* gene, which is required for germ cell development. Due to imperfect complementarity, piRNAs derived from the *AT-chX* locus are not able to repress *vasa* expression in *D. melanogaster* germ cells. Surprisingly, the *AT-chX* locus shares a high level of complementarity with the *vasa* gene from closely related species to *D. melanogaster*, leading to a repression of *vasa* in interspecies hybrid testes [165]. These studies suggest that piRNAs could play an important role in speciation.

#### 4.2. Potential Implications for Disease Development

Environmental stresses, such as pesticide and pollutant exposure, have been shown to drive epigenetic inheritance of obesity and many associated diseases, including diabetes, infertility, immune disorders, and polycystic ovarian syndrome in rat. These transgenerational effects are associated with the emergence of differential DNA methylation regions, which contain genes involved in several processes, such as signaling, metabolism, transcription, development, and transport [23–25,32,35,36,38,166–168].

##### 4.2.1. piRNA Pathway and Disease

To date, there is no evidence for piRNA-mediated transgenerational inheritance of disease risks. However, paramutation and small RNA inheritance could represent a part of ‘missing heritability’ in some rare diseases that genetic mutations cannot explain (discussed in [169,170]). Deregulation of piRNAs and PIWI proteins has been found in some diseases, including many types of cancer. Hence, piRNAs are considered as cancer biomarkers along with several circulating RNAs [171]. Transcriptome analyses highlight the deregulation of piRNAs in several malignant tumors [172–175]. For instance, piR-Hep1 is upregulated in a hepatocellular carcinoma cell line [173]. piR-Hep1 knockdown reduces cell motility and invasiveness, whereas piR-Hep1 overexpression strongly increases cell migration. Finally, the high expression of human PIWI proteins (PiwiL) has also been associated with various cancers. Interestingly, the expression level of PiwiL2 is correlated with tumors aggressiveness. Moreover, PiwiL2 knockdown decreases cellular migration, suggesting a role in invasion [176,177].

It also appears that PIWI proteins are involved in different cellular processes, such as stem-cell maintenance or cell differentiation, in the *Drosophila* germ line [178] and tumor cell proliferation through induction of *c-Myc* expression in human cells [179]. As a consequence, piRNA and PIWI protein deregulation could sustainably affect these essential processes and induce developmental defects. It is possible that piRNA deregulation could partly explain cases of family susceptibility to cancer. Indeed, if the piRNA pathway is impaired, TE regulation is affected, and genome integrity is compromised.

#### 4.2.2. TEs and Disease

In different models (*Drosophila*, mouse, zebrafish, and human cultured cells), *p53* mutation leads to TE derepression associated with a piRNA biogenesis defect, at least in *Drosophila* [180]. Accordingly, *p53*-derived cancers have been considered as ‘transposopathies’, diseases induced by TE mobilization resulting in genome instability [181]. However, even though TE insertions have been detected in several cases of cancers, it is still unclear whether they are the direct cause of cancer development or a consequence of another cellular defect causing cancer. For instance, in colon cancer, retrotransposon insertion into the *APC* tumor suppressor gene leads to its inactivation; in cancer cell lines, retrotransposon insertions are linked to long chimeric transcript emergence and thus, drive the expression of several genes [182,183]. Transcriptome-wide analyses of RNA-seq data for several cancer types and tumors samples identified TE-oncogene chimera transcripts, a phenomenon called a ‘TE onco-exaptation event’. Several TE onco-exaptation events lead to oncogene activation or overexpression. For instance, *AluJb*, which drives *LIN28B* expression, has been implicated in lung cancer. DNA methylation of the *AluJb* sequence leads to a decrease in *LIN28B* expression [184].

Other diseases are linked to TE activity, including hemophilia [185,186], psychiatric disorders [187], and Alzheimer’s disease [188]. In the particular case of Alzheimer’s disease, piRNA molecules seem to be more abundant in the brain and can be considered as a molecular marker of this disease [189]. Indeed, transcriptome analyses of samples isolated from human brain show that some piRNAs are upregulated in Alzheimer’s disease [190].

Altogether, control of TE expression, especially by piRNAs, might play a crucial role in genome stability. If impaired by stress, such as mutation or environmentally-induced piRNA biogenesis misregulation, aberrant gene expression can occur inducing cancer development or other diseases related to TEs in subsequent generations, leading to some of the unexplained family susceptibility.

## 5. Conclusions

There is increasing evidence of environmentally-induced epigenetic inheritance, and today we have a better understanding of their associated mechanisms. However, there are still many points in our understanding of the interaction between the environment and the epigenome that are lacking, in particular in metazoans. Three major mechanisms are usually reported as mediators of epigenetic inheritance, namely DNA methylation, histone modification, and non-coding RNAs.

One stress may cause multiple epigenetic changes. Among these changes, some of them are probably independent, whereas other modifications work together. Nature is a combination of numerous stressors including temperature, pressure variations, pollution, starvation, drought as well as competition for food, social stresses, and so on. It is unlikely that one stress can act only on one biological process, but even if this is the case, we can expect that the addition of all these stresses will impact several important pathways.

A balance between transgenerational epigenetic inheritance and epigenetic reprogramming occurs at each generation, and these processes are tightly regulated during development. Several factors are capable of tipping this epigenetic balance toward transgenerational inheritance, notably after stress exposure. piRNA molecules could be one of these factors. Molecular mechanisms of transgenerational piRNA inheritance have been partially studied, and some hypotheses to explain their maintenance have been proposed, which include other epigenetic mechanisms, such as chromatin modifications and several other potential factors discussed above. For the moment, piRNA function in environmentally-induced epigenetic inheritance has been little explored. TE activity is found in some diseases, and many environmental stresses are known to generate TE mobilization events. In this context, piRNA misregulation may have an important impact on genome integrity and structure. Changes in piRNA expression could also modify the regulation of several genes and influence the initiation and development of some diseases. Future work will surely improve our understanding of the piRNA transgenerational effect in response to environmental injury.

**Author Contributions:** Conceptualization, K.C. and L.T.; Writing—Original Draft Preparation, K.C.; Writing—Review and Editing, K.C., A.B., C.C., and L.T.; Visualization, K.C., A.B., and L.T.

**Funding:** This research was funded by University Pierre et Marie Curie, grant number EME1223 to L.T, Action Incitative 2018 and COST-Epitrans, grant number CA16120 to CC, by financial supports of Sorbonne Université and the CNRS to L.T. and C.C., and by the Ph.D. fellowships from the Ministère de l'Enseignement Supérieur et de la Recherche and la Ligue Nationale Contre le Cancer to KC. We also thank the Réseau André Picard and la Société Française de Biologie du Développement for their supports on travel fellowship grants to K.C.

**Acknowledgments:** We thank FlyBase.org for providing valuable databases. We thank Lori Pile for critical reading of the manuscript, all the members of the Transgenerational Epigenetics & Small RNA Biology (TErBIO) Team for helpful discussions and the two reviewers for their help and comments during the revision process of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Jacob, F.; Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **1961**, *3*, 318–356. [[CrossRef](#)]
- Kaati, G.; Bygren, L.O.; Edvinsson, S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur. J. Hum. Genet.* **2002**, *10*, 682–688. [[CrossRef](#)] [[PubMed](#)]
- Stanner, S.A.; Bulmer, K.; Andres, C.; Lantseva, O.E.; Borodina, V.; Poteen, V.V.; Yudkin, J.S. Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study. *BMJ* **1997**, *315*, 1342–1348. [[CrossRef](#)] [[PubMed](#)]
- Painter, R.C.; Osmond, C.; Gluckman, P.; Hanson, M.; Phillips, D.I.; Roseboom, T.J. Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *BJOG* **2008**, *115*, 1243–1249. [[CrossRef](#)] [[PubMed](#)]
- Veenendaal, M.V.; Painter, R.C.; de Rooij, S.R.; Bossuyt, P.M.; van der Post, J.A.; Gluckman, P.D.; Hanson, M.A.; Roseboom, T.J. Transgenerational effects of prenatal exposure to the 1944–45 Dutch famine. *BJOG* **2013**, *120*, 548–553. [[CrossRef](#)] [[PubMed](#)]
- Yehuda, R.; Daskalakis, N.P.; Bierer, L.M.; Bader, H.N.; Klengel, T.; Holsboer, F.; Binder, E.B. Holocaust exposure induced intergenerational effects on FKBP5 methylation. *Biol. Psychiatry* **2016**, *80*, 372–380. [[CrossRef](#)] [[PubMed](#)]
- Horsthemke, B. A critical view on transgenerational epigenetic inheritance in humans. *Nat. Commun.* **2018**, *9*, 2973. [[CrossRef](#)] [[PubMed](#)]
- Jablonka, E.; Raz, G. Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* **2009**, *84*, 131–176. [[CrossRef](#)] [[PubMed](#)]
- Skinner, M.K. What is an epigenetic transgenerational phenotype? F3 or F2. *Reprod. Toxicol.* **2008**, *25*, 2–6. [[CrossRef](#)] [[PubMed](#)]
- Balhorn, R. The protamine family of sperm nuclear proteins. *Genome Biol.* **2007**, *8*, 227. [[CrossRef](#)] [[PubMed](#)]
- Hammoud, S.S.; Nix, D.A.; Hammoud, A.O.; Gibson, M.; Cairns, B.R.; Carrell, D.T. Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. *Hum. Reprod.* **2011**, *26*, 2558–2569. [[CrossRef](#)] [[PubMed](#)]
- Lim, J.P.; Brunet, A. Bridging the transgenerational gap with epigenetic memory. *Trends Genet.* **2013**, *29*, 176–186. [[CrossRef](#)] [[PubMed](#)]
- Heard, E.; Martienssen, R.A. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* **2014**, *157*, 95–109. [[CrossRef](#)] [[PubMed](#)]
- Szyf, M. Nongenetic inheritance and transgenerational epigenetics. *Trends Mol. Med.* **2015**, *21*, 134–144. [[CrossRef](#)] [[PubMed](#)]
- Manjrekar, J. Epigenetic inheritance, prions and evolution. *J. Genet.* **2017**, *96*, 445–456. [[CrossRef](#)] [[PubMed](#)]
- Ruden, D.M.; Lu, X. Hsp90 affecting chromatin remodeling might explain transgenerational epigenetic inheritance in Drosophila. *Curr. Genom.* **2008**, *9*, 500–508. [[CrossRef](#)] [[PubMed](#)]
- Angelova, M.T.; Dimitrova, D.G.; Dinges, N.; Lence, T.; Worpenberg, L.; Carre, C.; Roignant, J.Y. The emerging field of epitranscriptomics in neurodevelopmental and neuronal disorders. *Front. Bioeng. Biotechnol.* **2018**, *6*, 46. [[CrossRef](#)] [[PubMed](#)]

18. Dimitrova, D.G.; Teyssset, L.; Carre, C. RNA 2'-O-methylation (Nm) modification in human diseases. *Genes (Basel)* **2019**, *10*, 117. [[CrossRef](#)] [[PubMed](#)]
19. Kiani, J.; Grandjean, V.; Liebers, R.; Tuorto, F.; Ghanbarian, H.; Lyko, F.; Cuzin, F.; Rassoulzadegan, M. RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. *PLoS Genet.* **2013**, *9*, e1003498. [[CrossRef](#)]
20. Ben Maamar, M.; Sadler-Riggelman, I.; Beck, D.; Skinner, M.K. Epigenetic transgenerational inheritance of altered sperm histone retention sites. *Sci. Rep.* **2018**, *8*, 5308. [[CrossRef](#)]
21. Skinner, M.K.; Ben Maamar, M.; Sadler-Riggelman, I.; Beck, D.; Nilsson, E.; McBirney, M.; Klukovich, R.; Xie, Y.; Tang, C.; Yan, W. Alterations in sperm DNA methylation, non-coding RNA and histone retention associate with DDT-induced epigenetic transgenerational inheritance of disease. *Epigenet. Chromatin* **2018**, *11*, 8. [[CrossRef](#)] [[PubMed](#)]
22. King, S.E.; McBirney, M.; Beck, D.; Sadler-Riggelman, I.; Nilsson, E.; Skinner, M.K. Sperm epimutation biomarkers of obesity and pathologies following DDT induced epigenetic transgenerational inheritance of disease. *Environ. Epigenet.* **2019**, *5*, dvz008. [[CrossRef](#)] [[PubMed](#)]
23. Klukovich, R.; Nilsson, E.; Sadler-Riggelman, I.; Beck, D.; Xie, Y.; Yan, W.; Skinner, M.K. Environmental toxicant induced epigenetic transgenerational inheritance of prostate pathology and stromal-epithelial cell epigenome and transcriptome alterations: Ancestral origins of prostate disease. *Sci. Rep.* **2019**, *9*, 2209. [[CrossRef](#)] [[PubMed](#)]
24. Kubsad, D.; Nilsson, E.E.; King, S.E.; Sadler-Riggelman, I.; Beck, D.; Skinner, M.K. Assessment of glyphosate induced epigenetic transgenerational inheritance of pathologies and sperm epimutations: Generational toxicology. *Sci. Rep.* **2019**, *9*, 6372. [[CrossRef](#)] [[PubMed](#)]
25. Nilsson, E.; Klukovich, R.; Sadler-Riggelman, I.; Beck, D.; Xie, Y.; Yan, W.; Skinner, M.K. Environmental toxicant induced epigenetic transgenerational inheritance of ovarian pathology and granulosa cell epigenome and transcriptome alterations: Ancestral origins of polycystic ovarian syndrome and primary ovarian insufficiency. *Epigenetics* **2018**, *13*, 875–895. [[CrossRef](#)] [[PubMed](#)]
26. Casier, K.; Delmarre, V.; Gueguen, N.; Hermant, C.; Viode, E.; Vaury, C.; Ronsseray, S.; Brasset, E.; Teyssset, L.; Boivin, A. Environmentally-induced epigenetic conversion of a piRNA cluster. *Elife* **2019**, *8*. [[CrossRef](#)] [[PubMed](#)]
27. Yu, R.; Wang, X.; Moazed, D. Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. *Nature* **2018**, *558*, 615–619. [[CrossRef](#)]
28. Guerrero-Bosagna, C.; Settles, M.; Lucker, B.; Skinner, M.K. Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. *PLoS ONE* **2010**, *5*, e13100. [[CrossRef](#)]
29. Skinner, M.K.; Nilsson, E.; Sadler-Riggelman, I.; Beck, D.; Ben Maamar, M.; McCarrey, J.R. Transgenerational sperm DNA methylation epimutation developmental origins following ancestral vinclozolin exposure. *Epigenetics* **2019**, *14*, 721–739. [[CrossRef](#)]
30. Schuster, A.; Skinner, M.K.; Yan, W. Ancestral vinclozolin exposure alters the epigenetic transgenerational inheritance of sperm small noncoding RNAs. *Environ. Epigenet* **2016**, *2*, dvw001.
31. Brieno-Enriquez, M.A.; Garcia-Lopez, J.; Cardenas, D.B.; Guibert, S.; Cleroux, E.; Ded, L.; Hourcade Jde, D.; Peknicova, J.; Weber, M.; Del Mazo, J. Exposure to endocrine disruptor induces transgenerational epigenetic deregulation of microRNAs in primordial germ cells. *PLoS ONE* **2015**, *10*, e0124296. [[CrossRef](#)] [[PubMed](#)]
32. Skinner, M.K.; Manikkam, M.; Tracey, R.; Guerrero-Bosagna, C.; Haque, M.; Nilsson, E.E. Ancestral dichlorodiphenyltrichloroethane (DDT) exposure promotes epigenetic transgenerational inheritance of obesity. *BMC Med.* **2013**, *11*, 228. [[CrossRef](#)] [[PubMed](#)]
33. Ben Maamar, M.; Nilsson, E.; Sadler-Riggelman, I.; Beck, D.; McCarrey, J.R.; Skinner, M.K. Developmental origins of transgenerational sperm DNA methylation epimutations following ancestral DDT exposure. *Dev. Biol.* **2019**, *445*, 280–293. [[CrossRef](#)] [[PubMed](#)]
34. Nilsson, E.; Larsen, G.; Manikkam, M.; Guerrero-Bosagna, C.; Savenkova, M.I.; Skinner, M.K. Environmentally induced epigenetic transgenerational inheritance of ovarian disease. *PLoS ONE* **2012**, *7*, e36129. [[CrossRef](#)] [[PubMed](#)]
35. Manikkam, M.; Guerrero-Bosagna, C.; Tracey, R.; Haque, M.M.; Skinner, M.K. Transgenerational actions of environmental compounds on reproductive disease and identification of epigenetic biomarkers of ancestral exposures. *PLoS ONE* **2012**, *7*, e31901. [[CrossRef](#)]

36. Manikkam, M.; Tracey, R.; Guerrero-Bosagna, C.; Skinner, M.K. Plastics derived endocrine disruptors (BPA, DEHP and DBP) induce epigenetic transgenerational inheritance of obesity, reproductive disease and sperm epimutations. *PLoS ONE* **2013**, *8*, e55387. [[CrossRef](#)] [[PubMed](#)]
37. Gely-Pernot, A.; Hao, C.; Legoff, L.; Multigner, L.; D’Cruz, S.C.; Kervarrec, C.; Jegou, B.; Tevosian, S.; Smagulova, F. Gestational exposure to chlordecone promotes transgenerational changes in the murine reproductive system of males. *Sci. Rep.* **2018**, *8*, 10274. [[CrossRef](#)]
38. Tracey, R.; Manikkam, M.; Guerrero-Bosagna, C.; Skinner, M.K. Hydrocarbons (jet fuel JP-8) induce epigenetic transgenerational inheritance of obesity, reproductive disease and sperm epimutations. *Reprod. Toxicol.* **2013**, *36*, 104–116. [[CrossRef](#)]
39. Camacho, J.; Truong, L.; Kurt, Z.; Chen, Y.W.; Morselli, M.; Gutierrez, G.; Pellegrini, M.; Yang, X.; Allard, P. The memory of environmental chemical exposure in *C. elegans* is dependent on the jumonji demethylases *jmjd-2* and *jmjd-3/utx-1*. *Cell Rep.* **2018**, *23*, 2392–2404. [[CrossRef](#)]
40. Ou, X.; Zhang, Y.; Xu, C.; Lin, X.; Zang, Q.; Zhuang, T.; Jiang, L.; von Wettstein, D.; Liu, B. Transgenerational inheritance of modified DNA methylation patterns and enhanced tolerance induced by heavy metal stress in rice (*Oryza sativa* L.). *PLoS ONE* **2012**, *7*, e41143. [[CrossRef](#)]
41. Franklin, T.B.; Russig, H.; Weiss, I.C.; Graff, J.; Linder, N.; Michalon, A.; Vizi, S.; Mansuy, I.M. Epigenetic transmission of the impact of early stress across generations. *Biol. Psychiatry* **2010**, *68*, 408–415. [[CrossRef](#)] [[PubMed](#)]
42. Dias, B.G.; Ressler, K.J. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* **2014**, *17*, 89–96. [[CrossRef](#)] [[PubMed](#)]
43. Hoile, S.P.; Lillycrop, K.A.; Thomas, N.A.; Hanson, M.A.; Burdge, G.C. Dietary protein restriction during F0 pregnancy in rats induces transgenerational changes in the hepatic transcriptome in female offspring. *PLoS ONE* **2011**, *6*, e21668. [[CrossRef](#)] [[PubMed](#)]
44. Rechavi, O.; Hourri-Ze’evi, L.; Anava, S.; Goh, W.S.S.; Kerk, S.Y.; Hannon, G.J.; Hobert, O. Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* **2014**, *158*, 277–287. [[CrossRef](#)] [[PubMed](#)]
45. Ashe, A.; Sapetschnig, A.; Weick, E.-M.; Mitchell, J.; Bagijn, M.P.; Cording, A.C.; Doebley, A.-L.; Goldstein, L.D.; Lehrbach, N.J.; Le Pen, J.; et al. piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **2012**, *150*, 88–99. [[CrossRef](#)] [[PubMed](#)]
46. Moore, R.S.; Kaletsky, R.; Murphy, C.T. Piwi/PRG-1 argonaute and TGF- $\beta$  mediate transgenerational learned pathogenic avoidance. *Cell* **2019**, *177*, 1827–1841.e12. [[CrossRef](#)]
47. Govorko, D.; Bekdash, R.A.; Zhang, C.; Sarker, D.K. Male germline transmits fetal alcohol adverse effect on hypothalamic proopiomelanocortin gene across generations. *Biol. Psychiatry* **2012**, *72*, 378–388. [[CrossRef](#)]
48. Kou, H.P.; Li, Y.; Song, X.X.; Ou, X.F.; Xing, S.C.; Ma, J.; Von Wettstein, D.; Liu, B. Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (*Oryza sativa* L.). *J. Plant. Physiol.* **2011**, *168*, 1685–1693. [[CrossRef](#)]
49. Zheng, X.; Chen, L.; Li, M.; Lou, Q.; Xia, H.; Wang, P.; Li, T.; Liu, H.; Luo, L. Transgenerational variations in DNA methylation induced by drought stress in two rice varieties with distinguished difference to drought resistance. *PLoS ONE* **2013**, *8*, e80253. [[CrossRef](#)]
50. Jeremias, G.; Barbosa, J.; Marques, S.M.; De Schamphelaere, K.A.C.; Van Nieuwerburgh, F.; Deforce, D.; Goncalves, F.J.M.; Pereira, J.L.; Asselman, J. Transgenerational inheritance of DNA hypomethylation in *daphnia magna* in response to salinity stress. *Environ. Sci. Technol.* **2018**, *52*, 10114–10123. [[CrossRef](#)]
51. Seong, K.-H.; Li, D.; Shimizu, H.; Nakamura, R.; Ishii, S. Inheritance of stress-induced, ATF-2-dependent epigenetic change. *Cell* **2011**, *145*, 1049–1061. [[CrossRef](#)] [[PubMed](#)]
52. Klosin, A.; Casas, E.; Hidalgo-Carcedo, C.; Vavouri, T.; Lehner, B. Transgenerational transmission of environmental information in *C. elegans*. *Science* **2017**, *356*, 320–323. [[CrossRef](#)] [[PubMed](#)]
53. Ni, J.Z.; Kalinava, N.; Chen, E.; Huang, A.; Trinh, T.; Gu, S.G. A transgenerational role of the germline nuclear RNAi pathway in repressing heat stress-induced transcriptional activation in *C. elegans*. *Epigenet. Chromatin* **2016**, *9*, 3. [[CrossRef](#)] [[PubMed](#)]
54. Schott, D.; Yanai, I.; Hunter, C.P. Natural RNA interference directs a heritable response to the environment. *Sci. Rep.* **2014**, *4*, 7387. [[CrossRef](#)] [[PubMed](#)]
55. Feng, S.; Jacobsen, S.E.; Reik, W. Epigenetic reprogramming in plant and animal development. *Science* **2010**, *330*, 622–627. [[CrossRef](#)]



56. Hackett, J.A.; Zyllicz, J.J.; Surani, M.A. Parallel mechanisms of epigenetic reprogramming in the germline. *Trends Genet.* **2012**, *28*, 164–174. [[CrossRef](#)]
57. Shen, H.; Xu, W.; Lan, F. Histone lysine demethylases in mammalian embryonic development. *Exp. Mol. Med.* **2017**, *49*, e325. [[CrossRef](#)]
58. Seisenberger, S.; Andrews, S.; Krueger, F.; Arand, J.; Walter, J.; Santos, F.; Popp, C.; Thienpont, B.; Dean, W.; Reik, W. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **2012**, *48*, 849–862. [[CrossRef](#)]
59. Chen, T.; Ueda, Y.; Dodge, J.E.; Wang, Z.; Li, E. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell Biol.* **2003**, *23*, 5594–5605. [[CrossRef](#)]
60. Gaydos, L.J.; Wang, W.; Strome, S. H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science* **2014**, *345*, 1515–1518. [[CrossRef](#)]
61. Maenohara, S.; Unoki, M.; Toh, H.; Ohishi, H.; Sharif, J.; Koseki, H.; Sasaki, H. Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos. *PLoS Genet.* **2017**, *13*, e1007042. [[CrossRef](#)] [[PubMed](#)]
62. Houri-Ze'evi, L.; Korem, Y.; Sheftel, H.; Faigenbloom, L.; Toker, I.A.; Dagan, Y.; Awad, L.; Degani, L.; Alon, U.; Rechavi, O. A tunable mechanism determines the duration of the transgenerational small RNA inheritance in *C. elegans*. *Cell* **2016**, *165*, 88–99. [[CrossRef](#)] [[PubMed](#)]
63. Alcazar, R.M.; Lin, R.; Fire, A.Z. Transmission dynamics of heritable silencing induced by double-stranded RNA in *Caenorhabditis elegans*. *Genetics* **2008**, *180*, 1275–1288. [[CrossRef](#)] [[PubMed](#)]
64. Rechavi, O.; Lev, I. Principles of transgenerational small RNA inheritance in *Caenorhabditis elegans*. *Curr. Biol.* **2017**, *27*, R720–R730. [[CrossRef](#)] [[PubMed](#)]
65. Minkina, O.; Hunter, C.P. Intergenerational transmission of gene regulatory information in *Caenorhabditis elegans*. *Trends Genet.* **2018**, *34*, 54–64. [[CrossRef](#)] [[PubMed](#)]
66. Xu, F.; Guang, S.; Feng, X. Distinct nuclear and cytoplasmic machineries cooperatively promote the inheritance of RNAi in *Caenorhabditis elegans*: The inheritance of RNAi. *Biol. Cell* **2018**, *110*, 217–224. [[CrossRef](#)] [[PubMed](#)]
67. Sapetschnig, A.; Sarkies, P.; Lehrbach, N.J.; Miska, E.A. Tertiary siRNAs mediate paramutation in *C. elegans*. *PLoS Genet.* **2015**, *11*, e1005078. [[CrossRef](#)] [[PubMed](#)]
68. Houri-Ze'evi, L.; Rechavi, O. Plastic germline reprogramming of heritable small RNAs enables maintenance or erasure of epigenetic memories. *RNA Biol.* **2016**, *13*, 1212–1217. [[CrossRef](#)]
69. Houri-Zeevi, L.; Rechavi, O. A matter of time: Small RNAs regulate the duration of epigenetic inheritance. *Trends Genet.* **2017**, *33*, 46–57. [[CrossRef](#)]
70. Zhuang, J.J.; Hunter, C.P. The influence of competition among *C. elegans* small RNA pathways on development. *Genes (Basel)* **2012**, *3*, 671–685. [[CrossRef](#)]
71. de Vanssay, A.; Bougé, A.-L.; Boivin, A.; Hermant, C.; Teyssset, L.; Delmarre, V.; Antoniewski, C.; Ronsseray, S. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* **2012**, *490*, 112–115. [[CrossRef](#)] [[PubMed](#)]
72. Hermant, C.; Boivin, A.; Teyssset, L.; Delmarre, V.; Asif-Laidin, A.; van den Beek, M.; Antoniewski, C.; Ronsseray, S. Paramutation in *Drosophila* requires both nuclear and cytoplasmic actors of the piRNA pathway and induces cis-spreading of piRNA production. *Genetics* **2015**, *201*, 1381–1396. [[CrossRef](#)] [[PubMed](#)]
73. Ghildiyal, M.; Zamore, P.D. Small silencing RNAs: An expanding universe. *Nat. Rev. Genet.* **2009**, *10*, 94–108. [[CrossRef](#)] [[PubMed](#)]
74. Brennecke, J.; Aravin, A.A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G.J. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **2007**, *128*, 1089–1103. [[CrossRef](#)] [[PubMed](#)]
75. Gunawardane, L.S.; Saito, K.; Nishida, K.M.; Miyoshi, K.; Kawamura, Y.; Nagami, T.; Siomi, H.; Siomi, M.C. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* **2007**, *315*, 1587–1590. [[CrossRef](#)]
76. Mohn, F.; Sienski, G.; Handler, D.; Brennecke, J. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **2014**, *157*, 1364–1379. [[CrossRef](#)]
77. Andersen, P.R.; Tirian, L.; Vunjak, M.; Brennecke, J. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **2017**, *549*, 54–59. [[CrossRef](#)]

78. ElMaghraby, M.F.; Andersen, P.R.; Pühringer, F.; Hohmann, U.; Meixner, K.; Lendl, T.; Tirian, L.; Brennecke, J. A heterochromatin-specific RNA export pathway facilitates piRNA production. *Cell* **2019**, *178*, 964–979. [\[CrossRef\]](#)
79. Kneuss, E.; Munafò, M.; Eastwood, E.L.; Deumer, U.-S.; Preall, J.B.; Hannon, G.J.; Czech, B. Specialization of the *Drosophila* nuclear export family protein Nxf3 for piRNA precursor export. *Genes Dev.* **2019**, *33*, 1208–1220. [\[CrossRef\]](#)
80. Ipsaro, J.J.; Haase, A.D.; Knott, S.R.; Joshua-Tor, L.; Hannon, G.J. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* **2012**, *491*, 279–283. [\[CrossRef\]](#)
81. Nishimasu, H.; Ishizu, H.; Saito, K.; Fukuhara, S.; Kamatani, M.K.; Bonnefond, L.; Matsumoto, N.; Nishizawa, T.; Nakanaga, K.; Aoki, J.; et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* **2012**, *491*, 284–287. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Zhang, F.; Wang, J.; Xu, J.; Zhang, Z.; Koppetsch, B.S.; Schultz, N.; Vreven, T.; Meignin, C.; Davis, I.; Zamore, P.D.; et al. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **2012**, *151*, 871–884. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Hur, J.K.; Luo, Y.; Moon, S.; Ninova, M.; Marinov, G.K.; Chung, Y.D.; Aravin, A.A. Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in *Drosophila*. *Genes Dev.* **2016**, *30*, 840–855. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Batki, J.; Schnabl, J.; Wang, J.; Handler, D.; Andreev, V.I.; Stieger, C.E.; Novatchkova, M.; Lampersberger, L.; Kauneckaitė, K.; Xie, W.; et al. The nascent RNA binding complex SFiNX licenses piRNA-guided heterochromatin formation. *Nat. Struct. Mol. Biol.* **2019**, *26*, 720–731. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Brower-Toland, B.; Findley, S.D.; Jiang, L.; Liu, L.; Yin, H.; Dus, M.; Zhou, P.; Elgin, S.C.; Lin, H. *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev.* **2007**, *21*, 2300–2311. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Fabry, M.H.; Ciabrelli, F.; Munafò, M.; Eastwood, E.L.; Kneuss, E.; Falcatori, I.; Falconio, F.A.; Hannon, G.J.; Czech, B. piRNA-guided co-transcriptional silencing coopts nuclear export factors. *eLife* **2019**, *8*, e47999. [\[CrossRef\]](#)
87. Klenov, M.S.; Sokolova, O.A.; Yakushev, E.Y.; Stolyarenko, A.D.; Mikhaleva, E.A.; Lavrov, S.A.; Gvozdev, V.A. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18760–18765. [\[CrossRef\]](#)
88. Klenov, M.S.; Lavrov, S.A.; Korbut, A.P.; Stolyarenko, A.D.; Yakushev, E.Y.; Reuter, M.; Pillai, R.S.; Gvozdev, V.A. Impact of nuclear Piwi elimination on chromatin state in *Drosophila melanogaster* ovaries. *Nucleic Acids Res.* **2014**, *42*, 6208–6218. [\[CrossRef\]](#)
89. Le Thomas, A.; Rogers, A.K.; Webster, A.; Marinov, G.K.; Liao, S.E.; Perkins, E.M.; Hur, J.K.; Aravin, A.A.; Tóth, K.F. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* **2013**, *27*, 390–399. [\[CrossRef\]](#)
90. Rangan, P.; Malone, C.D.; Navarro, C.; Newbold, S.P.; Hayes, P.S.; Sachidanandam, R.; Hannon, G.J.; Lehmann, R. piRNA production requires heterochromatin formation in *Drosophila*. *Curr. Biol.* **2011**, *21*, 1373–1379. [\[CrossRef\]](#)
91. Murano, K.; Iwasaki, Y.W.; Ishizu, H.; Mashiko, A.; Shibuya, A.; Kondo, S.; Adachi, S.; Suzuki, S.; Saito, K.; Natsume, T.; et al. Nuclear RNA export factor variant initiates piRNA-guided co-transcriptional silencing. *EMBO J.* **2019**, *38*. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Marie, P.P.; Ronsseray, S.; Boivin, A. From embryo to adult: piRNA-mediated silencing throughout germline development in *Drosophila*. *G3 (Bethesda)* **2017**, *7*, 505–516. [\[CrossRef\]](#) [\[PubMed\]](#)
93. Reiss, D.; Josse, T.; Anxolabehere, D.; Ronsseray, S. aubergine mutations in *Drosophila melanogaster* impair P cytotype determination by telomeric P elements inserted in heterochromatin. *Mol. Genet. Genom.* **2004**, *272*, 336–343. [\[CrossRef\]](#) [\[PubMed\]](#)
94. Klattenhoff, C.; Bratu, D.P.; McGinnis-Schultz, N.; Koppetsch, B.S.; Cook, H.A.; Theurkauf, W.E. *Drosophila* rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Dev. Cell* **2007**, *12*, 45–55. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Molla-Herman, A.; Valles, A.M.; Ganem-Elbaz, C.; Antoniewski, C.; Huynh, J.-R. tRNA processing defects induce replication stress and Chk2-dependent disruption of piRNA transcription. *EMBO J.* **2015**, *34*, 3009–3027. [\[CrossRef\]](#) [\[PubMed\]](#)

96. Grimson, A.; Srivastava, M.; Fahey, B.; Woodcroft, B.J.; Chiang, H.R.; King, N.; Degnan, B.M.; Rokhsar, D.S.; Bartel, D.P. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **2008**, *455*, 1193–1197. [[CrossRef](#)] [[PubMed](#)]
97. Zuo, L.; Wang, Z.; Tan, Y.; Chen, X.; Luo, X. piRNAs and their functions in the brain. *Int. J. Hum. Genet.* **2016**, *16*, 53–60. [[CrossRef](#)]
98. Lewis, S.H.; Quarles, K.A.; Yang, Y.; Tanguy, M.; Frezal, L.; Smith, S.A.; Sharma, P.P.; Cordaux, R.; Gilbert, C.; Giraud, I.; et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat. Ecol. Evol.* **2018**, *2*, 174–181. [[CrossRef](#)] [[PubMed](#)]
99. Jehn, J.; Gebert, D.; Pipilescu, F.; Stern, S.; Kiefer, J.S.T.; Hewel, C.; Rosenkranz, D. PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Commun. Biol.* **2018**, *1*, 137. [[CrossRef](#)]
100. Perera, B.P.U.; Tsai, Z.T.-Y.; Colwell, M.L.; Jones, T.R.; Goodrich, J.M.; Wang, K.; Sartor, M.A.; Faulk, C.; Dolinoy, D.C. Somatic expression of piRNA and associated machinery in the mouse identifies short, tissue-specific piRNA. *Epigenetics* **2019**, *14*, 504–521. [[CrossRef](#)]
101. Brennecke, J.; Malone, C.D.; Aravin, A.A.; Sachidanandam, R.; Stark, A.; Hannon, G.J. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **2008**, *322*, 1387–1392. [[CrossRef](#)] [[PubMed](#)]
102. Saito, K.; Inagaki, S.; Mituyama, T.; Kawamura, Y.; Ono, Y.; Sakota, E.; Kotani, H.; Asai, K.; Siomi, H.; Siomi, M.C. A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* **2009**, *461*, 1296–1299. [[CrossRef](#)] [[PubMed](#)]
103. Jones, B.C.; Wood, J.G.; Chang, C.; Tam, A.D.; Franklin, M.J.; Siegel, E.R.; Helfand, S.L. A somatic piRNA pathway in the *Drosophila* fat body ensures metabolic homeostasis and normal lifespan. *Nat. Commun.* **2016**, *7*, 13856. [[CrossRef](#)] [[PubMed](#)]
104. Sun, W.; Samimi, H.; Gamez, M.; Zare, H.; Frost, B. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. *Nat. Neurosci.* **2018**, *21*, 1038–1048. [[CrossRef](#)] [[PubMed](#)]
105. van den Beek, M.; da Silva, B.; Pouch, J.; Ali Chaouche, M.E.A.; Carre, C.; Antoniewski, C. Dual-layer transposon repression in heads of *Drosophila melanogaster*. *RNA* **2018**, *24*, 1749–1760. [[CrossRef](#)]
106. Robine, N.; Lau, N.C.; Balla, S.; Jin, Z.; Okamura, K.; Kuramochi-Miyagawa, S.; Blower, M.D.; Lai, E.C. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **2009**, *19*, 2066–2076. [[CrossRef](#)]
107. Rouget, C.; Papin, C.; Boureux, A.; Meunier, A.-C.; Franco, B.; Robine, N.; Lai, E.C.; Pelisson, A.; Simonelig, M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* **2010**, *467*, 1128–1132. [[CrossRef](#)]
108. Barckmann, B.; Pierson, S.; Dufourt, J.; Papin, C.; Armenise, C.; Port, F.; Grentzinger, T.; Chambeyron, S.; Baronian, G.; Desvignes, J.-P.; et al. Aubergine iCLIP reveals piRNA-dependent decay of mRNAs involved in germ cell development in the early embryo. *Cell Rep.* **2015**, *12*, 1205–1216. [[CrossRef](#)]
109. Rajasethupathy, P.; Antonov, I.; Sheridan, R.; Frey, S.; Sander, C.; Tuschl, T.; Kandel, E.R. A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* **2012**, *149*, 693–707. [[CrossRef](#)]
110. Kiuchi, T.; Koga, H.; Kawamoto, M.; Shoji, K.; Sakai, H.; Arai, Y.; Ishihara, G.; Kawaoka, S.; Sugano, S.; Shimada, T.; et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* **2014**, *509*, 633–636. [[CrossRef](#)]
111. Roche, S.E.; Rio, D.C. Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the *Drosophila* polycomb group gene, enhancer of zeste. *Genetics* **1998**, *149*, 1839–1855. [[PubMed](#)]
112. Josse, T.; Teyssset, L.; Todeschini, A.-L.; Sidor, C.M.; Anxolabéhère, D.; Ronsseray, S. Telomeric trans-silencing: An epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet.* **2007**, *3*, 1633–1643. [[CrossRef](#)] [[PubMed](#)]
113. Josse, T.; Maurel-Zaffran, C.; de Vanssay, A.; Teyssset, L.; Todeschini, A.-L.; Delmarre, V.; Chaminade, N.; Anxolabehere, D.; Ronsseray, S. Telomeric trans-silencing in *Drosophila melanogaster*: Tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS ONE* **2008**, *3*, e3249. [[CrossRef](#)] [[PubMed](#)]
114. Todeschini, A.-L.; Teyssset, L.; Delmarre, V.; Ronsseray, S. The epigenetic trans-silencing effect in *Drosophila* involves maternally-transmitted small RNAs whose production depends on the piRNA pathway and HP1. *PLoS ONE* **2010**, *5*, e11032. [[CrossRef](#)] [[PubMed](#)]

115. Poyhonen, M.; de Vanssay, A.; Delmarre, V.; Hermant, C.; Todeschini, A.L.; Teyssset, L.; Ronsseray, S. Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3 (Bethesda)* **2012**, *2*, 331–338. [[CrossRef](#)] [[PubMed](#)]
116. Muerdter, F.; Olovnikov, I.; Molaro, A.; Rozhkov, N.V.; Czech, B.; Gordon, A.; Hannon, G.J.; Aravin, A.A. Production of artificial piRNAs in flies and mice. *RNA* **2012**, *18*, 42–52. [[CrossRef](#)]
117. Shirayama, M.; Seth, M.; Lee, H.C.; Gu, W.; Ishidate, T.; Conte, D., Jr.; Mello, C.C. piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **2012**, *150*, 65–77. [[CrossRef](#)] [[PubMed](#)]
118. Chandler, V.L. Paramutation: From maize to mice. *Cell* **2007**, *128*, 641–645. [[CrossRef](#)]
119. Hollick, J.B. Paramutation: A trans-homolog interaction affecting heritable gene regulation. *Curr. Opin. Plant. Biol.* **2012**, *15*, 536–543. [[CrossRef](#)]
120. Ronsseray, S. Paramutation phenomena in non-vertebrate animals. *Semin. Cell Dev. Biol.* **2015**, *44*, 39–46. [[CrossRef](#)]
121. Stuwe, E.; Toth, K.F.; Aravin, A.A. Small but sturdy: Small RNAs in cellular memory and epigenetics. *Genes Dev.* **2014**, *28*, 423–431. [[CrossRef](#)] [[PubMed](#)]
122. Almeida, M.V.; Andrade-Navarro, M.A.; Ketting, R.F. Function and evolution of nematode RNAi pathways. *Noncoding RNA* **2019**, *5*, 8. [[CrossRef](#)] [[PubMed](#)]
123. Batista, P.J.; Ruby, J.G.; Claycomb, J.M.; Chiang, R.; Fahlgren, N.; Kasschau, K.D.; Chaves, D.A.; Gu, W.; Vasale, J.J.; Duan, S.; et al. PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell* **2008**, *31*, 67–78. [[CrossRef](#)] [[PubMed](#)]
124. Das, P.P.; Bagijn, M.P.; Goldstein, L.D.; Woolford, J.R.; Lehrbach, N.J.; Sapetschnig, A.; Buhecha, H.R.; Gilchrist, M.J.; Howe, K.L.; Stark, R.; et al. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell* **2008**, *31*, 79–90. [[CrossRef](#)] [[PubMed](#)]
125. Antoniewski, C.; Carré, C. New rules for regulation of genes by piRNAs in *C. elegans*. *Non-coding RNA Investig.* **2018**, *2*, 33. [[CrossRef](#)]
126. Shen, E.-Z.; Chen, H.; Ozturk, A.R.; Tu, S.; Shirayama, M.; Tang, W.; Ding, Y.-H.; Dai, S.-Y.; Weng, Z.; Mello, C.C. Identification of piRNA binding sites reveals the argonaute regulatory landscape of the *C. elegans* germline. *Cell* **2018**, *172*, 937–951. [[CrossRef](#)]
127. Zhang, D.; Tu, S.; Stubna, M.; Wu, W.-S.; Huang, W.-C.; Weng, Z.; Lee, H.-C. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. *Science* **2018**, *359*, 587–592. [[CrossRef](#)] [[PubMed](#)]
128. Buckley, B.A.; Burkhart, K.B.; Gu, S.G.; Spracklin, G.; Kershner, A.; Fritz, H.; Kimble, J.; Fire, A.; Kennedy, S. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **2012**, *489*, 447–451. [[CrossRef](#)]
129. Guang, S.; Bochner, A.F.; Burkhart, K.B.; Burton, N.; Pavelec, D.M.; Kennedy, S. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. *Nature* **2010**, *465*, 1097–1101. [[CrossRef](#)]
130. Kalinava, N.; Ni, J.Z.; Gajic, Z.; Kim, M.; Ushakov, H.; Gu, S.G. *C. elegans* heterochromatin factor SET-32 plays an essential role in transgenerational establishment of nuclear RNAi-mediated epigenetic silencing. *Cell Rep.* **2018**, *25*, 2273–2284. [[CrossRef](#)]
131. Pak, J.; Fire, A. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **2007**, *315*, 241–244. [[CrossRef](#)] [[PubMed](#)]
132. Woodhouse, R.M.; Buchmann, G.; Hoe, M.; Harney, D.J.; Low, J.K.K.; Larance, M.; Boag, P.R.; Ashe, A. Chromatin modifiers SET-25 and SET-32 are required for establishment but not long-term maintenance of transgenerational epigenetic inheritance. *Cell Rep.* **2018**, *25*, 2259–2272. [[CrossRef](#)] [[PubMed](#)]
133. Xu, F.; Feng, X.; Chen, X.; Weng, C.; Yan, Q.; Xu, T.; Hong, M.; Guang, S. A cytoplasmic argonaute protein promotes the inheritance of RNAi. *Cell Rep.* **2018**, *23*, 2482–2494. [[CrossRef](#)] [[PubMed](#)]
134. Lev, I.; Gingold, H.; Rechavi, O. H3K9me3 is required for inheritance of small RNAs that target a unique subset of newly evolved genes. *eLife* **2019**, *8*. [[CrossRef](#)] [[PubMed](#)]
135. Posner, R.; Toker, I.A.; Antonova, O.; Star, E.; Anava, S.; Azmon, E.; Hendricks, M.; Bracha, S.; Gingold, H.; Rechavi, O. Neuronal small RNAs control behavior transgenerationally. *Cell* **2019**, *177*, 1814–1826. [[CrossRef](#)] [[PubMed](#)]
136. Bagijn, M.P.; Goldstein, L.D.; Sapetschnig, A.; Weick, E.-M.; Bouasker, S.; Lehrbach, N.J.; Simard, M.J.; Miska, E.A. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* **2012**, *337*, 574–578. [[CrossRef](#)] [[PubMed](#)]

137. Lee, H.-C.; Gu, W.; Shirayama, M.; Youngman, E.; Conte, D.; Mello, C.C. *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. *Cell* **2012**, *150*, 78–87. [[CrossRef](#)]
138. Gent, J.I.; Lamm, A.T.; Pavelec, D.M.; Maniar, J.M.; Parameswaran, P.; Tao, L.; Kennedy, S.; Fire, A.Z. Distinct Phases of siRNA synthesis in an endogenous RNAi pathway in *C. elegans* soma. *Mol. Cell* **2010**, *37*, 679–689. [[CrossRef](#)]
139. Gapp, K.; Jawaid, A.; Sarkies, P.; Bohacek, J.; Pelczar, P.; Prados, J.; Farinelli, L.; Miska, E.; Mansuy, I.M. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* **2014**, *17*, 667–669. [[CrossRef](#)]
140. Grandjean, V.; Fourre, S.; De Abreu, D.A.F.; Derieppe, M.-A.; Remy, J.-J.; Rassoulzadegan, M. RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders. *Sci. Rep.* **2015**, *5*, 18193. [[CrossRef](#)]
141. de Castro Barbosa, T.; Ingerslev, L.R.; Alm, P.S.; Versteijhe, S.; Massart, J.; Rasmussen, M.; Donkin, I.; Sjogren, R.; Mudry, J.M.; Vetterli, L.; et al. High-fat diet reprograms the epigenome of rat spermatozoa and transgenerationally affects metabolism of the offspring. *Mol. Metab.* **2016**, *5*, 184–197. [[CrossRef](#)] [[PubMed](#)]
142. Belicard, T.; Jareosettasin, P.; Sarkies, P. The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress. *BMC Biol.* **2018**, *16*, 103. [[CrossRef](#)] [[PubMed](#)]
143. Meisel, J.D.; Panda, O.; Mahanti, P.; Schroeder, F.C.; Kim, D.H. Chemosensation of bacterial secondary metabolites modulates neuroendocrine signaling and behavior of *C. elegans*. *Cell* **2014**, *159*, 267–280. [[CrossRef](#)] [[PubMed](#)]
144. Richards, E.J. Inherited epigenetic variation—revisiting soft inheritance. *Nat. Rev. Genet.* **2006**, *7*, 395–401. [[CrossRef](#)] [[PubMed](#)]
145. Skinner, M.K. Environmental epigenetics and a unified theory of the molecular aspects of evolution: A neo-lamarckian concept that facilitates neo-darwinian evolution. *Genome Biol. Evol.* **2015**, *7*, 1296–1302. [[CrossRef](#)] [[PubMed](#)]
146. Tauffenberger, A.; Parker, J.A. Heritable transmission of stress resistance by high dietary glucose in *Caenorhabditis elegans*. *PLoS Genet.* **2014**, *10*, e1004346. [[CrossRef](#)] [[PubMed](#)]
147. Jobson, M.A.; Jordan, J.M.; Sandrof, M.A.; Hibshman, J.D.; Lennox, A.L.; Baugh, L.R. Transgenerational effects of early life starvation on growth, reproduction, and stress resistance in *Caenorhabditis elegans*. *Genetics* **2015**, *201*, 201–212. [[CrossRef](#)] [[PubMed](#)]
148. Kishimoto, S.; Uno, M.; Okabe, E.; Nono, M.; Nishida, E. Environmental stresses induce transgenerationally inheritable survival advantages via germline-to-soma communication in *Caenorhabditis elegans*. *Nat. Commun.* **2017**, *8*, 14031. [[CrossRef](#)] [[PubMed](#)]
149. Boffelli, D.; Martin, D.I.K. Epigenetic inheritance: A contributor to species differentiation? *DNA Cell Biol.* **2012**, *31*, S11. [[CrossRef](#)] [[PubMed](#)]
150. Brevik, K.; Lindstrom, L.; McKay, S.D.; Chen, Y.H. Transgenerational effects of insecticides—implications for rapid pest evolution in agroecosystems. *Curr. Opin. Insect Sci.* **2018**, *26*, 34–40. [[CrossRef](#)] [[PubMed](#)]
151. Daniels, S.B.; Peterson, K.R.; Strausbaugh, L.D.; Kidwell, M.G.; Chovnick, A. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* **1990**, *124*, 339–355. [[PubMed](#)]
152. Engels, W.R. Germ line aberrations associated with a case of hybrid dysgenesis in *Drosophila melanogaster* males. *Genet. Res.* **1979**, *33*, 137–146. [[CrossRef](#)] [[PubMed](#)]
153. Ronssey, S.; Anxolabehere, D.; Periquet, G. Hybrid dysgenesis in *Drosophila melanogaster*: Influence of temperature on cytotypic determination in the P-M system. *Mol. Gen. Genet.* **1984**, *196*, 17–23. [[CrossRef](#)] [[PubMed](#)]
154. Kofler, R.; Hill, T.; Nolte, V.; Betancourt, A.J.; Schlötterer, C. The recent invasion of natural *Drosophila simulans* populations by the P-element. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6659–6663. [[CrossRef](#)] [[PubMed](#)]
155. Kofler, R.; Senti, K.-A.; Nolte, V.; Tobler, R.; Schlötterer, C. Molecular dissection of a natural transposable element invasion. *Genome Res.* **2018**, *28*, 824–835. [[CrossRef](#)] [[PubMed](#)]
156. Ronssey, S.; Lehmann, M.; Anxolabehère, D. Copy number and distribution of P and I mobile elements in *Drosophila melanogaster* populations. *Chromosoma* **1989**, *98*, 207–214. [[CrossRef](#)] [[PubMed](#)]
157. Asif-Laidin, A.; Delmarre, V.; Laurentie, J.; Miller, W.J.; Ronssey, S.; Teyssset, L. Short and long-term evolutionary dynamics of subtelomeric piRNA clusters in *Drosophila*. *DNA Res.* **2017**, *24*, 459–472. [[CrossRef](#)]

158. Rebollo, R.; Romanish, M.T.; Mager, D.L. Transposable elements: An abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* **2012**, *46*, 21–42. [[CrossRef](#)]
159. van de Lagemaat, L.N.; Landry, J.R.; Mager, D.L.; Medstrand, P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **2003**, *19*, 530–536. [[CrossRef](#)]
160. Garcia Guerreiro, M.P. What makes transposable elements move in the Drosophila genome? *Heredity* **2012**, *108*, 461–468. [[CrossRef](#)]
161. Miousse, I.R.; Chalbot, M.-C.G.; Lumen, A.; Ferguson, A.; Kavouras, I.G.; Koturbash, I. Response of transposable elements to environmental stressors. *Mutat Res. Rev. Mutat. Res.* **2015**, *765*, 19–39. [[CrossRef](#)] [[PubMed](#)]
162. Feschotte, C.; Pritham, E.J. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* **2007**, *41*, 331–368. [[CrossRef](#)] [[PubMed](#)]
163. Jangam, D.; Feschotte, C.; Betrán, E. Transposable element domestication as an adaptation to evolutionary conflicts. *Trends Genet.* **2017**, *33*, 817–831. [[CrossRef](#)] [[PubMed](#)]
164. Parhad, S.S.; Tu, S.; Weng, Z.; Theurkauf, W.E. Adaptive evolution leads to cross-species incompatibility in the piRNA transposon silencing machinery. *Dev. Cell* **2017**, *43*, 60–70. [[CrossRef](#)] [[PubMed](#)]
165. Kotov, A.A.; Adashev, V.E.; Godneeva, B.K.; Ninova, M.; Shatskikh, A.S.; Bazylev, S.S.; Aravin, A.A.; Olenina, L.V. piRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in *Drosophila melanogaster*. *Nucleic Acids Res.* **2019**, *47*, 4255–4271. [[CrossRef](#)] [[PubMed](#)]
166. Skinner, M.K. Environmental stress and epigenetic transgenerational inheritance. *BMC Med.* **2014**, *12*, 153. [[CrossRef](#)] [[PubMed](#)]
167. Nilsson, E.E.; Sadler-Riggelman, I.; Skinner, M.K. Environmentally induced epigenetic transgenerational inheritance of disease. *Environ. Epigenet.* **2018**, *4*, dvy016. [[CrossRef](#)]
168. Perera, B.P.U.; Faulk, C.; Svoboda, L.K.; Goodrich, J.M.; Dolinoy, D.C. The role of environmental exposures and the epigenome in health and disease. *Environ. Mol. Mutagen.* **2019**. [[CrossRef](#)]
169. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorf, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)]
170. Rassoulzadegan, M.; Cuzin, F. From paramutation to human disease: RNA-mediated heredity. *Semin. Cell Dev. Biol.* **2015**, *44*, 47–50. [[CrossRef](#)]
171. Sole, C.; Arnaiz, E.; Manterola, L.; Otaegui, D.; Lawrie, C.H. The circulating transcriptome as a source of cancer liquid biopsy biomarkers. *Semin. Cancer Biol.* **2019**, *58*, 100–108. [[CrossRef](#)] [[PubMed](#)]
172. Huang, G.; Hu, H.; Xue, X.; Shen, S.; Gao, E.; Guo, G.; Shen, X.; Zhang, X. Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. *Clin. Transl. Oncol.* **2013**, *15*, 563–568. [[CrossRef](#)] [[PubMed](#)]
173. Law, P.T.Y.; Qin, H.; Ching, A.K.K.; Lai, K.P.; Co, N.N.; He, M.; Lung, R.W.M.; Chan, A.W.H.; Chan, T.-F.; Wong, N. Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. *J. Hepatol.* **2013**, *58*, 1165–1173. [[CrossRef](#)] [[PubMed](#)]
174. Li, Y.; Wu, X.; Gao, H.; Jin, J.M.; Li, A.X.; Kim, Y.S.; Pal, S.K.; Nelson, R.A.; Lau, C.M.; Guo, C.; et al. Piwi-interacting RNAs (piRNAs) are dysregulated in renal cell carcinoma and associated with tumor metastasis and cancer-specific survival. *Mol. Med.* **2015**, *21*, 381–388. [[CrossRef](#)] [[PubMed](#)]
175. Tan, L.; Mai, D.; Zhang, B.; Jiang, X.; Zhang, J.; Bai, R.; Ye, Y.; Li, M.; Pan, L.; Su, J.; et al. PIWI-interacting RNA-36712 restrains breast cancer progression and chemoresistance by interaction with SEPW1 pseudogene SEPW1P RNA. *Mol. Cancer* **2019**, *18*, 9. [[CrossRef](#)] [[PubMed](#)]
176. Li, D.; Sun, X.; Yan, D.; Huang, J.; Luo, Q.; Tang, H.; Peng, Z. Piwil2 modulates the proliferation and metastasis of colon cancer via regulation of matrix metalloproteinase 9 transcriptional activity. *Exp. Biol. Med.* **2012**, *237*, 1231–1240. [[CrossRef](#)] [[PubMed](#)]
177. Mei, Y.; Clark, D.; Mao, L. Novel dimensions of piRNAs in cancer. *Cancer Lett.* **2013**, *336*, 46–52. [[CrossRef](#)] [[PubMed](#)]
178. Rojas-Rios, P.; Chartier, A.; Pierson, S.; Simonelig, M. Aubergine and piRNAs promote germline stem cell self-renewal by repressing the proto-oncogene Cbl. *EMBO J.* **2017**, *36*, 3194–3211. [[CrossRef](#)]

179. Yao, Y.; Li, C.; Zhou, X.; Zhang, Y.; Lu, Y.; Chen, J.; Zheng, X.; Tao, D.; Liu, Y.; Ma, Y. PIWIL2 induces c-Myc expression by interacting with NME2 and regulates c-Myc-mediated tumor cell proliferation. *Oncotarget* **2014**, *5*, 8466–8477. [[CrossRef](#)]
180. Wylie, A.; Jones, A.E.; D’Brot, A.; Lu, W.-J.; Kurtz, P.; Moran, J.V.; Rakheja, D.; Chen, K.S.; Hammer, R.E.; Comerford, S.A.; et al. p53 genes function to restrain mobile elements. *Genes Dev.* **2016**, *30*, 64–77. [[CrossRef](#)]
181. Wylie, A.; Jones, A.E.; Abrams, J.M. p53 in the game of transposons. *Bioessays* **2016**, *38*, 1111–1116. [[CrossRef](#)] [[PubMed](#)]
182. Miki, Y.; Nishisho, I.; Horii, A.; Miyoshi, Y.; Utsunomiya, J.; Kinzler, K.W.; Vogelstein, B.; Nakamura, Y. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* **1992**, *52*, 643–645. [[PubMed](#)]
183. Cruickshanks, H.A.; Tufarelli, C. Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter. *Genomics* **2009**, *94*, 397–406. [[CrossRef](#)] [[PubMed](#)]
184. Jang, H.S.; Shah, N.M.; Du, A.Y.; Dailey, Z.Z.; Pehrsson, E.C.; Godoy, P.M.; Zhang, D.; Li, D.; Xing, X.; Kim, S.; et al. Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.* **2019**, *51*, 611–617. [[CrossRef](#)] [[PubMed](#)]
185. Kazazian, H.H., Jr.; Wong, C.; Youssoufian, H.; Scott, A.F.; Phillips, D.G.; Antonarakis, S.E. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* **1988**, *332*, 164–166. [[CrossRef](#)] [[PubMed](#)]
186. Ganguly, A.; Dunbar, T.; Chen, P.; Godmilow, L.; Ganguly, T. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum. Genet.* **2003**, *113*, 348–352. [[CrossRef](#)] [[PubMed](#)]
187. Guffanti, G.; Gaudi, S.; Fallon, J.H.; Sobell, J.; Potkin, S.G.; Pato, C.; Macciardi, F. Transposable elements and psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **2014**, *165B*, 201–216. [[CrossRef](#)]
188. Guo, C.; Jeong, H.-H.; Hsieh, Y.-C.; Klein, H.-U.; Bennett, D.A.; De Jager, P.L.; Liu, Z.; Shulman, J.M. Tau activates transposable elements in Alzheimer’s disease. *Cell Rep.* **2018**, *23*, 2874–2880. [[CrossRef](#)]
189. Qiu, W.; Guo, X.; Lin, X.; Yang, Q.; Zhang, W.; Zhang, Y.; Zuo, L.; Zhu, Y.; Li, C.-S.R.; Ma, C.; et al. Transcriptome-wide piRNA profiling in human brains of Alzheimer’s disease. *Neurobiol. Aging* **2017**, *57*, 170–177. [[CrossRef](#)]
190. Roy, J.; Sarkar, A.; Parida, S.; Ghosh, Z.; Mallick, B. Small RNA sequencing revealed dysregulated piRNAs in Alzheimer’s disease and their probable role in pathogenesis. *Mol. Biosyst.* **2017**, *13*, 565–576. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## IV. Question biologique

De manière générale, notre équipe s'intéresse à la biologie des petits ARN non-codants et aux phénomènes transgénérationnels qui y sont associés. Parmi nos thématiques de recherche, nous nous focalisons, entre autres, sur la biologie des piRNA clusters afin de comprendre comment ces séquences se forment, quelle est leur dynamique ou encore quelles sont leurs propriétés fonctionnelles et moléculaires. Nous pouvons facilement aborder ces questions grâce à des outils génétiques variés ainsi qu'à des approches de biologie moléculaire et de séquençage à haut-débit.

Dans la lignée germinale, les piRNA clusters sont des séquences régulatrices majeures de l'activité des ET. En 2012, l'utilisation d'un transgène mimant un piRNA cluster, *BX2*, a permis à notre équipe de montrer que l'héritage maternel de piRNA est capable d'activer la production de piRNA dans la descendance pour le cluster (de Vanssay *et al*, 2012). Ce cas de paramutation, identifié chez la drosophile, indique que l'activité des piRNA clusters peut être maintenue génération après génération grâce à l'héritage des piRNA maternels par l'embryon. Autrement dit, il faut des piRNA pour activer un piRNA cluster...et pour fabriquer de nouveaux piRNA. Cette situation évoque le paradoxe de l'œuf et de la poule puisqu'on peut s'interroger sur qui du piRNA ou du piRNA cluster apparaît le premier. Ainsi, on peut se demander comment produire les premiers piRNA, ou encore si l'on peut activer la production de piRNA en l'absence d'héritage de piRNA maternels.

Pour aborder cette question, j'ai étudié le rôle de facteurs environnementaux, en particulier celui de la température, et de facteurs génétiques dans l'activation *de novo* de la production de piRNA par le cluster *BX2*. En parallèle, j'ai participé à des travaux portant sur la dynamique des piRNA clusters au cours des générations et la mise en place de la répression d'un ET après son arrivée au sein du génome.





# **RESULTATS**

# I. Effet de facteurs environnementaux sur l'état épigénétique de *BX2*

## 1. Effet de la chaleur

### 1.1. Présentation de l'Article I

En 1984, Stéphane Ronsseray a montré que la régulation de l'élément *P* est sensible à la température. En effet, une température élevée augmente les capacités régulatrices de l'élément *P* dans le cas d'un croisement dysgénésique (Ronsseray *et al*, 1984). Ces travaux indiquent que lorsque la température augmente, la régulation de l'élément *P*, *via* la voie des piRNAs, peut se mettre en place et suggèrent ainsi que l'environnement joue un rôle important dans la mise en place de la régulation des ET de manière générale. Partant de cette hypothèse, Stéphane et Valérie Delmarre, technicienne au laboratoire, ont entrepris de premières expériences où ils ont laissé se développer des lignées *BX2* à une température élevée (29°C au lieu de 25°C). Ils ont constaté qu'après plusieurs générations passées à 29°C, le cluster *BX2<sup>OFF</sup>* pouvait être activé pour la production de piRNA et ce en l'absence d'héritage de piRNA maternels.

Le premier objectif de ma thèse était de préciser ce phénomène de conversion épigénétique de *BX2<sup>OFF</sup>* par une température élevée et d'identifier les mécanismes moléculaires impliqués. Nous avons mis en évidence que la conversion de *BX2<sup>OFF</sup>* à 29°C ne requiert qu'une seule génération et qu'une fois le cluster activé, son état épigénétique peut être maintenu aux générations suivantes même après un retour à une température normale (25°C). En effet, l'état stabilisé est associé à une production de piRNA et à un enrichissement en marques hétérochromatiques (H3K9me3 et Rhino). Nous avons également montré qu'à 29°C, on détecte plus de transcrits précurseurs antisens de *BX2* par rapport à 25°C. Par ailleurs, l'activation de *BX2<sup>OFF</sup>* à 29°C n'est possible que lorsqu'une séquence cible est présente dans le génome et transcrite. Ces résultats suggèrent que l'activation de *BX2* à 29°C repose sur un mécanisme dépendant de l'ARN. Ces travaux ont été publiés en mars 2019 dans la revue eLife.

**1.2. Article I: Environmentally-induced epigenetic conversion of a piRNA cluster**

# Environmentally-induced epigenetic conversion of a piRNA cluster

Karine Casier<sup>1</sup>, Valérie Delmarre<sup>1</sup>, Nathalie Gueguen<sup>2</sup>, Catherine Hermant<sup>1†</sup>, Elise Viodé<sup>1</sup>, Chantal Vaury<sup>2</sup>, Stéphane Ronsseray<sup>1</sup>, Emilie Brasset<sup>2</sup>, Laure Teyssset<sup>1\*</sup>, Antoine Boivin<sup>1\*</sup>

<sup>1</sup>Laboratoire Biologie du Développement, UMR7622, Sorbonne Université, CNRS, Institut de Biologie Paris-Seine, Paris, France; <sup>2</sup>GReD, Université Clermont Auvergne, CNRS, INSERM, BP 10448, Clermont-Ferrand, France

**Abstract** Transposable element (TE) activity is repressed in animal gonads by PIWI-interacting RNAs (piRNAs) produced by piRNA clusters. Current models in flies propose that germinal piRNA clusters are functionally defined by the maternal inheritance of piRNAs produced during the previous generation. Taking advantage of an inactive, but ready to go, cluster of *P*-element derived transgene insertions in *Drosophila melanogaster*, we show here that raising flies at high temperature (29°C) instead of 25°C triggers the stable conversion of this locus from inactive into actively producing functional piRNAs. The increase of antisense transcripts from the cluster at 29°C combined with the requirement of transcription of euchromatic homologous sequences, suggests a role of double stranded RNA in the production of *de novo* piRNAs. This report describes the first case of the establishment of an active piRNA cluster by environmental changes in the absence of maternal inheritance of homologous piRNAs.

**Editorial note:** This article has been through an editorial process in which the authors decide how to respond to the issues raised during peer review. The Reviewing Editor's assessment is that all the issues have been addressed (see decision letter).

DOI: <https://doi.org/10.7554/eLife.39842.001>

**\*For correspondence:**

laure.teyssset@upmc.fr (LT);  
antoine.boivin@upmc.fr (AB)

**Present address:** <sup>†</sup>Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS, INSERM, PSL Research University, Paris, France

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 19

**Received:** 04 July 2018

**Accepted:** 06 March 2019

**Published:** 15 March 2019

**Reviewing editor:** Timothy W Nilsen, Case Western Reserve University, United States

© Copyright Casier et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

Transposable element (TE) activity needs to be repressed to avoid severe genome instability and gametogenesis defects. In humans, growing evidence has implicated TE in several disorders such as cancers defining a new field of diseases called transposopathies (Wylie et al., 2016a; Wylie et al., 2016b). In the animal germline, TE activity is controlled at both transcriptional and post-transcriptional levels by small RNAs called piRNAs associated with the PIWI clade of germline Argonaute proteins (Piwi, Aub and Ago3 in *Drosophila*) (Brennecke et al., 2007; Gunawardane et al., 2007; Le Thomas et al., 2013; Sienski et al., 2012). piRNAs are processed from transcripts produced from specific heterochromatic loci enriched in TE fragments, called piRNA clusters (Brennecke et al., 2007; Gunawardane et al., 2007). These loci undergo non-canonical transcription, ignoring splicing and transcription termination signals, licensed by specific protein complexes such as Rhino-Deadlock-Cutoff (Mohn et al., 2014; Zhang et al., 2014) and Moonshiner-TRF2 (Andersen et al., 2017). Thus, when a new TE inserts into a naive genome, it will freely transpose until one copy gets inserted into a piRNA cluster leading to the production of homologous new TE piRNAs that will then repress transposition (Brennecke et al., 2008). In support of this idea, exogenous sequences inserted into preexisting piRNA clusters lead to the production of matching piRNAs (de Vanssay et al., 2012; Hermant et al., 2015; Marie et al., 2017; Muerdter et al., 2012; Pöyhönen et al., 2012). The specificity of the efficient repression mediated by piRNAs appears to be determined solely by the piRNA cluster sequences. Thus, it raises the question of how piRNA

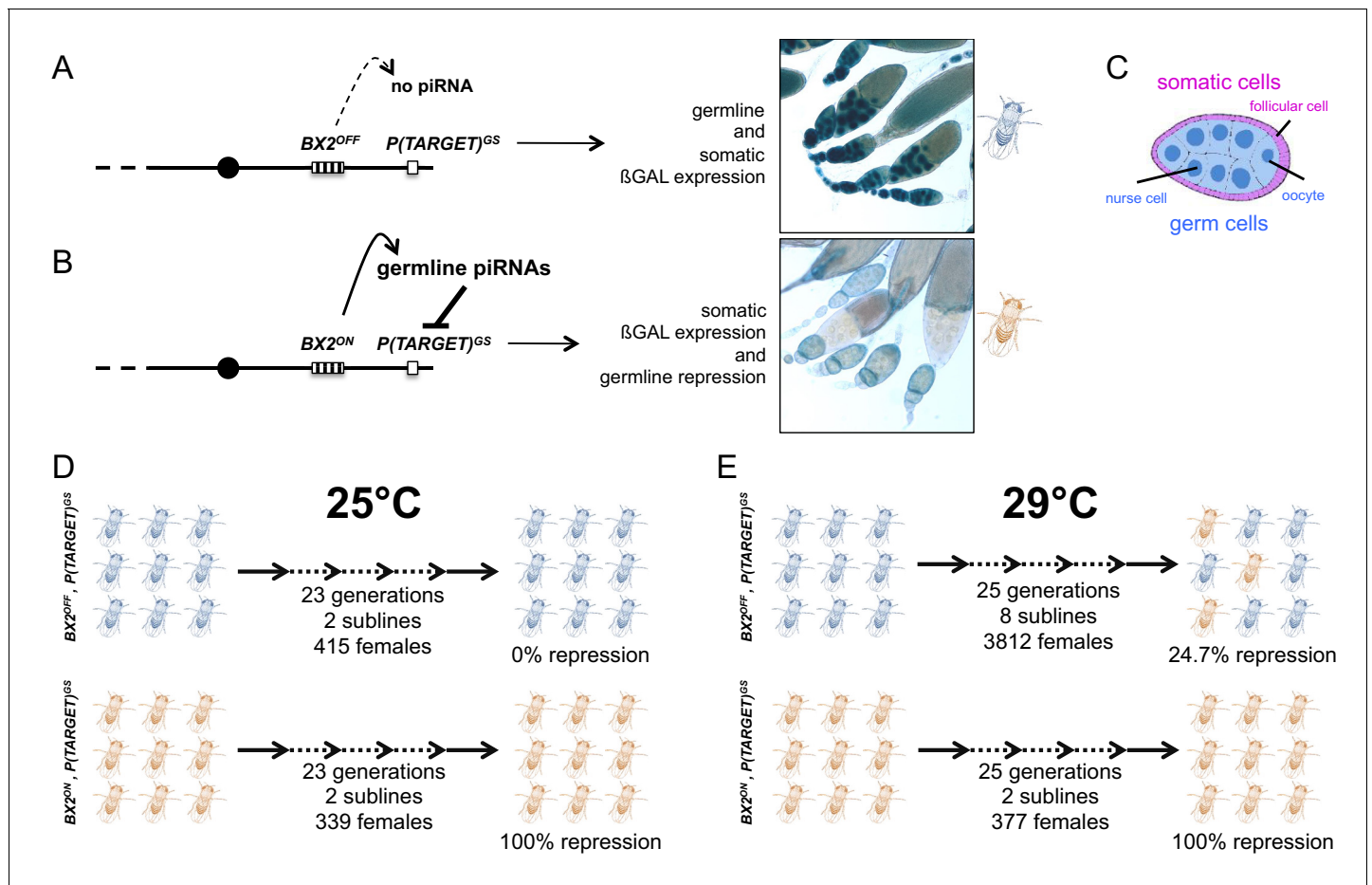
cluster loci are themselves specified. Histone H3 lysine nine tri-methylation (H3K9me3) that is recognized by Rhino, a paralog of heterochromatin protein HP1 (Klattenhoff et al., 2009), is a shared feature of piRNA clusters. Enrichment of H3K9me3, however, is not specific to piRNA clusters and tethering Rhino onto a transgene leads to the production of piRNAs only when both sense and anti-sense transcripts are produced (Zhang et al., 2014). This suggests that neither H3K9me3 marks nor having Rhino-bound is sufficient to induce piRNA production. One current model proposes that piRNA clusters are defined and activated at each generation by the deposition in the egg of their corresponding piRNAs from the mother (Huang et al., 2017). In support of this model, we previously described the first case of a stable transgenerational epigenetic conversion known as paramutation in animals (de Vanssay et al., 2012). This phenomenon was first described in plants and defined as "an epigenetic interaction between two alleles of a locus, through which one allele induces a heritable modification of the other allele without modifying the DNA sequence" (Brink, 1956; Chandler, 2007). In our previous study, we showed that an inactive non-producing piRNA cluster of *P* transgene insertions inherited from the father can be converted into a piRNA-producing cluster by piRNAs inherited from the mother (de Vanssay et al., 2012). This attractive model, however, does not answer the question of how the first piRNAs were produced.

To address this paradox, we used the same *BX2* cluster of seven *P(lacW)* transgenes, which resulted from multiple and successive *P(lacW)* transposition events, thus resembling the structure of natural piRNA clusters (Dorer and Henikoff, 1994). The key advantage of the *BX2* locus is that it can exist in two epigenetic states for the production of germline piRNAs: 1) the inactive state (*BX2*<sup>OFF</sup>) does not produce any piRNAs and thus is unable to repress the expression of homologous sequences, and 2) the active state (*BX2*<sup>ON</sup>) produces abundant piRNAs that functionally repress a homologous reporter transgene in the female germline (de Vanssay et al., 2012; Hermant et al., 2015). We therefore used *BX2* in an inactive state to search for conditions that would convert it into an active piRNA-producing locus, without pre-existing maternal piRNAs. In this report, we describe how culturing flies at high temperature, 29°C instead of 25°C, induces the conversion of an inactive *BX2* locus (*BX2*<sup>OFF</sup>) into a stable piRNA cluster exhibiting repression properties (*BX2*<sup>ON</sup>). It should be noted that flies in their natural habitat exhibit this range of temperature, especially in the context of global warming. These data provide the first report of a *de novo* piRNA cluster establishment independent of maternal inheritance of homologous piRNAs and highlight how environmental changes can stably induce transgenerational modification of the epigenome.

## Results

### Germline silencing induced at high temperature

Earlier studies of hybrid dysgenesis reported that high temperature enhances *P*-element repression (Ronsseray et al., 1984) and that thermic modification of *P* repression can persist over several generations (Ronsseray, 1986). Moreover, *P*-element repression in a strain carrying two *P*-elements inserted into a subtelomeric piRNA cluster can be stimulated by heat treatment (Ronsseray et al., 1991). Very recently, the tracking of natural invasion of *P* elements in *Drosophila simulans* confirmed the key role of high temperature in the establishment of repression through generations (Kofler et al., 2018). These results suggested that temperature may influence the activity of some piRNA clusters. To investigate whether high temperature (29°C) could affect the stability of *BX2* epialleles (*BX2*<sup>OFF</sup> and *BX2*<sup>ON</sup>) across generations, we generated flies carrying, on the same chromosome, each of the *BX2* epialleles and an euchromatic reporter transgene sharing *P* and *lacZ* sequences with *BX2* (made of seven *P(lacW)*, Figure 1—figure supplement 1A). This transgene promotes the expression of β-Galactosidase both in the germline and in the somatic cells of the ovary and thus will hereinafter be referred to as '*P(TARGET)*<sup>GS</sup>' (Figure 1—figure supplement 1B). As was previously described (de Vanssay et al., 2012), at 25°C *BX2*<sup>OFF</sup> does not synthesize functional piRNAs complementary to *P(TARGET)*<sup>GS</sup> resulting in β-Galactosidase expression in whole ovaries of *BX2*<sup>OFF</sup>, *P(TARGET)*<sup>GS</sup> lines (Figure 1A). Whereas in *BX2*<sup>ON</sup>, *P(TARGET)*<sup>GS</sup> lines, functional *lacZ* piRNAs are synthesized in the germline where they specifically repress the *P(TARGET)*<sup>GS</sup> β-Galactosidase expression (Figure 1B). Both *BX2*<sup>OFF</sup>, *P(TARGET)*<sup>GS</sup> and *BX2*<sup>ON</sup>, *P(TARGET)*<sup>GS</sup> lines incubated at 25°C for 23 generations maintained their epigenetic state, showing that both epialleles are stable (Figure 1D, Supplementary file 1). At 29°C, the repression capacity of *BX2*<sup>ON</sup>, *P(TARGET)*<sup>GS</sup> lines



**Figure 1.** Functional assay of the *BX2* epigenetic state. Females carrying either one of the *BX2* epialleles and *P(TARGET)<sup>GS</sup>* were analyzed. (A) When *BX2* is OFF for production of piRNAs (*BX2<sup>OFF</sup>*), no repression of *P(TARGET)<sup>GS</sup>* occurs, allowing expression of  $\beta$ -Galactosidase in both germline and in somatic lineages in ovaries ( $\beta$ -Galactosidase staining). *BX2<sup>OFF</sup>* is illustrated by a blue fly. (B) When *BX2* is ON for production of piRNAs (*BX2<sup>ON</sup>*), repression of *P(TARGET)<sup>GS</sup>* occurs only in the germline lineage. *BX2<sup>ON</sup>* is illustrated by a light brown fly. (C) Drawing of an intermediate egg chamber showing germ cells (nurse cells and oocyte in blue) surrounded by somatic follicular cells (in pink), adapted from **Figure 1A** from **Frydman and Spradling (2001)**. (D) At 25°C, *BX2<sup>OFF</sup>* and *BX2<sup>ON</sup>* are stable over generations. (E) At 29°C, *BX2<sup>OFF</sup>* can be converted into *BX2<sup>ON</sup>*, while *BX2<sup>ON</sup>* is stable over generations.

DOI: <https://doi.org/10.7554/eLife.39842.002>

The following figure supplements are available for figure 1:

**Figure supplement 1.** Schematic maps, drawn at scale, of the different transgenes used in this study.

DOI: <https://doi.org/10.7554/eLife.39842.003>

**Figure supplement 2.** Maternal (A) and paramutagenic (B) effect of *BX2<sup>ON</sup>* lines.

DOI: <https://doi.org/10.7554/eLife.39842.004>

**Figure supplement 3.** Dynamic of conversion across generations at 29°C.

DOI: <https://doi.org/10.7554/eLife.39842.005>

remained stable through 25 generations. Among the *BX2<sup>OFF</sup>*, *P(TARGET)<sup>GS</sup>* lines, 24.7% of females analyzed during 25 generations showed a complete and specific germline  $\beta$ -Galactosidase repression ( $n = 3812$ , **Figure 1E** and **Supplementary file 1**), suggesting a conversion of the *BX2<sup>OFF</sup>* epiallele into *BX2<sup>ON</sup>*. Interestingly, the appearance of females showing  $\beta$ -Galactosidase repression was gradual and stochastic, resulting in a global frequency that increased with the number of generations (**Supplementary file 1**). To test whether the temperature-induced conversion was stable, a set of five lines showing full repression capacity after 23 generations at 29°C, obtained from an independent experiment, were transferred to 25°C and tested for their silencing capacities for several generations. In all cases, the silencing capacities of the *BX2<sup>ON</sup>* epiallele induced at 29°C remained stable

during 50 additional generations at 25°C (**Supplementary file 2**). These stable  $BX2^{ON}$  lines converted by high temperature were named hereafter  $BX2^{\Theta}$  (Greek theta for temperature) to distinguish them from the  $BX2^*$  lines converted by maternally inherited piRNAs (*de Vanssay et al., 2012*). Taken together, our data show that  $BX2^{OFF}$  can be functionally converted by high temperature (29°C), strongly suggesting that *de novo* piRNA production can occur in the absence of maternal inheritance of homologous piRNAs.

### **$BX2$ lines converted by high temperature or by maternal homologous piRNA inheritance present identical functional and molecular properties**

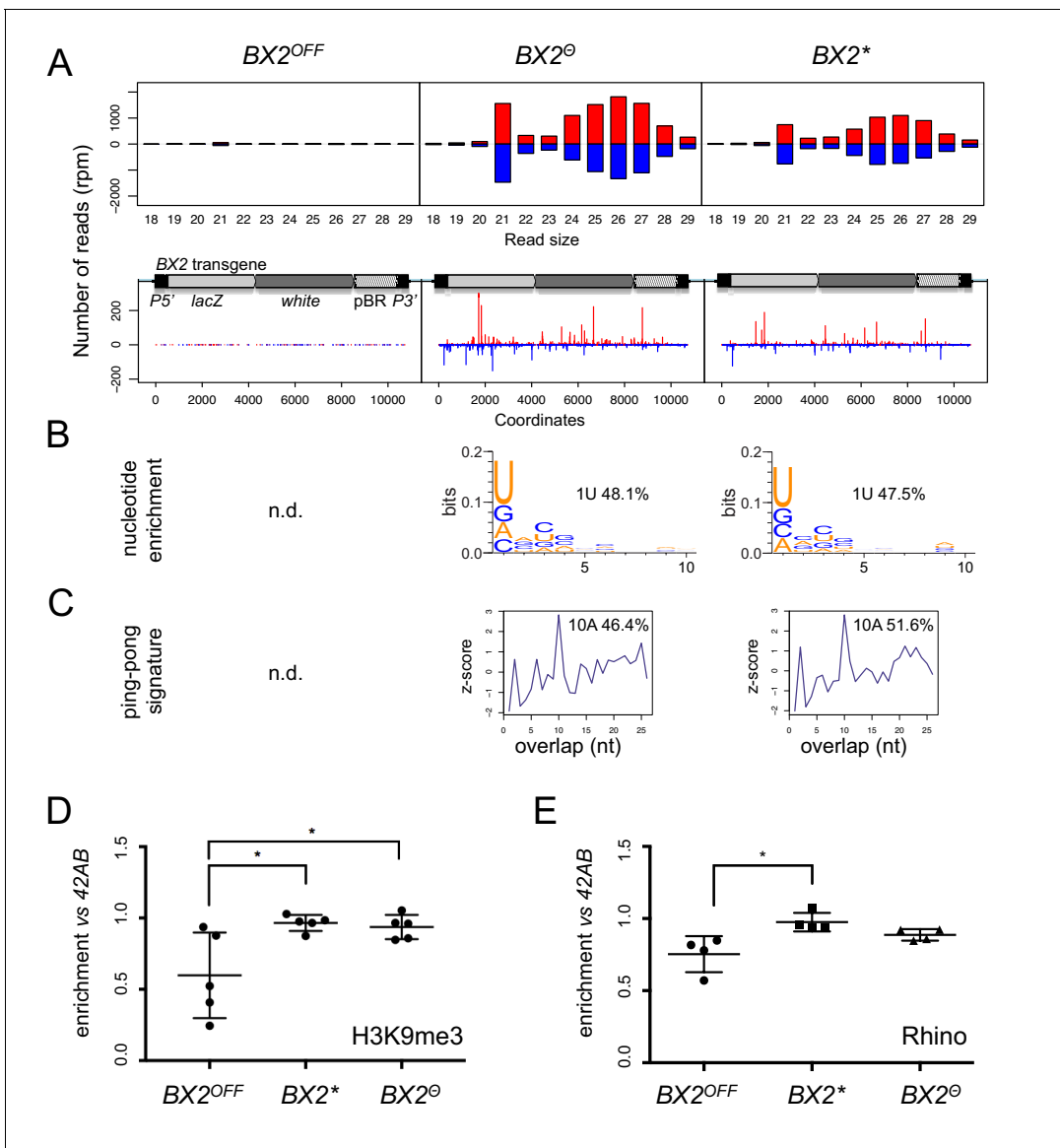
We further characterized the functional and molecular properties of  $BX2^{\Theta}$  activated by temperature and compared them to  $BX2^*$  activated by maternal inheritance of homologous piRNAs. Firstly, we compared the maternal and paternal  $BX2$  locus inheritance effect of three  $BX2^{\Theta}$  lines and three  $BX2^*$  at 25°C (**Figure 1—figure supplement 2A**). Maternal inheritance of either  $BX2^{\Theta}$  or  $BX2^*$  loci leads to complete and stable repression of  $\beta$ -Galactosidase expression (n flies = 152 and 159, respectively, **Supplementary file 3**), whereas paternal inheritance of either  $BX2^{\Theta}$  or  $BX2^*$  loci, that is in absence of maternal piRNA deposition, results in  $\beta$ -Galactosidase expression, and thus a definitive loss of  $BX2$  silencing capacities (n flies = 156 and 155, respectively, **Supplementary file 3**). Secondly, we previously showed that progeny with a paternally inherited  $BX2^{OFF}$  locus and maternally inherited  $BX2^*$  piRNAs, but lacking the maternally  $BX2^*$  genomic locus, have 100% conversion (*de Vanssay et al., 2012*). This process of recurrent conversions of an allele that is heritable without DNA modification is known as paramutation, thus  $BX2^*$  females are paramutagenic, that is able to trigger paramutation. To test this property on  $BX2$  lines converted by temperature,  $BX2^{\Theta}$  females were crossed with  $BX2^{OFF}$  males. The progeny that inherited the paternal  $BX2^{OFF}$  locus but not the maternal  $BX2^{\Theta}$  locus were selected and three independent lines were established (**Figure 1—figure supplement 2B**). Silencing measured over 20 generations revealed 100% of repression capacity showing that  $BX2^{\Theta}$  is also paramutagenic (n flies = 159, **Supplementary file 4**).

To determine whether the silencing capacities of  $BX2^{\Theta}$  involved piRNAs, small RNAs from  $BX2^{OFF}$ ,  $BX2^{\Theta}$  and  $BX2^*$  ovaries were extracted and sequenced (**Supplementary file 5**). Unique reads matching the  $P(lacW)$  sequences were identified only in the  $BX2^{\Theta}$  and  $BX2^*$  libraries (**Figure 2A**). Most of these small RNAs display all of the characteristics of *bona fide* germline piRNAs, that is a high proportion of 23–29 nt with a strong U bias on the first 5' nucleotide, an enrichment of a 10 nucleotide overlap between sense and antisense piRNAs, also known as the ping-pong signature, and a high proportion of reads with A at the tenth position among the 10 nt overlapped reads (**Figure 2B–C**) (*Brennecke et al., 2007*; *Gunawardane et al., 2007*). As a control, the 42AB piRNA cluster, a canonical germline dual-strand piRNA cluster, presented no significant difference between the three genotypes (**Figure 2—figure supplement 1**). Therefore, these results show that high temperature can initiate piRNA production from  $BX2$  naive sequences ( $BX2^{OFF}$ ) and strongly suggest that once a piRNA cluster is activated for piRNA production, the 'ON' state is maintained at each generation by maternal inheritance of piRNAs.

$BX2$  is inserted into the first intron of the  $AGO1$  gene (*de Vanssay et al., 2012*) and we looked at the piRNA production from this region in the different  $BX2$  epigenetic contexts. No significant amount of piRNAs coming from the  $AGO1$  gene region can be detected whatever the  $BX2$  state (**Figure 2—figure supplement 2**). These findings indicate that the  $AGO1$  gene region is not a natural piRNA cluster. To test whether other non-piRNA producing genomic loci have started to produce piRNAs following high temperature treatment, we looked for specific piRNAs (23–29 nt) matching at unique positions on *Drosophila* chromosomes and compared them between  $BX2^{\Theta}$  and  $BX2^{OFF}$ . The reads were then resampled per 50 kilobases windows. To eliminate background noise, only regions that produced more than five piRNAs per kilobase on average in both libraries were considered. Only exons of the *white* gene present in the  $P(lacW)$  transgenes of  $BX2$  showed differential piRNA expression (log2 ratio >8.5, **Figure 2—figure supplement 3**). This analysis revealed that the activation of piRNA production after thermic treatment is restricted to the  $BX2$  locus, suggesting that all other loci able to produce piRNAs are already active.

Previous studies had suggested that the chromatin state plays a role in the differential activity of  $BX2$  (*Le Thomas et al., 2014*). We therefore profiled H3K9me3 marks and Rhino binding on the  $P(lacW)$  transgene in ovaries from  $BX2^{OFF}$ ,  $BX2^{\Theta}$  and  $BX2^*$  strains by chromatin immunoprecipitation (ChIP) followed by quantitative PCR (qPCR). In both strains  $BX2^{\Theta}$  and  $BX2^*$ , H3K9me3 and Rhino





**Figure 2.** *BX2*<sup>O</sup> and *BX2*<sup>\*</sup> produce piRNAs and are enriched in H3K9me3. (A) Size distribution of ovarian small RNAs matching *BX2* transgene sequences reveals that both *BX2*<sup>O</sup> and *BX2*<sup>\*</sup> but not *BX2*<sup>OFF</sup> produce 21 nt siRNAs and 23–29-nt piRNAs (upper panels, pBR = backbone plasmid pBR322). Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. Unique 23–29 nt mappers are shown on the *BX2* transgene sequences (lower panels). (B) Percentage of 23–29 nt small RNAs from *BX2*<sup>O</sup> and *BX2*<sup>\*</sup> matching transgene sequence with a U at the first position are shown. n.d.: not determined due to low number of reads. (C) Relative frequency (z-score) of overlapping sense-antisense 23–29 nt RNA pairs reveals an enrichment of 10 nt overlapping corresponding to the ping-pong signature. (D) H3K9me3 and (E) Rhino binding on the *BX2* transgene in ovaries of *BX2*<sup>OFF</sup>, *BX2*<sup>O</sup> and *BX2*<sup>\*</sup> strains revealed by chromatin immunoprecipitation (ChIP) followed by quantitative PCR (qPCR) on specific *white* sequences. In both ‘ON’ strains, *BX2*<sup>O</sup> and *BX2*<sup>\*</sup>, H3K9me3 and Rhino levels over the transgene are very similar and higher than in the *BX2*<sup>OFF</sup> strain (unpaired *t*-test was used to calculate significance of the differences ( $p < 0.05$ ,  $n = 5$ )).

DOI: <https://doi.org/10.7554/eLife.39842.006>

The following figure supplements are available for figure 2:

**Figure supplement 1.** 42AB piRNA production is not altered in different *BX2* epigenetic contexts.

DOI: <https://doi.org/10.7554/eLife.39842.007>

**Figure supplement 2.** The *AGO1* gene region does not produce piRNA by itself.

DOI: <https://doi.org/10.7554/eLife.39842.008>

**Figure supplement 3.** Mapping the genome for regions stably producing new piRNAs.

DOI: <https://doi.org/10.7554/eLife.39842.009>

were similarly enriched over the *P(lacW)* transgene compared to the  $BX2^{OFF}$  strain, significantly for H3K9me3 (**Figure 2D–E**). Taken together, these results show that *de novo* activation of  $BX2^{OFF}$  by 29°C treatment ( $BX2^{\ominus}$ ) or paramutation by maternal inheritance of homologous piRNAs ( $BX2^*$ ) lines leads to similar functional and molecular properties.

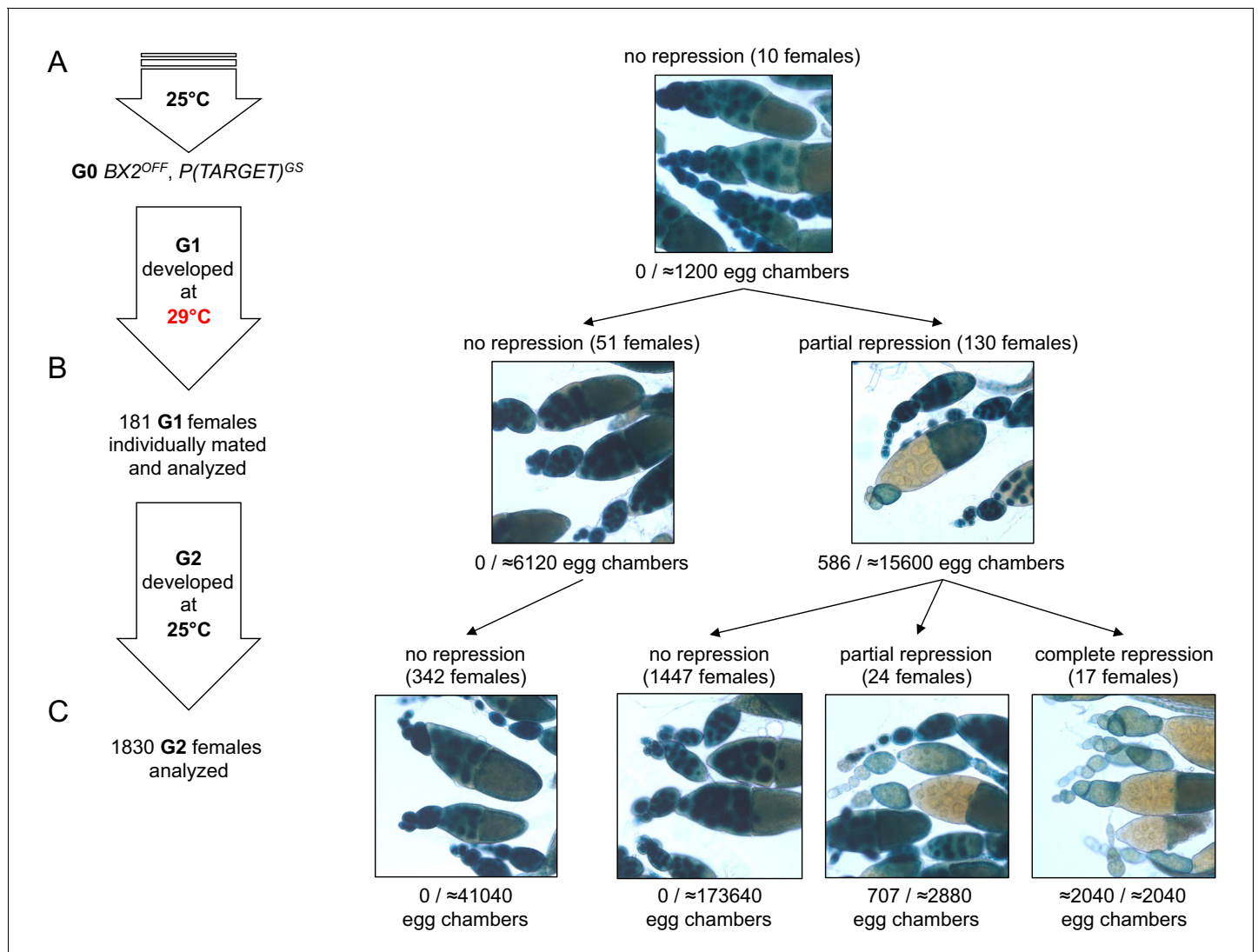
### Epigenetic conversion at 29°C occurs at a low rate from the first generation

To explain the low occurrence and the generational delay of  $BX2$  conversion at 29°C (**Supplementary file 1**), we propose that conversion is a complete but rare event occurring in a small number of egg chambers at each generation. Under this hypothesis, the sampling size of tested females should be crucial to observe such stochastic events. We therefore increased the number of analyzed females raised at 29°C during one generation. For this, eggs laid by females maintained at 25°C carrying the *P(TARGET)<sup>GS</sup>* reporter transgene and the  $BX2^{OFF}$  locus were collected during three days. These eggs were then transferred at 29°C until adults emerged (**Figure 3A**). To follow their offspring, we individually crossed 181 G1 females with two siblings and let them lay eggs for three days at 25°C. These 181 G1 females were then stained for  $\beta$ -Galactosidase expression. Strikingly, repression occurred only after one generation at 29°C in a few of the egg chambers of 130 G1 females ( $\approx 2.7\%$  of the estimated total number of G1 egg chambers  $n \approx 21700$ , **Figure 3B**, right panel). These results support our hypothesis whereby epigenetic conversion of  $BX2^{OFF}$  into  $BX2^{\ominus}$  is an instantaneous and complete event occurring at a low frequency per egg chamber and at each generation that is kept at 29°C.

To test the stability of the epigenetic  $BX2^{\ominus}$  states observed in G1 females, offspring daughters (G2) were raised at 25°C and their ovaries examined for  $\beta$ -Galactosidase expression. Among G2 females, partial ( $n = 24$ ) or complete ( $n = 17$ ) repression of  $\beta$ -Galactosidase expression in the germline was observed only in the progeny of those 130 G1 females in which partial repression was previously detected (**Figure 3C**). The proportion of 2.2% of converted G2 females (41/1830) is reminiscent with the proportion of repressed egg chambers observed in G1. The progeny of the 51 G1 females that did not present repression (**Figure 3B** left panel) did not show spontaneous conversion (**Figure 3C**). Taken together, these observations strongly suggest that newly converted  $BX2^{ON}$  egg chambers give rise to adult females with complete or partial silencing capacities. The low conversion rate observed in thousands of flies after one generation raised at high temperature and its stability through the next generation might explain the apparent delay of  $BX2^{ON}$  conversion of dozens of flies continuously raised at 29°C observed in the first set of experiments (see **Figure 1E** and **Supplementary file 1**). We more finely analyzed the silencing capacities of eight independent  $BX2^{OFF}$  lines throughout generations at 29°C (**Supplementary file 1**) by monitoring *P(TARGET)<sup>GS</sup>* repression in each egg chamber instead of whole ovaries (**Supplementary file 6**).  $BX2$  conversion occurred in each tested line with various dynamics (**Figure 1—figure supplement 3A**), likely reflecting a genetic drift due to the low conversion occurrence coupled to an important sampling effect at each generation. Globally, the mean of repression frequencies seems to indicate a progressive increase of  $BX2$  conversion over generations (**Figure 1—figure supplement 3B**). Altogether, our results illustrate how environmental modifications like high temperature experienced during one generation might stably modify the epigenome of the future ones. Such a newly acquired epigenetic state may spread in a given population within a few generations.

### High temperature increases $BX2$ antisense RNA but not piRNAs

Previously, we showed that  $BX2^{OFF}$  and  $BX2^{ON}$  produce similar amounts of sense and antisense transcripts (*de Vanssay et al., 2012*). However, these transcripts do not lead to  $\beta$ -Galactosidase expression in the germline nor piRNA production in the  $BX2^{OFF}$  line. We wondered if high temperature might change the RNA steady-state level of  $BX2$ . To ensure that we were detecting RNA specifically from  $BX2$ , qRT-PCR experiments targeting the *lacZ* gene were carried out on ovarian RNA extracted from the  $BX2^{OFF}$  line that did not contain *P(TARGET)<sup>GS</sup>*. We observed a significant increase in the steady-state  $BX2$  RNA levels at 29°C compared to 25°C (**Figure 4A**). Remarkably, strand-specific qRT-PCR experiments revealed that only  $BX2$  antisense transcripts, corresponding to antisense *lacZ* transcripts, increase at 29°C (**Figure 4B**). As  $BX2$  is inserted into the *AGO1* gene in a convergent transcription manner (*de Vanssay et al., 2012*), **Figure 4C**), we compared *AGO1* steady-state RNA



**Figure 3.**  $BX2$  conversion at 29°C occurs in one generation at a low rate. (A) G0 females carrying the  $P(TARGET)^{GS}$  reporter and  $BX2^{OFF}$  laid eggs at 25°C during three days. The  $BX2^{OFF}$  state of these females was confirmed after the three days at 25°C by  $\beta$ -Galactosidase staining (number of egg chambers  $\geq 1200$ ). (B) Their eggs were allowed to develop at 29°C until emergence of the next generation. G1 females ( $n = 181$ ) were individually mated with two siblings and left to lay for three days at 25°C. G1 females were then individually stained for  $\beta$ -Galactosidase expression. Strikingly, 130 females (71.8%) show  $\beta$ -Galactosidase repression in some egg chambers (586 among  $\approx 21700$  - estimation of the total egg chamber number among 181 females). The  $BX2^{OFF}$  into  $BX2^{ON}$  conversion frequency is  $\approx 2.7\%$ . (C) Analysis of each G1 female progeny developed at 25°C by  $\beta$ -Galactosidase staining. The progeny of the 51 G1 females that did not present repression maintained  $BX2^{OFF}$  state ( $n$  flies = 342). Most of the progeny of the 130 G1 females presenting conversion show no repression (97.2%,  $n$  flies = 1488) while 41 females present partial ( $n = 24$ ) or complete ( $n = 17$ ) repression of the germline expression of  $\beta$ -Galactosidase.

DOI: <https://doi.org/10.7554/eLife.39842.010>

The following figure supplements are available for figure 3:

**Figure supplement 1.** No  $BX2$  conversion at 29°C without the  $P(TARGET)^{GS}$  reporter transgene.

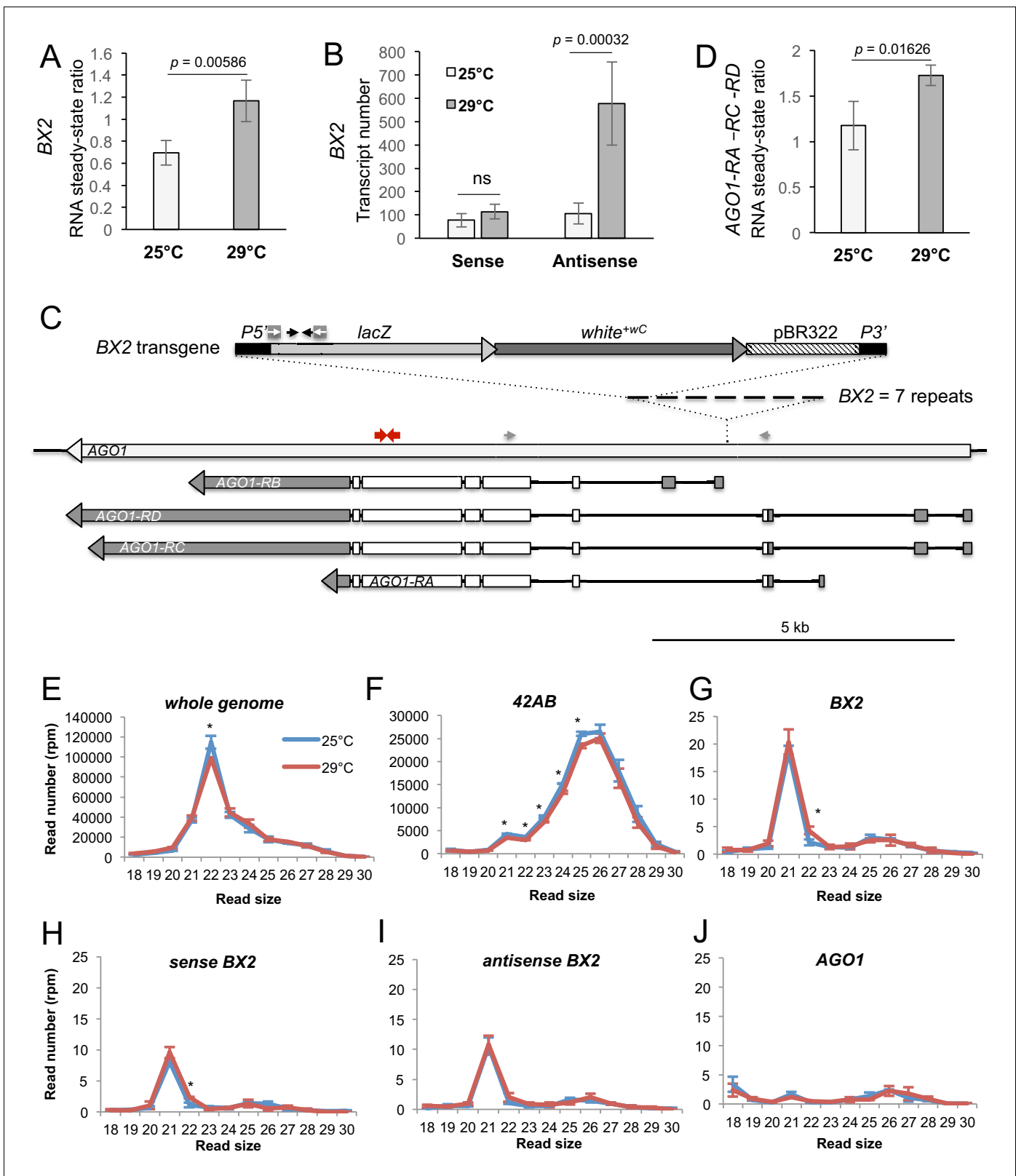
DOI: <https://doi.org/10.7554/eLife.39842.011>

**Figure supplement 2.** No  $BX2$  conversion at 29°C with a transgene that is not expressed in the germline.

DOI: <https://doi.org/10.7554/eLife.39842.012>

**Figure supplement 3.**  $BX2^{ON}$  silencing depends on *moonshiner*.

DOI: <https://doi.org/10.7554/eLife.39842.013>



**Figure 4.**  $BX2^{OFF}$  antisense RNA increase at 29°C. (A) RT-qPCR experiments revealed that the steady-state level of ovarian *lacZ* RNAs from BX2 is more abundant at 29°C (n = 5) compared to 25°C (n = 6). (B) Sense-specific RT-qPCR experiments revealed that only antisense transcripts from BX2, corresponding to antisense *lacZ* transcripts, are increased (25°C n = 6, 29°C n = 4). Significant p-values are given (bilateral Student's t-test). ns: not significant. (C) Map of the BX2 locus containing seven  $P(lacW)$  transgenes inserted into the AGO1 gene.  $P(lacW)$  and AGO1 are drawn to scale. The *lacZ* Figure 4 continued on next page

Figure 4 continued

gene contained in *P(lacW)* and *AGO1* are transcriptionally in opposite directions. Black arrows show *lacZ* primers used for (A) and (B) experiments. White arrows show primers used for sense-specific reverse transcription in (B) experiment. Grey arrows show *AGO1* primers used for (D) experiment: these primers are specific for *AGO1* transcripts (RA, RC and RD) that originate from promoters located upstream the *BX2* insertion point and, thus, are potentially convergent to *BX2*. Red arrows show primers used to measure steady-state of all *AGO1* isoforms (see **Figure 4—figure supplement 1D**). (D) RT-qPCR experiments performed on flies devoid of *P(lacW)* transgenes (*w<sup>1118</sup>* context) revealed that the steady-state level of ovarian *AGO1*-RA, -RC and -RD RNA isoforms is more abundant at 29°C (n = 5) compared to 25°C (n = 6). (E–I) To compare small RNAs at 25 versus 29°C, total RNAs were extracted from *BX2<sup>OFF</sup>* ovaries dissected from adults incubated at 25°C or 29°C. Three samples were tested for each temperature. Small RNAs from 18 to 30 nt were deep sequenced. For each library, normalization has been performed for 1 million reads matching the *Drosophila* genome (rpm, **Supplementary file 7**). Size distributions of unique reads that match reference sequences are given. (E) Small RNAs matching the *Drosophila* genome present similar profiles in both temperatures except for 22 nt RNAs that are more represented at 25°C. (F) The 21 to 25 nt reads matching the 42AB piRNA cluster that range from 21 to 25 nt are slightly more abundant at 25°C. (G) Strikingly, almost only 21 nt RNAs match *BX2* sequence. They are equally distributed among sense (H) and antisense (I) sequences at both temperatures. (J) No small RNAs corresponding to the *AGO1* gene can be detected whatever the temperature. \*= $p < 0.05$ , bilateral Student's t-test.

DOI: <https://doi.org/10.7554/eLife.39842.014>

The following figure supplement is available for figure 4:

**Figure supplement 1.** Complementary RNA steady-state measurements by quantitative RT-PCR.

DOI: <https://doi.org/10.7554/eLife.39842.015>

level at 29°C and 25°C. *AGO1* transcript isoforms that are initiated upstream the *BX2* insertion point are significantly increased at 29°C (**Figure 4D** and **Figure 4—figure supplement 1A–B**). Thus, it is possible that, at 29°C, an increase of transcription from the *AGO1* promoters located upstream the *BX2* insertion point could lead to an increase of *BX2* antisense RNA transcription.

We then examined whether the increase of *BX2* antisense RNAs leads to an increase of antisense small RNAs. Ovarian small RNAs (18 to 30 nucleotides) of *BX2<sup>OFF</sup>* flies (without *P(TARGET)<sup>GS</sup>*) raised at 25°C and at 29°C for one generation were sequenced and the read numbers normalized (**Supplementary file 7**). A slight, yet statistically significant, decrease is observed at 29°C for small RNAs matching the whole genome (**Figure 4E**) and for the 42AB piRNA cluster (**Figure 4F**). Strikingly, no piRNAs were produced from the *BX2* locus at 25°C nor after one generation at 29°C. Thus, the increase of *BX2* antisense transcripts observed at 29°C (**Figure 4B**) did not correlate with an increase of corresponding antisense piRNAs. At 25°C and 29°C, *BX2<sup>OFF</sup>* produced the same low amount of 21 nt small RNAs, equally distributed between sense and antisense (**Figure 4G,H,I**), suggesting that *BX2* transcripts are processed into siRNAs. These results confirm that *AGO1* is not a piRNA producing locus (as shown in **Figure 2—figure supplement 2**) and showed that, at 29°C, *AGO1* is still not producing small RNAs (**Figure 4J**). These data indicate that 21 nt small RNA production was restricted to *BX2* sequences.

## Heat conversion requires a transcribed homologous sequence in trans

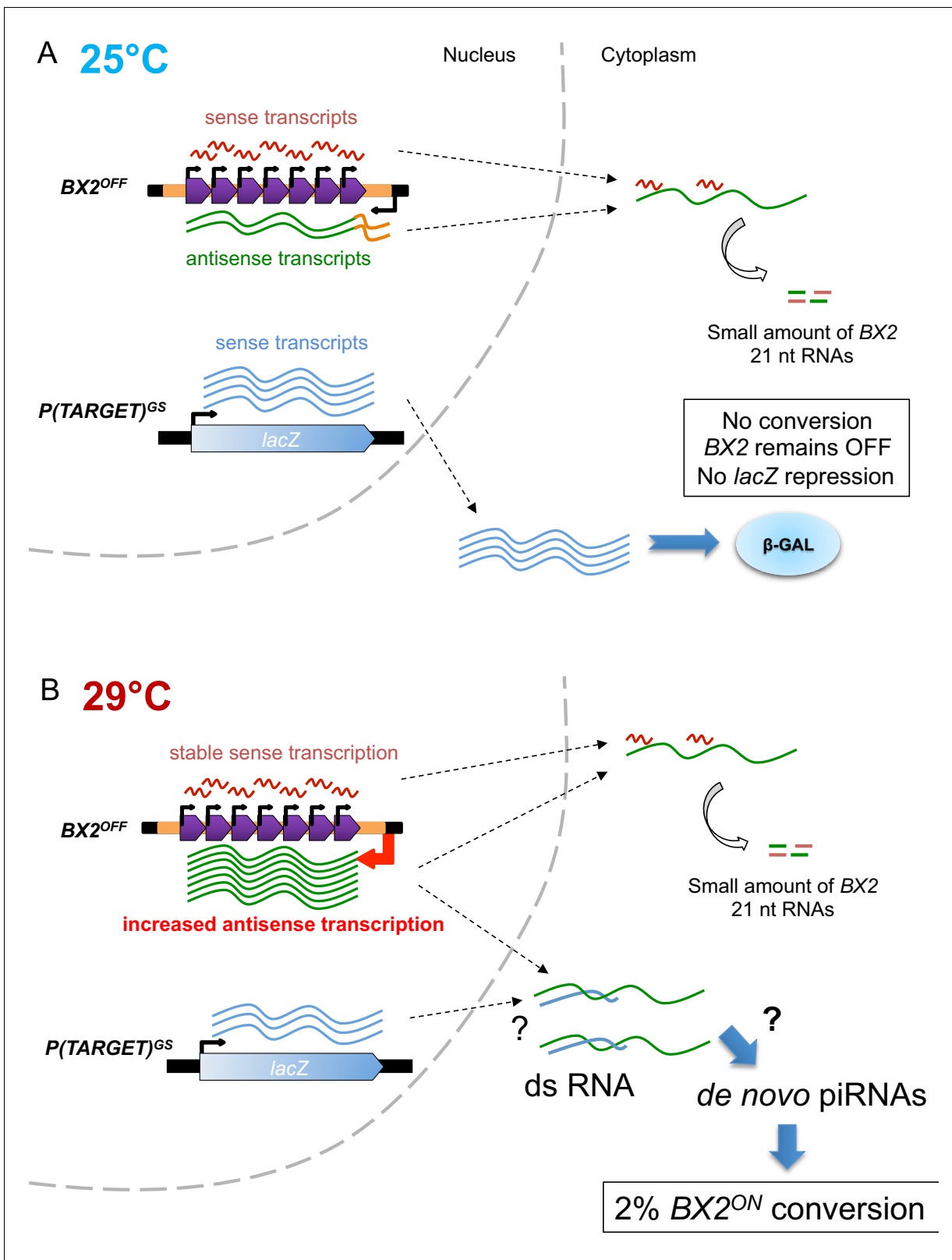
Quantitative RT-PCR experiments described above were carried out on flies bearing only the *BX2* locus while all conversion experiments at 29°C were done with flies bearing the *BX2* and the *P(TARGET)<sup>GS</sup>* locus. Interestingly, the amount of *P(TARGET)<sup>GS</sup>* transcripts is affected by temperature but in the opposite way to *BX2*, as less transcripts were measured at 29°C compared to 25°C (**Figure 4—figure supplement 1C**). We next asked whether the *P(TARGET)<sup>GS</sup>* transgene could participate in the conversion process of *BX2*. For this, 'heat-activated-conversion' experiments of *BX2* were done in flies not carrying the *P(TARGET)<sup>GS</sup>* (**Figure 3—figure supplement 1**). To assess the *BX2* epigenetic state of the G1 raised at 29°C, 157 G1 females were individually crossed at 25°C with males harboring the *P(TARGET)<sup>GS</sup>* transgene. Among the 1137 G2 females analyzed, only one female presented partial repression of the  $\beta$ -Galactosidase expression and none presented complete repression (**Figure 3—figure supplement 1**). If we compare these results with those obtained with the *BX2*, *P(TARGET)<sup>GS</sup>* lines (**Figure 3**), the difference was highly significant ( $p=8.5 \times 10^{-6}$ , homogeneity  $\chi^2 = 23.35$  with 2 degrees of freedom, **Supplementary file 8**). To further validate the requirement of the euchromatic homologous transgene *P(TARGET)<sup>GS</sup>* in establishing the temperature-dependent *BX2* conversion, we generated eight independent lines in which *BX2* was recombined into the same *P(TARGET)<sup>GS</sup>* genetic background but without the *P(TARGET)<sup>GS</sup>* transgene. After 30 generations at 29°C, no female showing  $\beta$ -Galactosidase repression was observed (**Supplementary file 9**). A

homogeneity  $\chi^2$  test comparing, at G13, the repression occurrence in *BX2*,  $P(\text{TARGET})^{\text{GS}}$  lines (31/161, **Supplementary file 1**) and in recombined *BX2* lines (0/975, **Supplementary file 9**) is highly significant ( $p=7.04 \times 10^{-44}$ , homogeneity  $\chi^2 = 192.9$  with 1 degree of freedom), arguing against a background effect of the  $P(\text{TARGET})^{\text{GS}}$  line in the conversion phenomenon. Additionally, we reproduced the experiment with an euchromatic  $P(\text{TARGET})$  expressed only in the germline and referred to hereinafter as ' $P(\text{TARGET})^{\text{G}}$ ' (**Figure 1—figure supplement 1A–B**). In the presence of  $P(\text{TARGET})^{\text{G}}$ , *BX2* was converted at 29°C at a rate comparable with that observed with  $P(\text{TARGET})^{\text{GS}}$  (**Supplementary file 10**). These data show that *BX2* conversion by temperature cannot be attributed to any specificity linked to the  $P(\text{TARGET})^{\text{GS}}$  insertion. To know if the transcription of the  $P(\text{TARGET})^{\text{GS}}$  (or of the  $P(\text{TARGET})^{\text{G}}$ ) is required, we reproduced the same experiment with another euchromatic transgene that is not expressed in the germline but in the somatic cells surrounding the germ cells and therefore referred to hereinafter as ' $P(\text{TARGET})^{\text{S}}$ ' (**Figure 1—figure supplement 1A–B**). In the presence of  $P(\text{TARGET})^{\text{S}}$ , *BX2* was not converted at 29°C (0/784, **Figure 3—figure supplement 2**). A homogeneity  $\chi^2$  test comparing these results with those obtained with  $P(\text{TARGET})^{\text{GS}}$  considering only the complete *BX2*<sup>ON</sup> G2 females (17/1464 females, **Figure 3**) is significant ( $p=0.0055$ , homogeneity  $\chi^2 = 7.7$  with 1 degree of freedom). We conclude that the transcription of a reporter transgene sharing homologous sequences (*i.e.* *lacZ*) with *BX2* is required in the germline for *BX2* conversion at 29°C.

To summarize, in the absence of  $P(\text{TARGET})$  sequences, at 29°C *BX2*<sup>OFF</sup> produces an elevated number of antisense transcripts, no piRNAs and is unable to be converted to *BX2*<sup>ON</sup>. In contrast, when a  $P(\text{TARGET})$  is present and transcribed in the germline, *BX2* conversion and piRNA production are observed at 29°C. Although much of the mechanistic aspects of *BX2* conversion remain unknown, these findings lead us to propose that, at 29°C, double strand RNA (dsRNA) made of the excess of *BX2* antisense transcripts and the sense  $P(\text{TARGET})$  transcripts could be a prerequisite for the production of *de novo* piRNAs and the conversion of *BX2* into an active piRNA cluster (see recapitulative model **Figure 5**).

## Discussion

Here, we report on the heritable establishment of a new piRNA cluster associated with silencing properties induced by high temperature during development. The epigenetic response to heat exposure has been studied in several model species: in *Arabidopsis* for instance, increasing temperature induces transcriptional activation of repetitive elements (*Ito et al., 2011; Pecinka et al., 2010; Tittel-Elmer et al., 2010*). Whether these changes involve chromatin modifications is not clear but none of these modifications have been found to be heritable through generations except in mutants for siRNA biogenesis where high frequency of new TE insertions was observed in the progeny of stressed plants (*Ito et al., 2011*). In animals, response to heat can result in modification of DNA methylation at specific loci in reef building coral (*Dimond and Roberts, 2016*), chicken (*Yossifoff et al., 2008*) and wild guinea pigs (*Weyrich et al., 2016*). In the latter, modifications affecting  $\approx 50$  genes are inherited in G1 progeny (*Weyrich et al., 2016*). The mechanisms of this heritability, however, are not yet understood. In *Drosophila*, heat-shock treatment of 0–2 hr embryos for one hour at 37°C or subjecting flies to osmotic stress induce phosphorylation of dATF-2 and its release from heterochromatin (*Seong et al., 2011*). This defective chromatin state is maintained for several generations before returning to the original state. We have tested if such stresses were able to convert *BX2*<sup>OFF</sup> into *BX2*<sup>ON</sup> in one generation but neither heat-shock nor osmotic stress induces *BX2* conversion (**Supplementary file 11**), suggesting that *BX2* activation does not depend on dATF-2. In a more recent paper, *Fast et al.* (*Fast et al., 2017*) found less piRNAs at 29°C than at 18°C. However, RNAseq analyses of differentially expressed genes involved in the piRNA pathway were not conclusive, as some piRNA genes (*ago3, aub, zuc, armi*) were more expressed at 29°C while others (*shu, hsp83, Yb*) were less expressed as compared to the levels at 18°C (*Fast et al., 2017*). Overall, the enhancement of the piRNA ping-pong amplification loop observed at 29°C was attributed to RNA secondary structures, because of a lack of specificity for any particular class of TE (*Fast et al., 2017*). Furthermore, in contrast to our RT-qPCR results obtained at 25°C and 29°C (**Figure 4D** and **Figure 4—figure supplement 1A–B**), AGO1 was not differentially expressed between 18°C and 29°C (*Fast et al., 2017*). This difference can be explained, because, in our experiments, we checked for specific spliced transcripts of AGO1 (RA, RC and RD) that originate upstream



**Figure 5.** Model of *BX2* activation at 29°C. (A) At 25°C, a low bidirectional transcription of the *BX2* cluster leads to a small production of 21 nt RNAs. *BX2<sup>OFF</sup>* is stable at 25°C, no conversion event is observed and the *lacZ* reporter gene from *P(TARGET)<sup>GS</sup>* remains active throughout generations. (B) At 29°C, a specific increase in antisense transcription occurs, presumably due to a higher activity of the promoter of the *AGO1* gene (orange box). This excess of *BX2* antisense RNA could interact with sense *P(TARGET)<sup>GS</sup>* transcripts to produce double stranded RNA. Through a yet unknown mechanism, Figure 5 continued on next page

Figure 5 continued

such dsRNA could lead to the formation of *de novo* BX2 piRNAs. These piRNAs could in turn trigger the conversion of BX2<sup>OFF</sup> into an active piRNA cluster, a phenomenon observed based on the repression of the *lacZ* reporter gene of *P(TARGET)*<sup>GS</sup>. The BX2 conversion is a rare event ( $\approx 2\%$  per generation) but once achieved, BX2<sup>ON</sup> remains active throughout generations due to the maternal inheritance of homologous piRNAs and the paramutation of the paternal BX2<sup>OFF</sup> allele.

DOI: <https://doi.org/10.7554/eLife.39842.016>

the BX2 insertion point (Figure 4C). In agreement with Fast *et al.*, no statistical difference in the global amount of AGO1 transcripts between 25°C and 29°C is observed using primers located downstream the BX2 insertion point (Figure 4—figure supplement 1D). This suggests that the AGO1 promoter located downstream the BX2 insertion point, and responsible for the production of the RB isoform (see Figure 4), is not sensitive to temperature. Taken together, these observations show that temperature modification might induce epigenetic changes in several species but the underlying mechanisms remain largely unsolved.

Whole genome comparison of small RNA sequencing between BX2<sup>OFF</sup> and newly BX2<sup>ON</sup> heat-converted flies (BX2<sup>o</sup>) did not reveal additional regions stably converted for piRNA production (Figure 2—figure supplement 3). This suggests that no other loci are metastable for piRNA production, that is all potential piRNA clusters are already active at 25°C. This observation raises the question of what makes BX2 locus competent for piRNA activation at high temperature. The BX2 locus is the result of successive induced transpositions of *P(lacW)* used to screen for *white-variegating* phenotype (Dorer and Henikoff, 1994). Thus, BX2 resembles natural TE clusters where TEs have the capacity to transpose into each other assembling structures named *nested TE*, as described in numerous genomes (Gao *et al.*, 2012; Liu *et al.*, 2007; Zanni *et al.*, 2013). Some *nested TE* loci might have the capacity to be activated and respond to a new TE invasion. In *Drosophila*, BX2-like tandemly inserted transgenes were shown to be new sites of HP1 enrichment in larval salivary glands, emphasizing a heterochromatic structure at the BX2 locus in somatic cells (Fanti *et al.*, 1998), and to cause pairing-dependent silencing (Dorer and Henikoff, 1997). When tested for repressing capacities, however, this strain is inactive for BX2 piRNA production (de Vanssay *et al.*, 2012; Josse *et al.*, 2008). We have shown that maternally inherited *P(lacW)* piRNAs are able to paramutate with complete and stable penetrance from an inactive BX2 locus into an active locus for piRNA production (de Vanssay *et al.*, 2012). The paramutated BX2 locus appeared to be a genuine piRNA cluster since it is sensitive to a number of factors known to be involved in piRNA biogenesis such as *aub*, *rhi*, *cuff*, *zuc* (Hermant *et al.*, 2015) and *moonshiner* (Figure 3—figure supplement 3). The number of transgene copies appears to be crucial in the process since smaller number of transgenes results in less somatic heterochromatinization (Dorer and Henikoff, 1994; Fanti *et al.*, 1998), less pairing-dependent silencing (Dorer and Henikoff, 1997) and unstable paramutation (de Vanssay *et al.*, 2012). Taken together, these data suggest that the heterochromatic structure of a cluster precedes piRNA production. This is supported by our ChIP experiments showing that H3K9me3 levels on BX2<sup>OFF</sup> are slightly below the H3K9me3 level of piRNA producing states (BX2<sup>ON</sup> and BX2<sup>o</sup>, Figure 2D). The same observation can be made for Rhino (Figure 2E), suggesting that Rhino may be already present on the BX2<sup>OFF</sup> locus but below the threshold required for piRNA production as suggested by Akulenko *et al.* (2018). Thus, a locus made of repeated sequences and being likely heterochromatic (H3K9me3, Rhino) is a necessary but not sufficient condition to specify an active piRNA cluster.

In the germline, piRNA clusters produce piRNAs from both strands and it was recently shown that, in most cases, transcription initiates within clusters on both strands through the interaction of Rhino and Moonshiner (Andersen *et al.*, 2017). In few cases, however, piRNA cluster transcription may take advantage of the read-through from a flanking promoter (Andersen *et al.*, 2017). Zhang *et al.* (Zhang *et al.*, 2014) have shown that tethering Rhino onto a transgene leads to its repression but the production of piRNA depends on the presence of another transgene producing antisense RNA. Moreover, in the context of the *Pld* promoter deletion, a gene flanking the 42AB piRNA cluster, flies can produce *Pld* piRNAs only if a *Pld* cDNA is expressed in trans (Andersen *et al.*, 2017). From all of these observations, a model emerges predicting that simultaneous production of sense and antisense RNA is a shared requirement for piRNA production. However, even if BX2<sup>OFF</sup> is transcribed on both strands, without additional signals, it still remains inactive for piRNA production.



In addition to having a number of heterochromatic repeats and a double stranded transcription, the production of *de novo* piRNAs from BX2 requires a triggering signal. From our experiments, BX2 conversion relies on the simultaneous increase of both sense and antisense RNAs. An active role of euchromatic copies in the establishment of new piRNA clusters by high temperature appears to be consistent with what would naturally happen during the invasion of a naive genome by new TEs or when chromosomal breakages occur leading to the loss of piRNA cluster loci (*Asif-Laidin et al., 2017*). At first, uncontrolled euchromatic TE transposition takes place before the establishment of repression. Such repression would occur after a copy integrates into a preexisting piRNA cluster or by the generation of a new cluster made by successive insertion of nested copies. Consequently, clusters of elements cannot exist without transcriptionally active euchromatic copies. The increase of germline antisense transcripts upon stress or environmental factors, depending on the neighboring genomic environment, and the concomitant presence of numerous sense transcripts from euchromatic active copies, appear to be the starting signals for new piRNA production. These piRNAs can then be inherited at the next generation where they will stably paramutate the corresponding DNA locus with repetitive nature. At that time, the triggering signal is no longer necessary since BX2 remains activated once flies get back at 25°C. Future generations thus remember what was once considered a threat only through the legacy of maternal piRNAs.

## Materials and methods

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Gene ( <i>Drosophila melanogaster</i> )	AGO1	NA	FLYB: FBgn0262739	
Gene ( <i>D. melanogaster</i> )	RpL32	NA	FLYB: FBgn0002626	
Gene ( <i>D. melanogaster</i> )	eEF5	NA	FLYB: FBgn0285952	
Gene ( <i>D. melanogaster</i> )	Moonshiner	NA	FLYB: FBgn0030373	
Strain, strain background ( <i>D. melanogaster</i> )	w1118	Laboratory Stock	FLYB: FBal0018186	
Strain, strain background ( <i>D. melanogaster</i> )	BX2	<i>Dorer and Henikoff, 1994</i> PMID:8020105	FLYB: FBti0016766	
Strain, strain background ( <i>D. melanogaster</i> )	P(TARGET) <sup>GS</sup>	Bloomington Drosophila Stock Center	FLYB: FBst0011039	also called P-1039
Strain, strain background ( <i>D. melanogaster</i> )	P(TARGET) <sup>G</sup>	Bloomington Drosophila Stock Center	FLYB: FBti0003435	also called BQ16
Strain, strain background ( <i>D. melanogaster</i> )	P(TARGET) <sup>S</sup>	Bloomington Drosophila Stock Center	FLYB: FBti0003418	also called BA37
Strain, strain background ( <i>D. melanogaster</i> )	nosGAL4	Bloomington Drosophila Stock Center	FLYB: FBti0131635, RRID:BDSC_32180	
Genetic reagent ( <i>D. melanogaster</i> )	P(lacW)	PMID: 2558049	FLYB: FBtp0000204	
Genetic reagent ( <i>D. melanogaster</i> )	P(PZ)	<i>Mlodzik and Hiromi, 1992</i> doi: 10.1016/B978-0-12-185267-2.50030-1	FLYB: FBtp0000210	
Genetic reagent ( <i>D. melanogaster</i> )	P(A92)	PMID: 2827169	FLYB: FBtp0000154	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent ( <i>D. melanogaster</i> )	Moon shRNA PA61	Andersen et al., 2017 doi:10.1038/nature23482		Dr. Julius Brennecke (Institute of Molecular Biotechnology, Vienna)
Genetic reagent ( <i>D. melanogaster</i> )	Moon shRNA PA62	Andersen et al., 2017 doi:10.1038/nature23482		Dr. Julius Brennecke (Institute of Molecular Biotechnology, Vienna)
Antibody	Mouse IgG polyclonal antibody	Merck (ex-Millipore)	Cat# 12-371B, RRID:AB_2617156	
Antibody	Rabbit polyclonal antibody against H3K9me3	Merck (ex-Millipore)	Cat# 07-442	
Antibody	Rabbit polyclonal antibody against Rhino	PMID: 19732946		Dr. William Theurkauf (University of Massachusetts Medical School, Worcester)
Sequence-based reagent	RT-qPCR primers	Sigma-Aldrich		
Sequence-based reagent	RT-qPCR primers	Eurogentech		
Commercial assay or kit	RNeasy kit	Qiagen	Cat# 74104	
Commercial assay or kit	Illumina TruSeq Small RNA library preparation kits	Fasteris	<a href="http://www.fasteris.com">http://www.fasteris.com</a>	
Commercial assay or kit	Revertaid RT	Thermo Scientific	EP0442	
Commercial assay or kit	Random Hexamers	Invitrogen	N8080127	
Commercial assay or kit	DNaseI (Rnase free)	New England Biolabs	M0303S	
Commercial assay or kit	dNTPs solution Mix	New England Biolabs	N0447S	
Commercial assay or kit	Ribolock RNA inhibitor	Thermo Scientific	EO0381	
Commercial assay or kit	Ssofast Evagreen Supermix	Biorad	Cat# 172-5204	
Commercial assay or kit	qPCR kit	Roche	Cat# 04887352001	
Chemical compound, drug	TRIzol	Invitrogen	Cat# 15596026	
Chemical compound, drug	Chloroform	Carlo Erba Reagents	Cat# 438601	
Chemical compound, drug	Chloroform	Sigma-Aldrich	C2432	
Chemical compound, drug	Ethanol (EtOH)	Merck millipore	Cat# 100983	
Chemical compound, drug	Ethanol (EtOH)	Honeywell	Cat# 32221	
Chemical compound, drug	Glycerol	VWR AnalaR NORMAPUR	Cat# 24388.295	
Chemical compound, drug	Glutaraldehyde	Sigma Aldrich	G-5882	
Chemical compound, drug	Potassium hexacyanoferrate(III)	Sigma Aldrich	P3667	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Chemical compound, drug	Potassium hexacyanoferrate(II) trihydrate	Sigma Aldrich	P3289	
Chemical compound, drug	X-Gal	Dutscher	Cat# 895014	
Chemical compound, drug	NaCl	VWR AnalaR NORMAPUR	Cat# 27810.295	
Chemical compound, drug	NaCl	Sigma-Aldrich	Cat# 31432	
Chemical compound, drug	Formaldehyde	Sigma-Aldrich	Cat# 252549	
Chemical compound, drug	Schneider Medium	Gibco	Cat# 21720-024	
Chemical compound, drug	Insulin	Sigma-Aldrich	I4011	
Chemical compound, drug	PBS	Ambion	AM9625	
Chemical compound, drug	Triton	Sigma-Aldrich	T8787	
Chemical compound, drug	KCl	Ambion	AM9640G	
Chemical compound, drug	HEPES	Fisher Scientific	BP299	
Chemical compound, drug	IPEGAL	Sigma-Aldrich	Cat# 18896	
Chemical compound, drug	DTT	Fisher Scientific	R0861	
Chemical compound, drug	Na Butyrate	Sigma-Aldrich	Cat# 07596	
Chemical compound, drug	EDTA free protease inhibitor	Roche	Cat# 04693159001	
Chemical compound, drug	N lauryl sarkosyl	Sigma-Aldrich	L-5125	
Chemical compound, drug	BSA	Fisher Scientific	BP9703	
Chemical compound, drug	SDS 20%	Euromedex	EU0660-B	
Chemical compound, drug	Tris HCl	Invitrogen	Cat# 15504-020	
Chemical compound, drug	Dynabeads A	Invitrogen	10002D	
Chemical compound, drug	Glycine	Sigma-Aldrich	G8898	
Chemical compound, drug	Isopropanol	VWR	Cat# 20842.298	
Software, algorithm	Galaxy Server	ARTBIO	<a href="https://mississippi.snv.jussieu.fr/">https://mississippi.snv.jussieu.fr/</a>	
Software, algorithm	Weblogo	<i>Crooks et al., 2004</i> doi:10.1101/gr.849004		

### Transgenes and strains

All transgenes are in the  $w^{1118}$  background. The BX2 line carries seven *P-lacZ-white* transgenes, (*P-lacW*), FBtp0000204) inserted in tandem and in the same orientation at cytological site 50C on the

second chromosome (*Dorer and Henikoff, 1994*). The transgene insertion site is located in an intron of the *AGO1* gene (*de Vanssay et al., 2012*). Homozygous individuals are rare and sterile and the stock is maintained in heterozygous state with a Cy-marked balancer chromosome.  $\beta$ -Galactosidase activity from these transgenes cannot be detected in the germline.  $P(TARGET)^{GS}$  corresponds to  $P(PZ)$  (FBtp0000210), a  $P$ -*lacZ*-*rosy* enhancer-trap transgene inserted into the *eEF5* gene at 60B7 and expressing  $\beta$ -Galactosidase in the germline and somatic cells of the female gonads (Bloomington stock number 11039 (FBst0011039)). Homozygous flies are not viable and the stock is maintained over a Cy-marked balancer chromosome.  $P(A92)$  (FBtp0000154) is another  $P$ -*lacZ*-*rosy* enhancer-trap transgene that has been used in this study:  $P(TARGET)^G$  corresponds to  $BQ16$  (FBti0003435) expressing *lacZ* only in the germline and  $P(TARGET)^S$  corresponds to  $BA37$  (FBti0003418) expressing *lacZ* only in the somatic follicle cells that surround the nurse cells. In both lines, homozygous flies are viable. The *nosGAL4* transgene used is from the  $w[*]$ ;  $PBac[w[+mW.hs]=GreenEye.nosGAL4]Dmel6$  line (FBti0131635). Modified miRNA against *moonshiner* (lines *PA61* and *PA62*) were a kind gift from Julius Brennecke (*Andersen et al., 2017*). Additional information about stocks are available at Flybase: 'http://flybase.bio.indiana.edu/'.

### Thermic and osmotic treatments

Since maintaining flies at high temperature (29°C) decreases viability, we used the following procedure at each generation: fertilized adult females (G0) were allowed to lay eggs for three days at 25°C on standard cornmeal medium. Adults were then discarded or tested for ovarian  $\beta$ -Galactosidase expression. Vials containing progeny were transferred at 29°C for the rest of the development until complete emergence of G1 adults. Young adults were then transferred into a new vial where they were allowed to lay eggs for three days at 25°C. For heat-shock treatment, embryos (0–2 hr) were incubated at 37°C for 1 hr as described in *Seong et al. (2011)* and then transferred at 25°C until adult emergence. For osmotic treatments, culturing flies on 300 mM NaCl, as described in *Seong et al. (2011)*, leads to either a large increased time of development or lethality and did not allow us to perform conversion measurements. Accordingly, flies were incubated on 150 mM NaCl for one generation before dissection and  $\beta$ -Galactosidase staining.

### $\beta$ -Galactosidase staining

Ovarian *lacZ* expression assays were carried out using X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) overnight staining at 37°C as previously described (*Lemaitre et al., 1993*), except that ovaries were fixed afterwards for 10 min. After mounting in glycerol/ethanol (50/50), the germline *lacZ* repression was then calculated by dividing the number of repressed egg chambers by the total number of egg chambers. Most of the time, the total number of egg chambers was estimated by multiplying the number of mounted ovaries by 60, corresponding to an average of three to four egg chambers per ovariole and 16 to 18 ovarioles per ovary. Images were acquired with an Axio-ApoTome (Zeiss) and ZEN2 software.

### Fly dissection and RNA extraction

For each genotype tested, 20 pairs of ovaries were manually dissected in 1X PBS. For small RNA sequencing, total RNA was extracted using TRIzol (Life Technologies) as described in the reagent manual ([http://tools.lifetechnologies.com/content/sfs/manuals/trizol\\_reagent.pdf](http://tools.lifetechnologies.com/content/sfs/manuals/trizol_reagent.pdf)). For the RNA precipitation step, 100% ethanol was used instead of isopropanol. For RT-qPCR experiments, total RNA was extracted using TRIzol for *BX2* and  $w^{1118}$  females or RNeasy kit (Qiagen) for  $P(TARGET)$  females. Up to six biological replicates were used for each genotype.

### Small RNA sequencing analyses

A small RNA fraction of 18 nt to 30 nt in length was obtained following separation of total RNA extracted from dissected ovaries on a denaturing polyacrylamide gel. This fraction was used to generate multiplexed libraries with Illumina TruSeq Small RNA library preparation kits (RS-200–0012, RS200-0024, RS-200–036 or RS-200–048) at *Fasteris* (<http://www.fasteris.com>). A house protocol based on TruSeq, which reduces 2S RNA (30 nt) contamination in the final library, was performed. Libraries were sequenced using Illumina HiSeq 2000 and 2500. Sequence reads in fastq format were trimmed from the adapter sequence 5'-TGAATTCTCGGGTGCCAAG-3' and matched to the *D.*

*melanogaster* genome release 5.49 using Bowtie (Langmead et al., 2009). Only 18–29 nt reads matching the reference sequences with 0 or one mismatch were retained for subsequent analyses. For global annotation of the libraries (Supplementary files 5 and 7), we used the release 5.49 of fasta reference files available in Flybase, including transposon sequences (dmel-all-transposon\_r5.49.fasta) and the release 20 of miRNA sequences from miRBase (<http://www.mirbase.org/>).

Sequence length distributions, small RNA mapping and small RNA overlap signatures were generated from bowtie alignments using Python and R (<http://www.r-project.org/>) scripts, which were wrapped and run in Galaxy instance publicly from ARTbio platform available at <http://mississippi.fr>. Tools and workflows used in this study may be downloaded from this Galaxy instance. For library comparisons, read counts were normalized to one million miRNA (Supplementary files 5 and 7). A second normalization, performed using the total number of small RNAs matching the *D. melanogaster* genome (release 5.49), gave similar results (Supplementary files 5 and 7). For small RNA mapping (Figures 2 and 4, Figure 2—figure supplements 1 and 2), we took into account only 23–29 nt RNA reads that uniquely aligned to reference sequences. Logos were calculated using Weblogo (Crooks et al., 2004) from 3' trimmed reads (23 nt long) matching either *P(lacW)* (Figure 2B) or 42AB (Figure 2—figure supplement 1B). The percentage of reads containing a 'U' at the first position was calculated with all 23–29 nt RNA matching the reference sequence (BX2 transgene in Figure 2B and 42AB in Figure 2—figure supplement 1B). Distributions of piRNA overlaps (ping-pong signatures, Figure 2C and Figure 2—figure supplement 1C) were computed as first described in Klattenhoff et al. (2009) and detailed in Antoniewski (2014). Thus, for each sequencing dataset, we collected all of the 23–29 nt RNA reads matching *P(lacW)* or the 42AB locus whose 5' ends overlapped with another 23–29 nt RNA read on the opposite strand. Then, for each possible overlap of 1 to 29 nt, the number of read pairs was counted. To plot the overlap signatures, a z-score was calculated by computing, for each overlap of 1 to *i* nucleotides, the number *O(i)* of read pairs and converting the value using the formula  $z(i) = (O(i) - \text{mean}(O)) / \text{standard deviation}(O)$ . The percentage of reads containing a 'A' at the tenth position was calculated within the paired 23–29 nt RNA matching the reference sequence as described in de Vanssay et al. (2012) (BX2 transgene in Figure 2C and 42AB in Figure 2—figure supplement 1B). GRH49 (BX2\*) was previously analyzed in Hermant et al. (2015). Small RNA sequences and project have been deposited at the GEO under accession number GSE116122.

## ChIP experiments

100 ovaries were dissected in Schneider medium supplemented with insulin at room temperature. Cross-linking was performed for 10 min at room temperature in 1X PBS 1% formaldehyde (Sigma). The cross-linking reaction was stopped by adding glycine to a final concentration of 0.125 mM in PBS 0.1% Triton and incubating 5 min on ice. The cross-linked ovaries were washed with 1 ml of PBS 0.1% Triton and crushed in a dounce A potter 20 times. Then a centrifugation at 400 g for 1 min was performed. The pellet was suspended with 1 ml of cell lysis buffer (KCl 0.085 M, HEPES 5 mM, IGEPAL 0.5%, DTT 0.5 mM, Na butyrate 10 mM, 0.01 M EDTA free protease inhibitor cocktail Roche) and crushed in a dounce B potter 20 times, then 2 mL of cell lysis buffer were added. Centrifugation at 2000 g for 5 min was performed and the pellet was resuspended in 0.5 mL nucleus buffer (HEPES 0.05 M, EDTA 0.01 M, N lauryl Sarcosyl 0.5%, Na butyrate 0.01 M EDTA free protease inhibitor cocktail Roche) and incubated 15 min in a cold room on a rotator. Sonication was performed with Bioruptor (Diagenode) set to high power for 10 cycles (15 s on and 15 s off). A centrifugation was performed 15 min at 16000 g at 4°C. Five µg of chromatin was used for each immunoprecipitation. A pre-clear of 4 hr was performed with 25 µL of dynabead Protein A. The immunoprecipitation reaction was performed with 50 µL of dynabead Protein A coated with 5 µg of antibodies (H3K9me3 polyclonal antibody C1540030 diagenode or Normal Mouse IgG polyclonal antibody 12–371 Millipore), or 20 µL of serum for the Rhino antibody (kindly provided by Dr W. Theurkauf) over night in the cold room on a rotator. Three washes of 10 min in a high salt buffer (Tris HCl pH 7.5 0.05 M, NaCl 0.5 M, Triton 0.25%, IGEPAL 0.5%, BSA 0.5%, EDTA 5 mM) were performed and the elution of chromatin was performed 30 min with 500 µL of elution buffer (Tris pH 7.5 0.05 M, NaCl 0.05 M, EDTA 5 mM, SDS 1%); RNase treatment was omitted; H3K9me3 and Rhino ChIP were respectively done on 5 and 4 independent biological samples followed by qPCR (Roche light Cycler) on each sample. Values were normalized to respective inputs and to two genomic regions known to be enriched in H3K9me3 and Rhino (42AB): region 1 (chr2R: 6449409–6449518) and region 2 (chr2R:

6288809–6288940). An unpaired t-test was used to calculate significance of the differences ( $p < 0.05$ ). Error bars represent the standard deviation.

### RT-qPCR experiments

For each sample, 10  $\mu\text{g}$  of total RNA was treated with DNase (Fermentas). For classical RT-qPCR experiments, 1  $\mu\text{g}$  of DNase-treated RNA was used for reverse transcription using random hexamer primers (Fermentas). Real-time qPCR was performed on triplicates of each sample. *RpL32* was used as reference. The same series of dilution of a mix of different RT preparations was used to normalize the quantity of transcripts in all RT preparations leading to standard quantity (Sq) values. Variations between technical triplicates was very low when compared to variations between biological replicates. The mean of the three technical replicates was then systematically used (meanSq). For each biological sample, we calculated the ratio  $\text{meanSq}(\text{gene})/\text{meanSq}(\text{RpL32})$  to normalize the transcript quantity. Then, the mean of each sample ratio was used to compare the two conditions. For sense-specific RT-qPCR experiments, three reverse transcription were performed using 1  $\mu\text{g}$  of DNase-treated RNA (Fermentas): first without primer (control RT), second with a *lacZ* sense primer (antisense transcript specific RT) and third with a *lacZ* antisense primer (sense transcript specific RT). qPCR was then performed on technical triplicates of each RT using a primer pair specific for *lacZ* sequence. A series of dilutions - ranging from  $50 \times 10^{-15} \text{ g} \cdot \mu\text{l}^{-1}$  to  $0.08 \times 10^{-15} \text{ g} \cdot \mu\text{l}^{-1}$  - of a plasmid containing the *P(lacW)* transgene was used as reference to normalize the quantity of *lacZ* transcripts (Sq values). The number of molecules was estimated by considering that *P(lacW)* is 11191 bp long and that the average weight of a base pair is 650 g/mol. Using Avogadro's number, the number of copies was estimated as equal to the dsDNA amount (in g) times  $6.022 \times 10^{23}$  divided by the dsDNA length times 650. For example,  $50 \times 10^{-15} \text{ g}$  corresponds to approximately 4139 molecules. Variations between technical triplicates were very low when compared to variations between biological replicates. The mean of the three technical replicates was then systematically used (meanSq). The measure of the quantity of transcripts (sense or antisense) for a biological sample was then calculated as the  $(\text{meanSq}(\text{sense or antisense specific}) - \text{meanSq}(\text{control}))$ . This allowed us to eliminate background noise due to unspecific RT amplification for both sense or antisense without specific primer. The mean of each sample ratio was used to compare the two conditions.

### Primer sequences

For classical RT-qPCR experiments, primers used were for *w* (ChIP experiment): 5'-GTCAATG TCCGCTTCAGTT-3' and 5'-GGAGTTTTGGCACAGCACTT-3', these primers are specific of the *P(lacW)* transgene in a *w*<sup>1118</sup> background; for 42AB regions, 5'-TGGAGTTTGGTGCAGAAGC-3' and 5'-AGCCGTGCTTTATGCTTTACT-3' (region 1) and 5'-AAGACCAATTTTTGCGTCGC-3' and 5'-CAAGGATAGGGATTTGGTCC-3' (region 2); for *RpL32*: 5'-CCGCTTCAAGGGACAGTATCTG-3' and 5'-ATCTCGCCGACAGTAAACGC-3'; for *lacZ*: 5'-GAGAATCCGACGGTTGTTA-3' and 5'-AAATTCA-GACGGCAAACGAC-3'; for *eEF5*: 5'-TAACATGGATGTGCCCAATG-3' and 5'-AACGCAATTG TTCACCCAAT-3'; for *AGO1*, primers have been chosen in order to detect spliced forms of transcripts coming upstream of the insertion point of *BX2* and encoding *AGO1*-RA, -RD and -RC isoforms (Figure 4C): 5'-GGATCTCCAGATGACCTCCA-3' and 5'-GGACACTTGTCCGGCTGTAT-3'. For detecting all *AGO1* transcripts isoforms, including the *AGO1*-RB isoform that originates from a promoter located downstream the *BX2* insertion point: 5'-ATGAGCCGGTCATCTTTTTG-3' and 5'-GGCAATCGATGGTTTCTTGT-3'. For sense-specific RT-qPCR experiments, we used specific primers during the reverse transcription step: 5'-AGTACGAAATGCGTCGTTAGAGC-3' for detection of antisense *lacZ* transcripts and 5'-AATGCGCTCAGGTCAAATTC-3' for detection of sense *lacZ* transcripts.

### Acknowledgements

We thank Doug Dorer, Steve Henikoff, Julius Brennecke and the *Bloomington Stock Center* for providing stocks. We thank Bill Theurkauf for providing antibodies. We thank flybase.org for providing databases. We thank Ritha Zamy for technical assistance. We thank Clément Carré, Lori Pile, Ana Maria Vallès and Jean-René Huynh for critical reading of the manuscript. We thank Christophe Antoinewski for helpful advices and development of the ARTbio platform (<http://artbio.fr/>).

## Additional information

### Funding

Funder	Grant reference number	Author
Fondation ARC pour la Recherche sur le Cancer	SFI20131200470	Stéphane Ronsseray
Fondation pour la Recherche Médicale	DEP20131128532	Stéphane Ronsseray
Agence Nationale de la Recherche	plastisipi	Stéphane Ronsseray
Pierre and Marie Curie University	EME1223	Laure Teysset

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Karine Casier, Conceptualization, Data curation, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—review and editing; Valérie Delmarre, Data curation, Validation, Investigation; Nathalie Gueguen, Catherine Hermant, Elise Viodé, Validation, Investigation; Chantal Vaury, Funding acquisition, Writing—review and editing; Stéphane Ronsseray, Conceptualization, Resources, Funding acquisition, Investigation, Writing—review and editing; Emilie Brasset, Validation, Investigation, Writing—review and editing; Laure Teysset, Supervision, Funding acquisition, Investigation, Visualization, Project administration, Writing—review and editing; Antoine Boivin, Conceptualization, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing

### Author ORCIDs

Karine Casier  <http://orcid.org/0000-0001-6825-8057>

Elise Viodé  <https://orcid.org/0000-0002-5058-2710>

Laure Teysset  <https://orcid.org/0000-0002-7413-1850>

Antoine Boivin  <https://orcid.org/0000-0001-8671-1599>

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.39842.032>

Author response <https://doi.org/10.7554/eLife.39842.033>

## Additional files

### Supplementary files

- Supplementary file 1 .Silencing capacities of  $BX2^{ON}$  and  $BX2^{OFF}$  lines throughout generations at 25°C and at 29°C.  $BX2^{OFF}$  and  $BX2^{ON}$  are recombined lines carrying the  $P(TARGET)^{GS}$  and the  $BX2$  locus transgenes on the same chromosome. Numbers show the fraction of females harboring complete germline repression of  $P(TARGET)^{GS}$  at each generation. Complete stability of the initial epigenetic state was observed at 25°C for  $BX2^{OFF}$  and  $BX2^{ON}$  lines, 0% repression (n = 415) and 100% repression (n = 339), respectively. At 29°C, all  $BX2^{OFF}$  lines showed emergence of silencing capacities, 24.7% (n = 3812).  $BX2^{ON}$  lines maintained their silencing capacities over generations at 29°C, 100% (n = 377). nt: not tested.

DOI: <https://doi.org/10.7554/eLife.39842.017>

- Supplementary file 2. Stability of  $BX2^O$  lines. Five  $BX2$ ,  $P(TARGET)^{GS}$  lines showing full repression capacities after 23 generations kept at 29°C were transferred at 25°C and tested for their silencing capacities throughout generations. Numbers show females presenting full  $P(TARGET)^{GS}$  repression and the total number of tested flies. In all cases, the  $BX2^{ON}$  epiallele induced by high temperature remains completely stable during 50 additional generations at 25°C.

DOI: <https://doi.org/10.7554/eLife.39842.018>

- Supplementary file 3. Maternal effect of  $BX2^{ON}$  lines. Reciprocal crosses were performed at 25°C between  $BX2^{ON}$ ,  $P(TARGET)^{GS}$  ( $BX2^O$  or  $BX2^*$ ) individuals and flies carrying a balancer of the second chromosome (Cy). For both the maternal and paternal inheritances (named MI and PI, respectively), lines were established and maintained at 25°C by crossing G1 individuals carrying the  $BX2^O$ ,  $P(TARGET)^{GS}$  chromosome over Cy (**Figure 1—figure supplement 2A**). Silencing capacities of these lines was tested over generations by intra-strain ovarian  $\beta$ -Galactosidase staining. Numbers represent the fraction of females showing complete repression of  $P(TARGET)^{GS}$ . In all cases, maternal transmission of the  $BX2^O$  cluster results in progeny showing complete repression capacities which are stable over generations whereas paternal transmission of the  $BX2^O$  cluster results in the definitive loss of  $BX2$  silencing capacities similarly to the  $BX2^*$  epigenetic state.

DOI: <https://doi.org/10.7554/eLife.39842.019>

- Supplementary file 4. Paramutagenic effect of  $BX2^{ON}$  lines. The capacity of the cytoplasm of  $BX2^{ON}$  females to activate a  $BX2^{OFF}$  cluster was tested as shown in the mating scheme.  $BX2^{ON}$  females (either  $BX2^O$  or  $BX2^*$ ) were crossed with  $BX2^{OFF}$  males, incubated at 25°C. Lines were established with G1 individuals, which have maternally inherited piRNAs and paternally inheritance of the  $BX2$  cluster (**Figure 1—figure supplement 2B**). These lines were maintained at 25°C and their silencing capacities were tested over generations by crossing females with  $P(TARGET)^{GS}$  males. Numbers represent the fraction of females showing complete repression of  $P(TARGET)^{GS}$ . All of the derived lines showed complete silencing capacities over generations, revealing that the cytoplasm of  $BX2^{ON}$  females, either  $BX2^O$  or  $BX2^*$ , can fully activate a  $BX2^{OFF}$  cluster.

DOI: <https://doi.org/10.7554/eLife.39842.020>

- Supplementary file 5. Annotation of small RNA libraries. Small RNAs were prepared from ovaries of females of the indicated genotype. Values for the different categories of sequences are the total number of sequence reads that matched reference libraries. For comparisons, libraries were normalized (normalization factor) to 1 million miRNA (miRNA rpm) or to 1 million Dmel reads (Dmel rpm).

DOI: <https://doi.org/10.7554/eLife.39842.021>

- Supplementary file 6. Silencing capacities of  $BX2^{ON}$  and  $BX2^{OFF}$  lines across generations cultured at 25°C and at 29°C. Same as **Supplementary file 1** except that egg chambers were monitored for  $P(TARGET)^{GS}$  repression instead of whole ovaries. Numbers show the fraction of repressed egg chamber per generation.

DOI: <https://doi.org/10.7554/eLife.39842.022>

- Supplementary file 7. Annotation of small RNA libraries from  $BX2^{OFF}$  raised at 25°C or 29°C. Small RNAs were prepared from ovaries of females of the indicated genotype. Values for the different categories of sequences are the total number of sequence reads that matched reference libraries. For comparisons, libraries were normalized (normalization factor) to 1 million miRNA (miRNA rpm) or to 1 million Dmel reads (Dmel rpm).

DOI: <https://doi.org/10.7554/eLife.39842.023>

- Supplementary file 8.  $P(TARGET)^{GS}$  requirement in the  $BX2$  conversion process. Comparison of the conversion frequency in one generation between  $BX2$ ,  $P(TARGET)^{GS}$  (**Figure 3**) and  $BX2$  (**Figure 3—figure supplement 1**) genotypes. The difference between the presence and absence of the  $P(TARGET)^{GS}$  transgene is highly significant ( $p=8.5 \times 10^{-6}$ , homogeneity  $\chi^2 = 23.35$  with 2 degrees of freedom).

DOI: <https://doi.org/10.7554/eLife.39842.024>

- Supplementary file 9. Silencing capacities of  $BX2^{OFF}$  lines recombined in a  $P(TARGET)^{GS}$  background throughout generations developed at 29°C.  $BX2^{OFF}$  was initially recombined with a line carrying the  $P(TARGET)^{GS}$  transgene to obtain the  $BX2^{OFF}$ ,  $P(TARGET)^{GS}$  lines. From these crosses, eight independent recombinants without the  $P(TARGET)^{GS}$  transgene were recovered and were further cultured at 29°C. To test if some of them acquired silencing capacities, females were crossed with males harboring the  $P(TARGET)^{GS}$  transgene and their progeny was stained for  $\beta$ -Galactosidase expression. Numbers show the fraction of females harboring complete germline repression of  $P(TARGET)^{GS}$  at each generation. A complete stability of the initial epigenetic OFF state was observed for all recombinant lines.

DOI: <https://doi.org/10.7554/eLife.39842.025>



• Supplementary file 10. Silencing capacities of  $BX2^{OFF}$ ;  $P(TARGET)^G$  lines throughout generations at 25°C and at 29°C. Numbers show the fraction of females harboring complete germline repression of  $P(TARGET)^G$  at each generation. Complete stability of the initial epigenetic state was observed at 25°C for  $BX2^{OFF}$ , 0% repression (n = 189). At 29°C,  $BX2^{OFF}$ ;  $P(TARGET)^G$  lines showed emergence of silencing capacities, 19.79% (n = 6766).

DOI: <https://doi.org/10.7554/eLife.39842.026>

• Supplementary file 11. Heat shock and saline stresses do not induce conversion of  $BX2^{OFF}$ .  $BX2^{OFF}$ ,  $P(TARGET)^{GS}$  flies were raised during one generation either on classical cornmeal medium (control) at 25°C or were heat shocked for 1 hr at 37°C at the 0–2 hr embryo stage or were cultured on medium supplemented with 150 mM NaCl. The female progeny were then stained for  $\beta$ -Galactosidase expression and egg chambers were individually monitored in order to detect any possible conversion event. No repressed egg chambers were observed (0/total number of egg chambers). Compared to results obtained at 29°C (from data observed in G1 in **Figure 3**), differences are significant: for heat shock experiment,  $p=7.2 \times 10^{-25}$ , homogeneity  $\chi^2 = 106.03$  with 2 degrees of freedom and for NaCl experiment,  $p=4.9 \times 10^{-27}$ , homogeneity  $\chi^2 = 115.93$  with 2 degrees of freedom.

DOI: <https://doi.org/10.7554/eLife.39842.027>

• Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.39842.028>

### Data availability

Small RNA sequences and project have been deposited at the GEO under accession number GSE116122.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Casier K, Boivin A	2018	Environmentally-induced epigenetic conversion of a piRNA cluster	<a href="https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116122">https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116122</a>	NCBI Gene Expression Omnibus, GSE116122

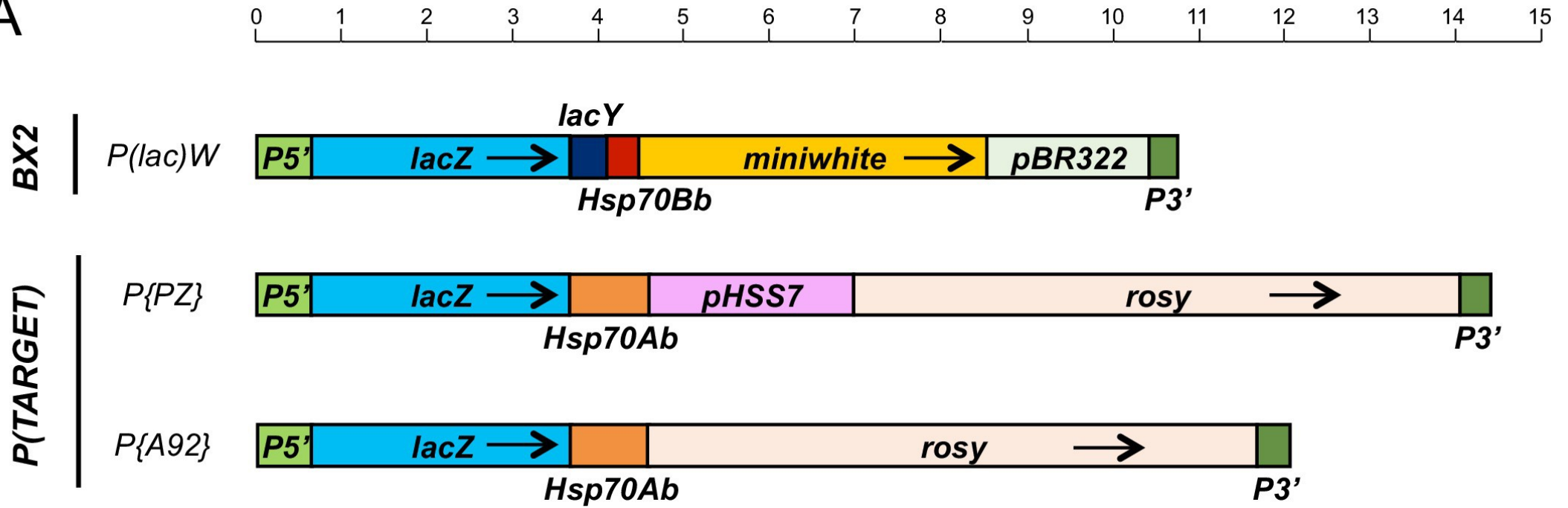
## References

- Akulenko N, Ryazansky S, Morgunova V, Komarov PA, Olovnikov I, Vaury C, Jensen S, Kalmykova A. 2018. Transcriptional and chromatin changes accompanying de novo formation of transgenic piRNA clusters. *Rna* **24**: 574–584. DOI: <https://doi.org/10.1261/rna.062851.117>, PMID: 29358235
- Andersen PR, Tirian L, Vunjak M, Brennecke J. 2017. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549**:54–59. DOI: <https://doi.org/10.1038/nature23482>, PMID: 28847004
- Antoniewski C. 2014. Computing siRNA and piRNA overlap signatures. *Methods in Molecular Biology* **1173**: 135–146. DOI: [https://doi.org/10.1007/978-1-4939-0931-5\\_12](https://doi.org/10.1007/978-1-4939-0931-5_12), PMID: 24920366
- Asif-Laidin A, Delmarre V, Laurentie J, Miller WJ, Ronsseray S, Teyssset L. 2017. Short and long-term evolutionary dynamics of subtelomeric piRNA clusters in *Drosophila*. *DNA Research* **24**:459–472. DOI: <https://doi.org/10.1093/dnares/dsx017>, PMID: 28459978
- 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: <https://doi.org/10.1016/j.cell.2007.01.043>, PMID: 17346786
- Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**:1387–1392. DOI: <https://doi.org/10.1126/science.1165171>, PMID: 19039138
- Brink RA. 1956. A genetic change associated with the R locus in maize which is directed and potentially reversible. *Genetics* **41**:872–889. PMID: 17247669
- Chandler VL. 2007. Paramutation: from maize to mice. *Cell* **128**:641–645. DOI: <https://doi.org/10.1016/j.cell.2007.02.007>, PMID: 17320501
- Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Research* **14**:1188–1190. DOI: <https://doi.org/10.1101/gr.849004>, PMID: 15173120
- de Vanssay A, Bougé AL, Boivin A, Hermant C, Teyssset L, Delmarre V, Antoniewski C, Ronsseray S. 2012. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* **490**:112–115. DOI: <https://doi.org/10.1038/nature11416>, PMID: 22922650
- Dimond JL, Roberts SB. 2016. Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. *Molecular Ecology* **25**:1895–1904. DOI: <https://doi.org/10.1111/mec.13414>, PMID: 26454152

- Dorer DR**, Henikoff S. 1994. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**:993–1002. DOI: [https://doi.org/10.1016/0092-8674\(94\)90439-1](https://doi.org/10.1016/0092-8674(94)90439-1), PMID: 8020105
- Dorer DR**, Henikoff S. 1997. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* **147**:1181–1190. PMID: 9383061
- Fanti L**, Dorer DR, Berloco M, Henikoff S, Pimpinelli S. 1998. Heterochromatin protein 1 binds transgene arrays. *Chromosoma* **107**:286–292. DOI: <https://doi.org/10.1007/s004120050310>, PMID: 9880761
- Fast I**, Hewel C, Wester L, Schumacher J, Gebert D, Zischler H, Berger C, Rosenkranz D. 2017. Temperature-responsive miRNAs in *Drosophila* orchestrate adaptation to different ambient temperatures. *Rna* **23**:1352–1364. DOI: <https://doi.org/10.1261/rna.061119.117>, PMID: 28630141
- Frydman HM**, Spradling AC. 2001. The receptor-like tyrosine phosphatase lar is required for epithelial planar polarity and for axis determination within *Drosophila* ovarian follicles. *Development* **128**:3209–3220. PMID: 11688569
- Gao C**, Xiao M, Ren X, Hayward A, Yin J, Wu L, Fu D, Li J. 2012. Characterization and functional annotation of nested transposable elements in eukaryotic genomes. *Genomics* **100**:222–230. DOI: <https://doi.org/10.1016/j.ygeno.2012.07.004>, PMID: 22800764
- 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: <https://doi.org/10.1126/science.1140494>, PMID: 17322028
- Hermant C**, Boivin A, Teyssset L, Delmarre V, Asif-Laidin A, van den Beek M, Antoniewski C, Ronsseray S. 2015. Paramutation in *Drosophila* requires both nuclear and cytoplasmic actors of the piRNA pathway and induces Cis-spreading of piRNA production. *Genetics* **201**:1381–1396. DOI: <https://doi.org/10.1534/genetics.115.180307>, PMID: 26482790
- Huang X**, Fejes Tóth K, Aravin AA. 2017. piRNA biogenesis in *Drosophila melanogaster*. *Trends in Genetics* **33**:882–894. DOI: <https://doi.org/10.1016/j.tig.2017.09.002>, PMID: 28964526
- Ito H**, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. 2011. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **472**:115–119. DOI: <https://doi.org/10.1038/nature09861>, PMID: 21399627
- Josse T**, Maurel-Zaffran C, de Vanssay A, Teyssset L, Todeschini AL, Delmarre V, Chaminade N, Anxolabéhère D, Ronsseray S. 2008. Telomeric trans-silencing in *Drosophila melanogaster*: tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS ONE* **3**:e3249. DOI: <https://doi.org/10.1371/journal.pone.0003249>, PMID: 18813361
- Klattenhoff C**, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, Weng Z, Theurkauf WE. 2009. The *Drosophila* HP1 homolog rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**:1137–1149. DOI: <https://doi.org/10.1016/j.cell.2009.07.014>, PMID: 19732946
- Kofler R**, Senti K-A, Nolte V, Tobler R, Schlötterer C. 2018. Molecular dissection of a natural transposable element invasion. *Genome Research* **28**:824–835. DOI: <https://doi.org/10.1101/gr.228627.117>
- Langmead B**, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**:R25. DOI: <https://doi.org/10.1186/gb-2009-10-3-r25>, PMID: 19261174
- 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 & Development* **27**:390–399. DOI: <https://doi.org/10.1101/gad.209841.112>, PMID: 23392610
- Le Thomas A**, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen YC, Luo Y, Sachidanandam R, Toth KF, Patel D, Aravin AA. 2014. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes & Development* **28**:1667–1680. DOI: <https://doi.org/10.1101/gad.245514.114>, PMID: 25085419
- Lemaitre B**, Ronsseray S, Coen D. 1993. Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: a model for the P cytotype. *Genetics* **135**:149–160. PMID: 8224815
- Liu R**, Vitte C, Ma J, Mahama AA, Dhliwayo T, Lee M, Bennetzen JL. 2007. A GeneTrek analysis of the maize genome. *PNAS* **104**:11844–11849. DOI: <https://doi.org/10.1073/pnas.0704258104>, PMID: 17615239
- Marie PP**, Ronsseray S, Boivin A. 2017. From embryo to adult: piRNA-mediated silencing throughout germline development in *Drosophila*. *G3 & Genes/Genomes/Genetics* **7**:505–516. DOI: <https://doi.org/10.1534/g3.116.037291>
- Mlodzik M**, Hiromi Y. 1992. Enhancer trap method in *Drosophila*: its application to neurobiology. *Methods in Neurosciences*:397–414. DOI: <https://doi.org/10.1016/b978-0-12-185267-2.50030-1>
- Mohn F**, Sienski G, Handler D, Brennecke J. 2014. The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **157**:1364–1379. DOI: <https://doi.org/10.1016/j.cell.2014.04.031>, PMID: 24906153
- Muerdter F**, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ, Aravin AA. 2012. Production of artificial piRNAs in flies and mice. *Rna* **18**:42–52. DOI: <https://doi.org/10.1261/rna.029769.111>
- Pecinka A**, Dinh HQ, Baubec T, Rosa M, Lettner N, Mittelsten Scheid O. 2010. Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in arabidopsis. *The Plant Cell* **22**:3118–3129. DOI: <https://doi.org/10.1105/tpc.110.078493>, PMID: 20876829
- Pöyhönen M**, de Vanssay A, Delmarre V, Hermant C, Todeschini AL, Teyssset L, Ronsseray S. 2012. Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3 & Genes/Genomes/Genetics* **2**:331–338. DOI: <https://doi.org/10.1534/g3.111.001925>

- Ronsseray S**, Anxolabéhère D, Périquet G. 1984. Hybrid dysgenesis in *Drosophila melanogaster*: influence of temperature on cytotype determination in the P-M system. *MGG Molecular & General Genetics* **196**:17–23. DOI: <https://doi.org/10.1007/BF00334086>, PMID: 6090865
- Ronsseray S**. 1986. P-M system of hybrid dysgenesis in *Drosophila melanogaster*: thermic modifications of the cytotype can be detected for several generations. *MGG Molecular & General Genetics* **205**:23–27. DOI: <https://doi.org/10.1007/BF02428028>
- Ronsseray S**, Lehmann M, Anxolabéhère D. 1991. The maternally inherited regulation of P elements in *Drosophila melanogaster* can be elicited by two P copies at cytological site 1A on the X chromosome. *Genetics* **129**:501–512. PMID: 1660427
- Seong KH**, Li D, Shimizu H, Nakamura R, Ishii S. 2011. Inheritance of stress-induced, ATF-2-dependent epigenetic change. *Cell* **145**:1049–1061. DOI: <https://doi.org/10.1016/j.cell.2011.05.029>, PMID: 21703449
- 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: <https://doi.org/10.1016/j.cell.2012.10.040>, PMID: 23159368
- Tittel-Elmer M**, Bucher E, Broger L, Mathieu O, Paszkowski J, Vaillant I. 2010. Stress-induced activation of heterochromatic transcription. *PLoS Genetics* **6**:e1001175. DOI: <https://doi.org/10.1371/journal.pgen.1001175>, PMID: 21060865
- Weyrich A**, Lenz D, Jeschek M, Chung TH, Rübensam K, Göritz F, Jewgenow K, Fickel J. 2016. Paternal intergenerational epigenetic response to heat exposure in male wild guinea pigs. *Molecular Ecology* **25**:1729–1740. DOI: <https://doi.org/10.1111/mec.13494>, PMID: 26686986
- Wylie A**, Jones AE, Abrams JM. 2016a. p53 in the game of transposons. *BioEssays* **38**:1111–1116. DOI: <https://doi.org/10.1002/bies.201600115>, PMID: 27644006
- Wylie A**, Jones AE, D’Brot A, Lu WJ, Kurtz P, Moran JV, Rakheja D, Chen KS, Hammer RE, Comerford SA, Amatruda JF, Abrams JM. 2016b. p53 genes function to restrain mobile elements. *Genes & Development* **30**:64–77. DOI: <https://doi.org/10.1101/gad.266098.115>, PMID: 26701264
- Yossifoff M**, Kislouk T, Meiri N. 2008. Dynamic changes in DNA methylation during thermal control establishment affect CREB binding to the brain-derived neurotrophic factor promoter. *European Journal of Neuroscience* **28**:2267–2277. DOI: <https://doi.org/10.1111/j.1460-9568.2008.06532.x>, PMID: 19046370
- 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. *PNAS* **110**:19842–19847. DOI: <https://doi.org/10.1073/pnas.1313677110>, PMID: 24248389
- Zhang Z**, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z, Theurkauf WE. 2014. The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**:1353–1363. DOI: <https://doi.org/10.1016/j.cell.2014.04.030>, PMID: 24906152

A

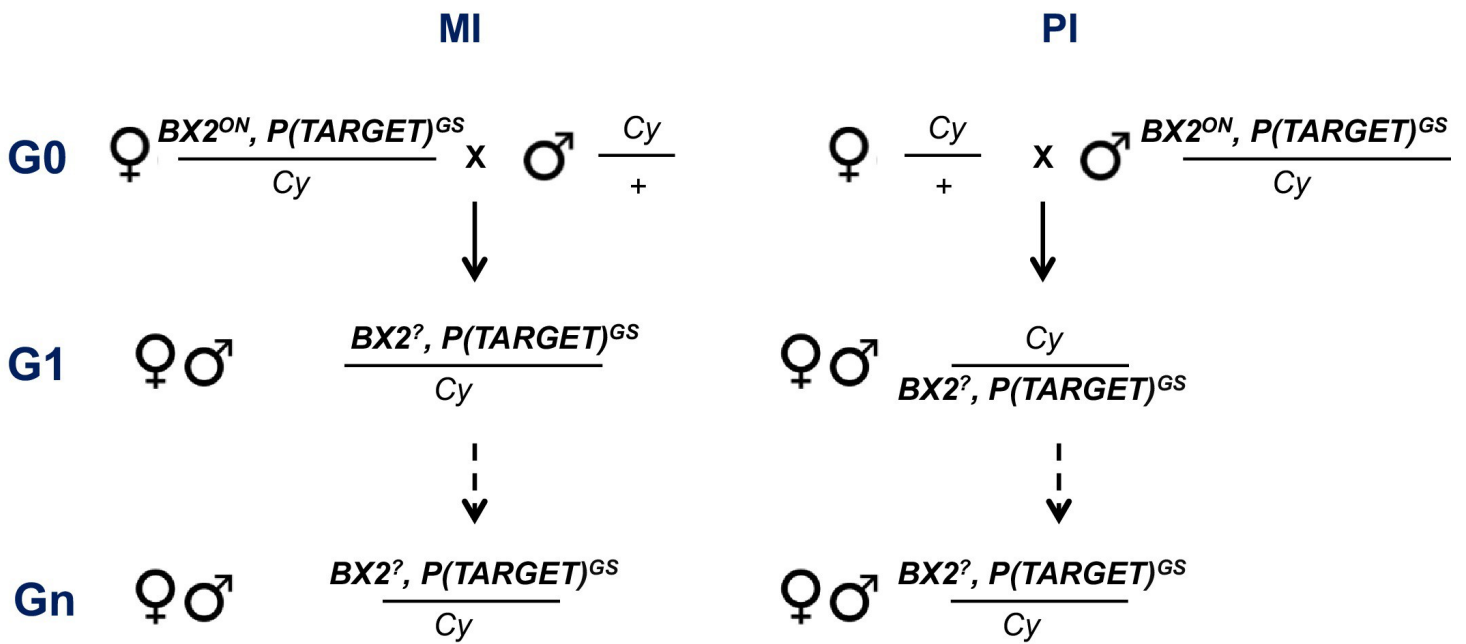


B

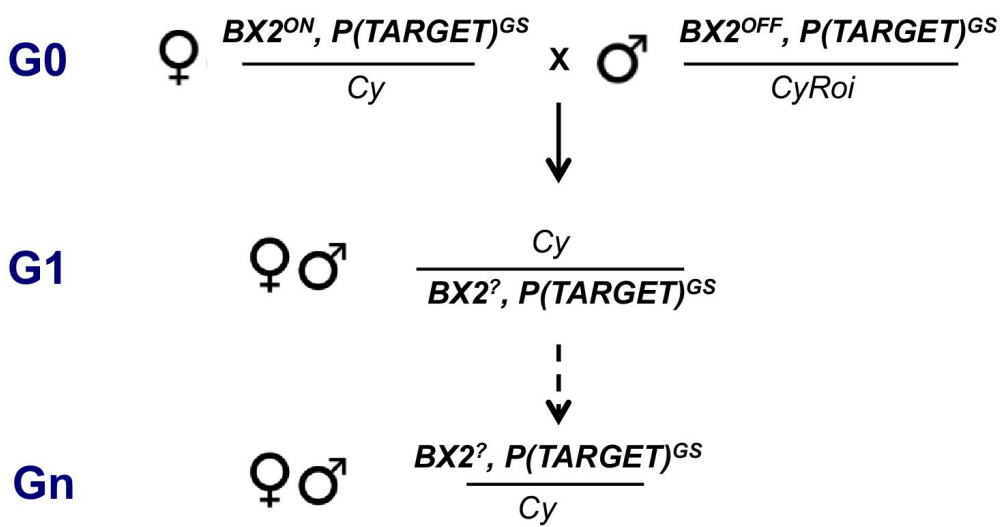
Name used in the article	Transgene structure	Tissue expression	Original line
<i>P(TARGET)<sup>GS</sup></i>	<i>P(PZ)</i>	Germline + Soma	11039
<i>P(TARGET)<sup>G</sup></i>	<i>P(A92)</i>	Germline	BQ16
<i>P(TARGET)<sup>S</sup></i>	<i>P(A92)</i>	Soma	BA37

**Figure 1 – Supplementary figure 1. Schematic maps, drawn at scale, of the different transgenes used in this study.** *BX2* is a cluster of seven *P(lac)W* insertion in tandem. *P(TARGET)* corresponds to euchromatic transgenes used as the reporter. The common transcribed sequence between *BX2* and *P(TARGET)* transgenes is 3547 bp long, from the transcription start at position 87 in *P5'* to the end of *lacZ* sequence. *Hsp70* sequences corresponds to transcription termination signal and are different between *P(lac)W* and *P(TARGET)* transgenes. A very few small RNAs matched these *Hsp70* sequences, and their number is not affected by temperature. Black arrows indicate transcription orientation. Upper

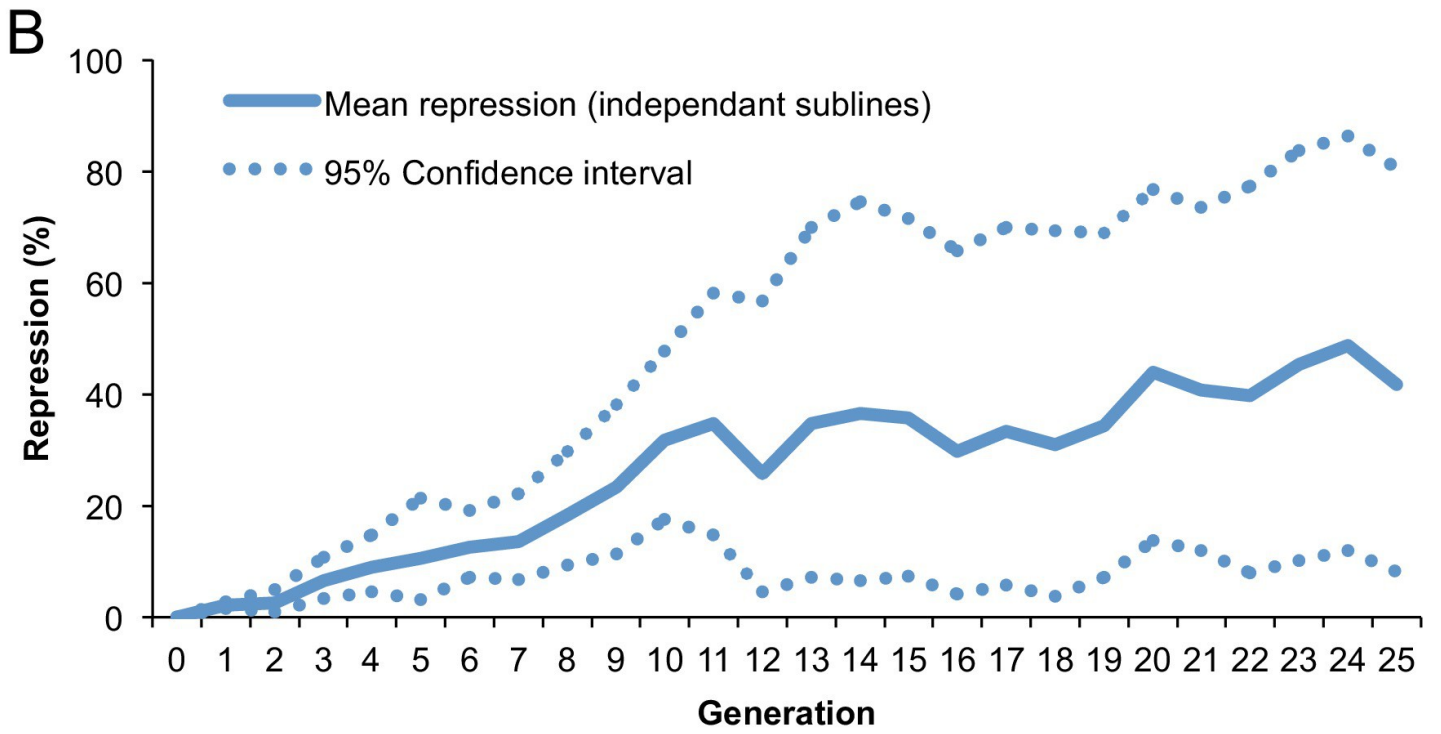
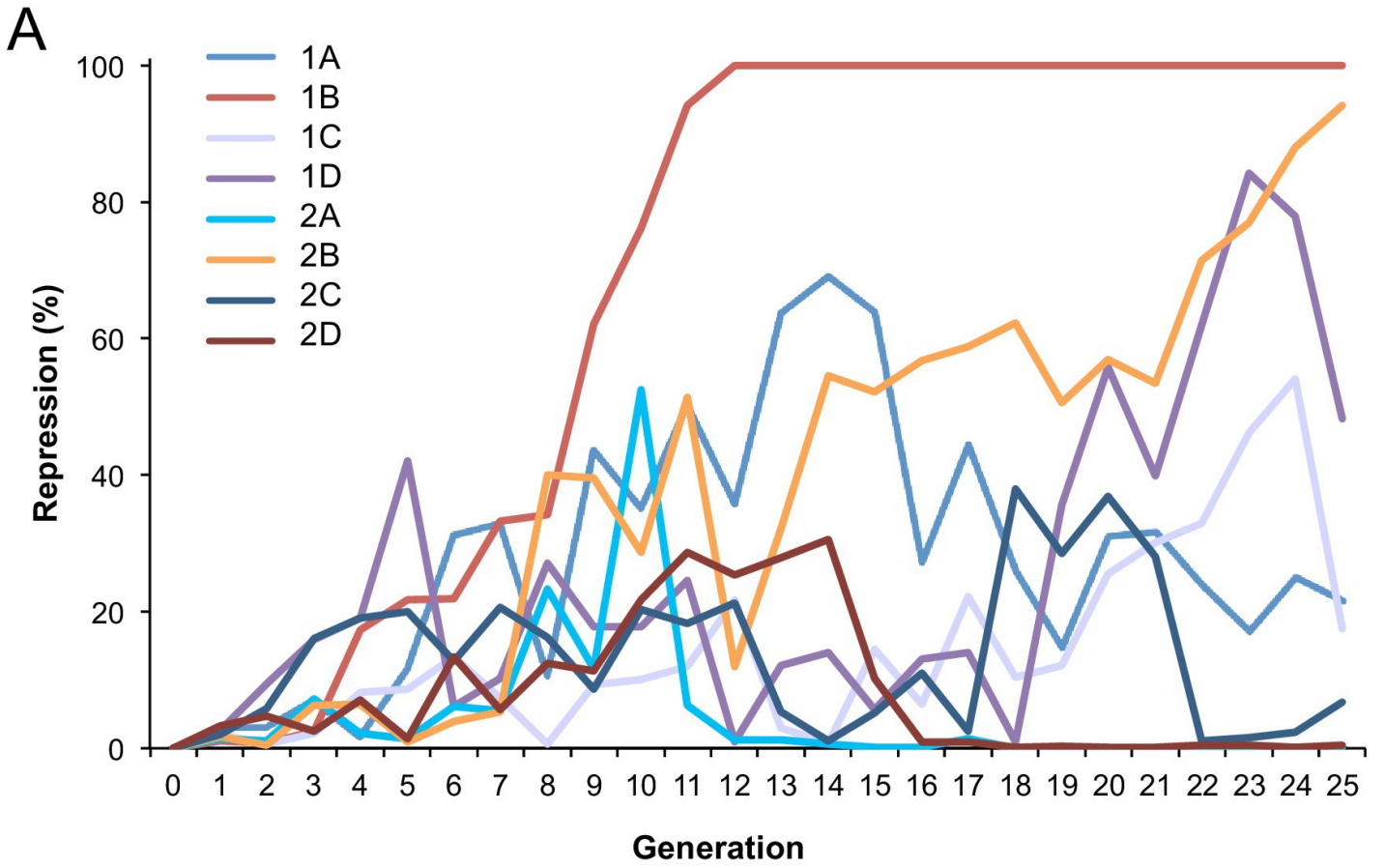
A



B

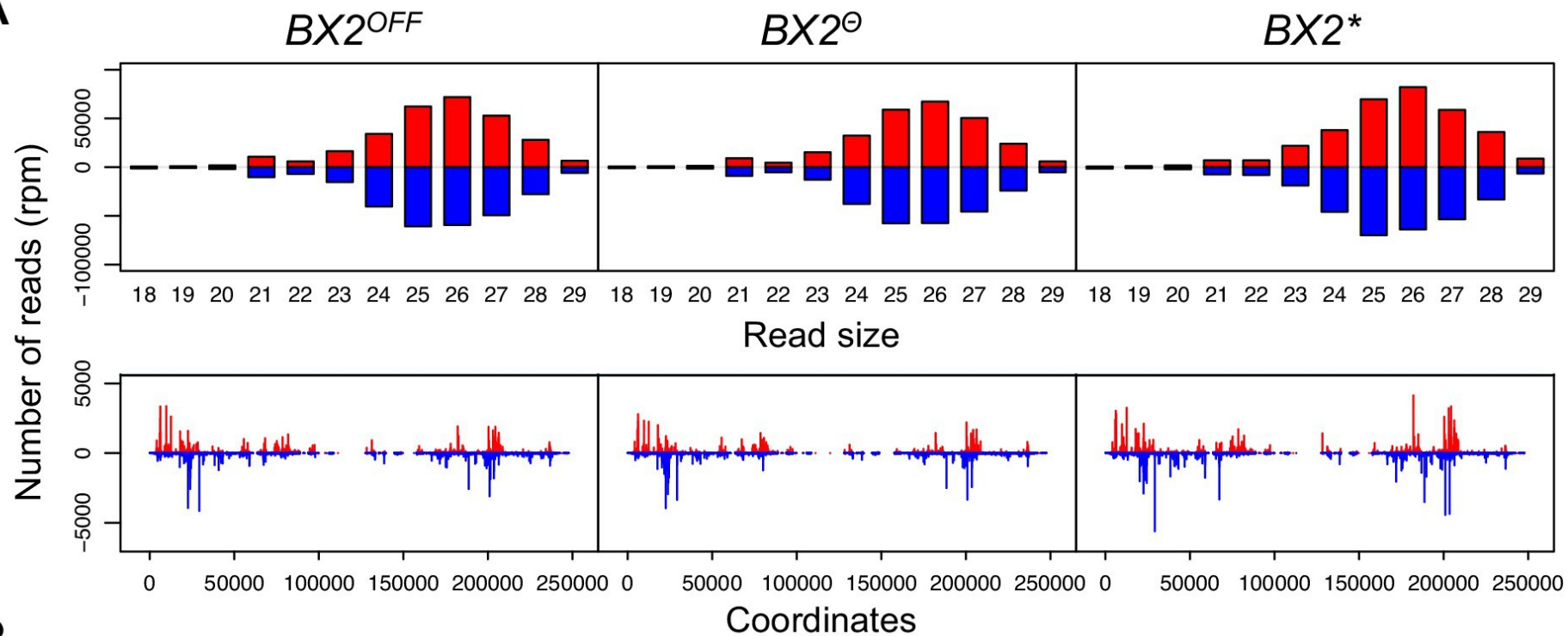


**Figure 1 – Supplementary figure 2. Maternal (A) and paramutagenic (B) effect of  $BX2^{ON}$  lines.** (A) Reciprocal crosses were performed at 25°C between  $BX2^{ON}$ ,  $P(TARGET)^{GS}$  flies (either  $BX2^{\theta}$  or  $BX2^*$ ) and flies carrying a balancer of the second chromosome ( $Cy$ ). Lines were established with G1 individuals of maternal or paternal inheritance (named MI or PI, respectively) that carry the  $BX2$ ,  $P(TARGET)^{GS}$  chromosome, associated with  $white^+$  phenotype conferred by  $BX2$  transgenes, over  $Cy$  and maintained at 25°C. Silencing capacities of these lines were tested over generations by ovarian  $\beta$ -Galactosidase staining ([Supplementary file 3](#)). (B) The capacity of maternally inherited  $BX2^{ON}$  piRNAs to activate a  $BX2^{OFF}$  cluster was tested as shown on the mating scheme.  $BX2^{ON}$ ,  $P(TARGET)^{GS}$  females (either  $BX2^{\theta}$  or  $BX2^*$ ) were crossed at 25°C with  $BX2^{OFF}$ ,  $P(TARGET)^{GS}$  males and lines were established with G1 individuals which maternally inherited piRNAs and paternally inherited the  $BX2$  cluster. Lines were distinguished thanks to two distinct balancers,  $Cy$  and  $CyRoi$  that are devoid of transgenes and carrying different dominant markers. These lines were maintained at 25°C and their silencing capacities were tested over generations by intra-strain ovarian  $\beta$ -Galactosidase staining ([Supplementary file 4](#)).

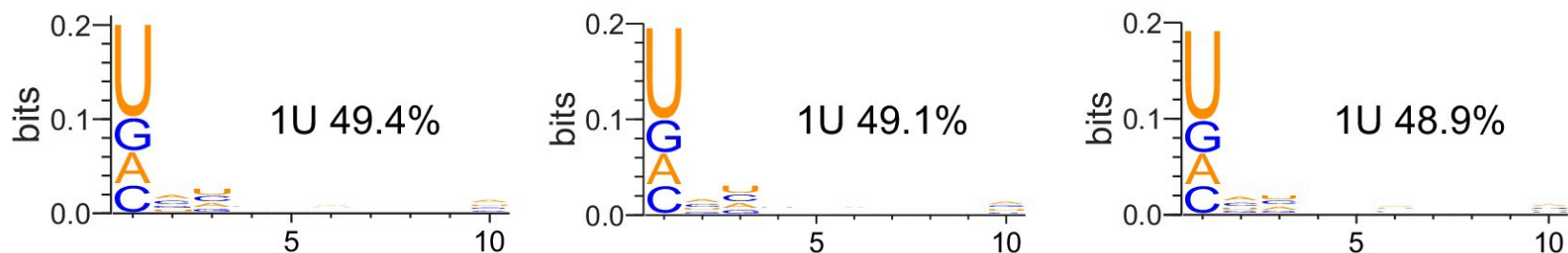




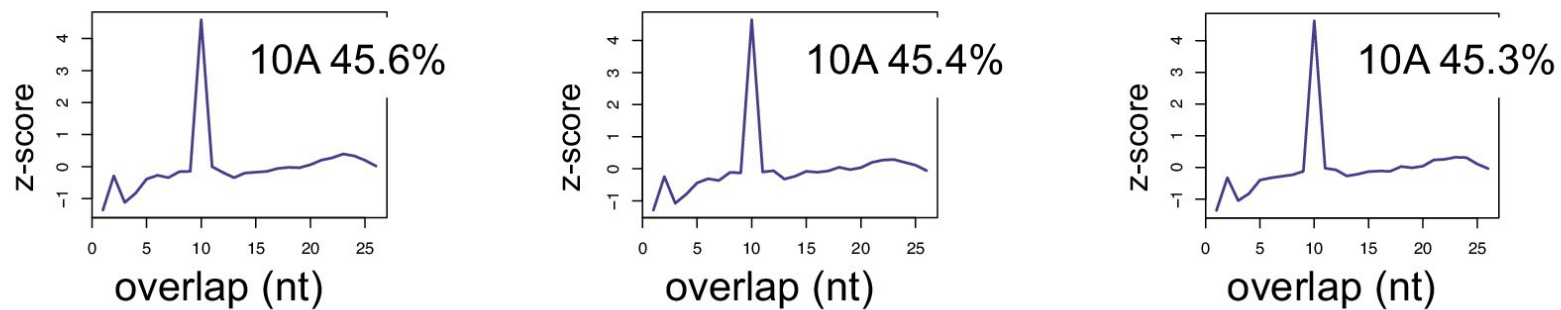
**Figure 1 – Supplementary figure 3. Dynamic of conversion across generations at 29°C.** (A) Fraction of repression across generations for eight independent *BX2*, *P(TARGET)<sup>GS</sup>* lines incubated at 29°C. The fraction of repressed egg chambers was determined at each generation ([Supplementary file 6](#)). *BX2* conversion occurred in each line tested with various dynamics: one of them appears to be homogeneous and stable from the 12<sup>th</sup> generation (1B) while others still fluctuate. 2A subline was lost at the 18th generation. (B) Mean repression and 95% confidence interval of repression frequency for the eight independent lines. Frequencies were modified (arcsine square root) prior to calculate means and confidence intervals. Data were then back transformed (sine-squared) for graphical representation.

**A****B**

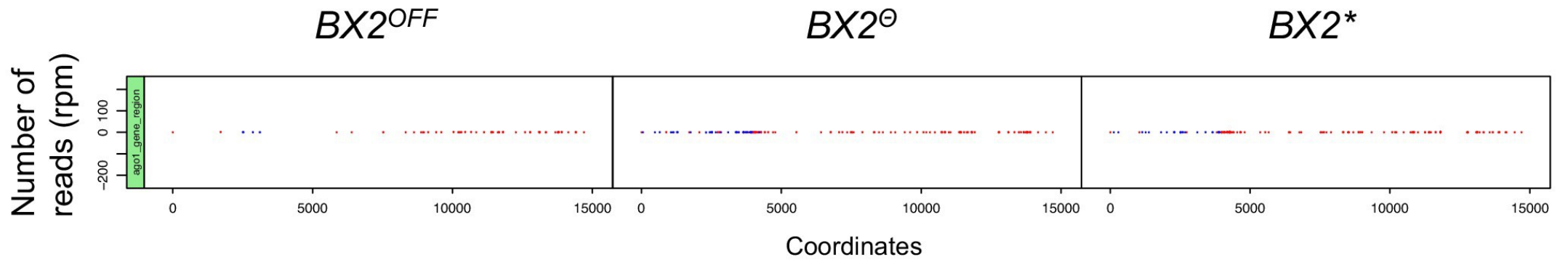
Nucleotide enrichment

**C**

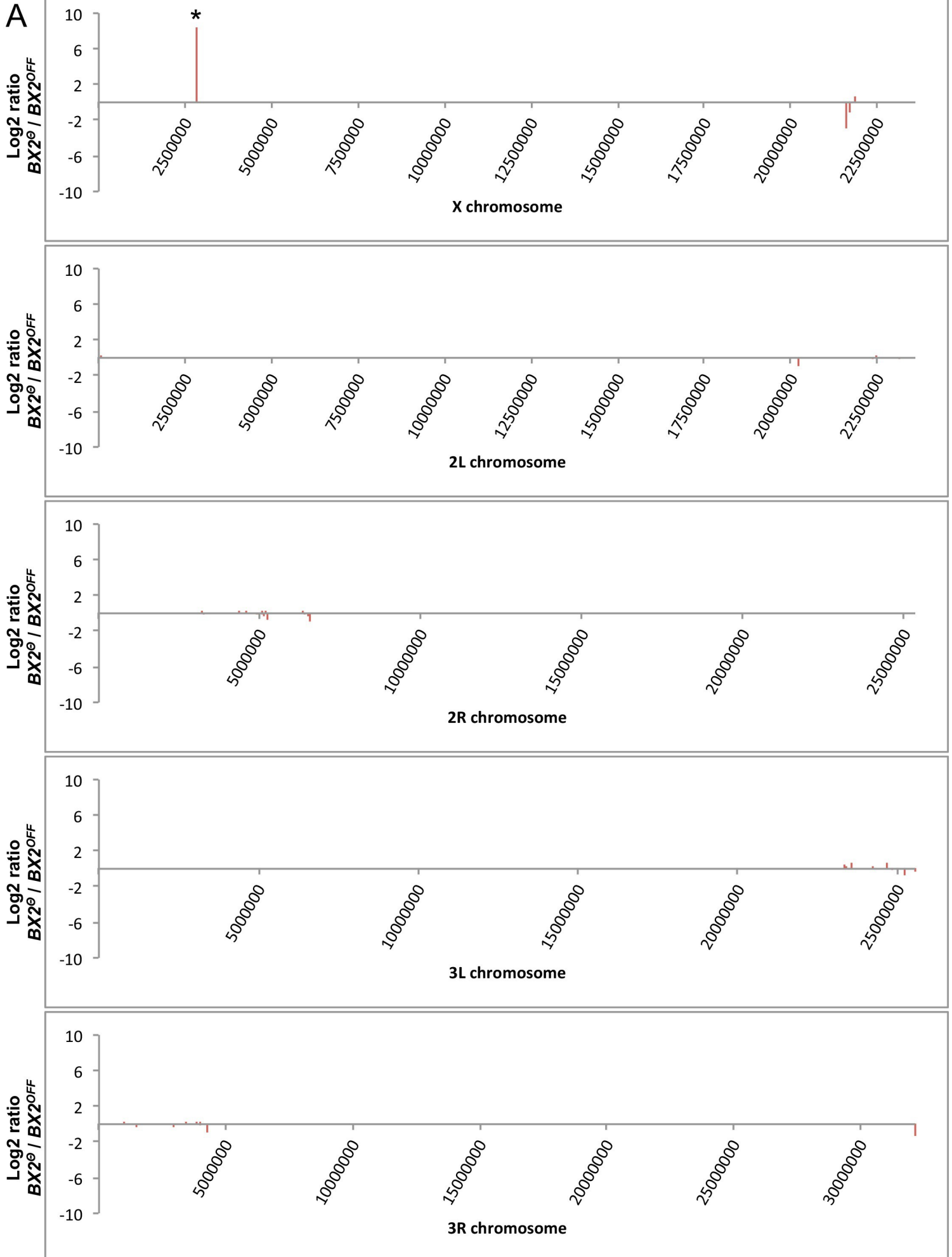
ping-pong signature



**Figure 2 – Supplementary figure 1. *42AB* piRNA production is not altered in different *BX2* epigenetic contexts.** (A) The size distribution and alignments of unique reads that match the *42AB* piRNA cluster show that all genotypes produce similar profile of small RNAs corresponding to *42AB* sequence (upper panels). Unique 23–29 nt mappers are shown on the *42AB* sequence (lower panels). Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. (B) Percentage of 1U-containing small RNAs from each genotype that match *42AB* sequence. (C) Relative frequency (z-score) of overlapping sense-antisense 23–29 nt RNA pairs reveals an enrichment of 10-nucleotides overlapping small RNAs, which corresponds to the ping-pong signature.

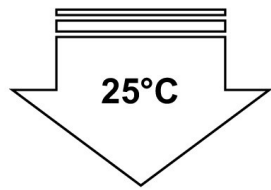


**Figure 2 – Supplementary figure 2. The *AGO1* gene region does not produce piRNA by itself.** Alignments of unique 23–29 nt reads that match the *AGO1* gene region (from nucleotide 13943387 to 13958089 on the 2R chromosome) show that all genotypes produce no significant levels of piRNAs. Compare this result to that of piRNAs matching *BX2* sequences ([Figure 2](#), same y-axis scale). Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. This observation demonstrates that the *AGO1* gene region is not a natural piRNA producing locus.

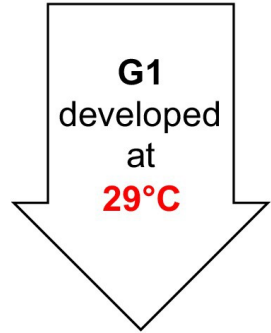
**A**

**Figure 2 – Supplementary figure 3. Mapping the genome for regions stably producing new piRNAs.** (A) Unique piRNAs matching *Drosophila* chromosomes were kept only if they correspond to specific positions in  $BX2^{\theta}$  compared to  $BX2^{OFF}$ . The number of reads per specific positions was then resampled per windows of 50 kilobases. Background noise was reduced by considering only regions that match more than five reads per kilobase on average on both libraries. Log2 ratio of the number of specific reads from  $BX2^{\theta}$  as compared to those from  $BX2^{OFF}$  was plotted along chromosomes (50 kb window). Only one region on the X chromosome presents a high increase (log2 ratio >8.5, marked by an asterisk). (B) These small RNAs match exons of the *white* gene and correspond to piRNAs produced by *P(lacW)* transgenes corresponding to the  $BX2^{\theta}$  locus. No other regions than  $BX2$  stably produce new piRNAs.

A

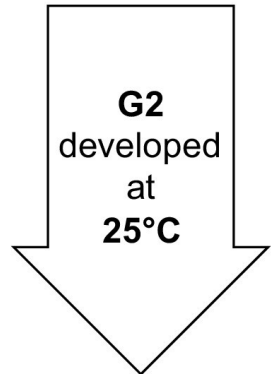


**G0** *BX2<sup>OFF</sup>*



B

157 **G1** females individually mated with *P(TARGET)<sup>GS</sup>* males

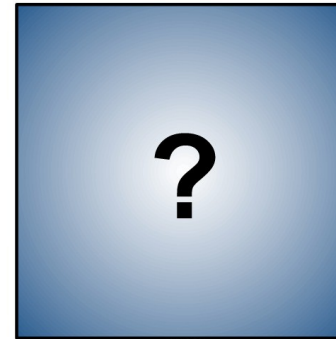


C

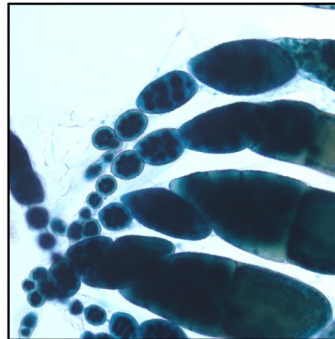
1137 **G2** females analyzed

repression = 0%  
determined by staining of the progeny issued from a cross between female of this strain and *P(TARGET)<sup>GS</sup>* males

unknown state

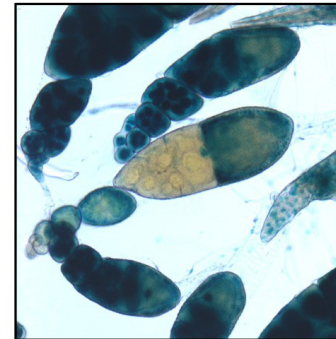


not repressed



n=1136 flies

partially repressed

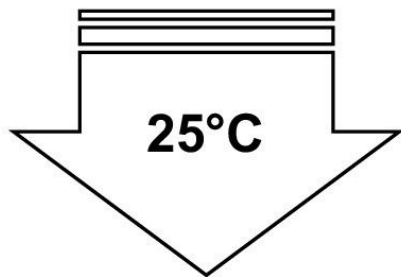
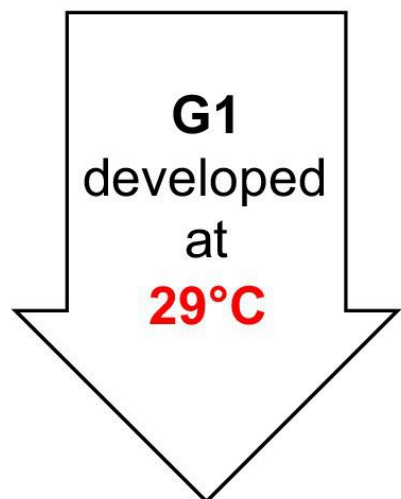
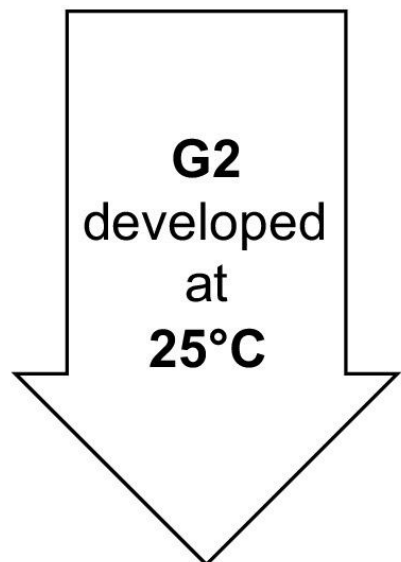


n=1 fly

no fully repressed fly



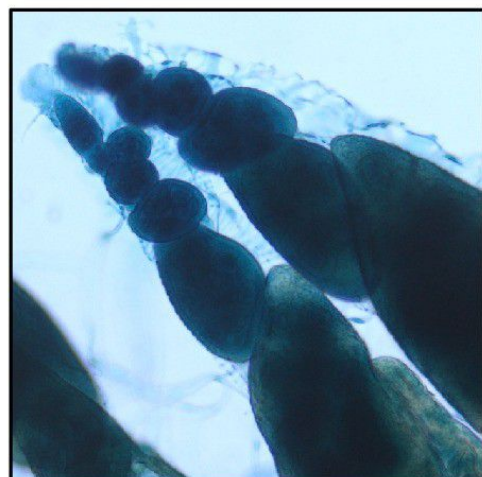
**Figure 3 – Supplementary figure 1. No *BX2* conversion at 29°C without the *P(TARGET)<sup>GS</sup>* reporter transgene.** This experiment is similar to the one in [Figure 3](#), except that *P(TARGET)<sup>GS</sup>* is absent in these flies incubated at 29°C. (A) The *BX2<sup>OFF</sup>* state of G0 females was confirmed by staining the progeny of a cross between females from this strain and *P(TARGET)<sup>GS</sup>* males. G0 females maintained at 25°C lay eggs at 25°C during three days. (B) Eggs were then transferred at 29°C until emergence of the next generation. 157 G1 females were individually mated with two siblings and left to lay for three days at 25°C. It was not possible to know the epigenetic state of *BX2* in these females because of the absence of a reporter transgene. (C) Analysis of the progeny of G1 females reveals that only one female among 1137 presents a partial repression (only one ovariole). Complete repression was never observed.

**A****G0** *BX2*<sup>OFF</sup> ; *P(TARGET)*<sup>S</sup>**B**163 **G1** females mated  
with *P(TARGET)*<sup>GS</sup> males**C**784 **G2** females  
analyzed

repression = 0%  
(n=43 flies)  
determined by  
staining of the  
progeny issued  
from a cross  
between female of  
this strain and  
*P(TARGET)*<sup>GS</sup>  
males

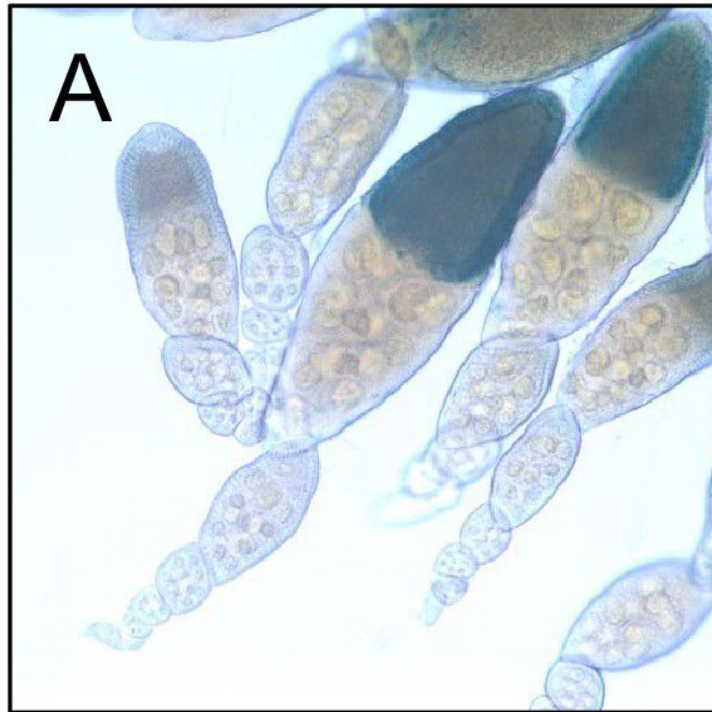
unknown *BX2* state

not repressed

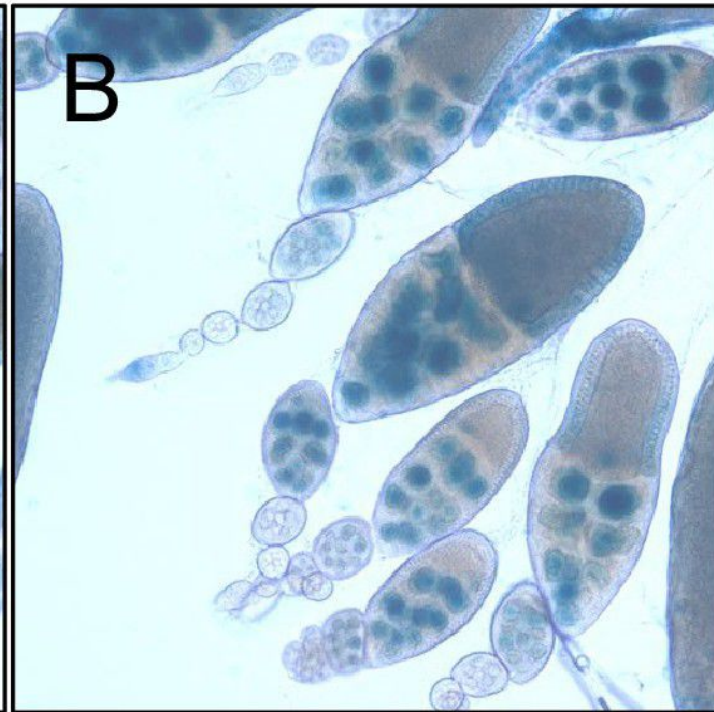


n=784 flies

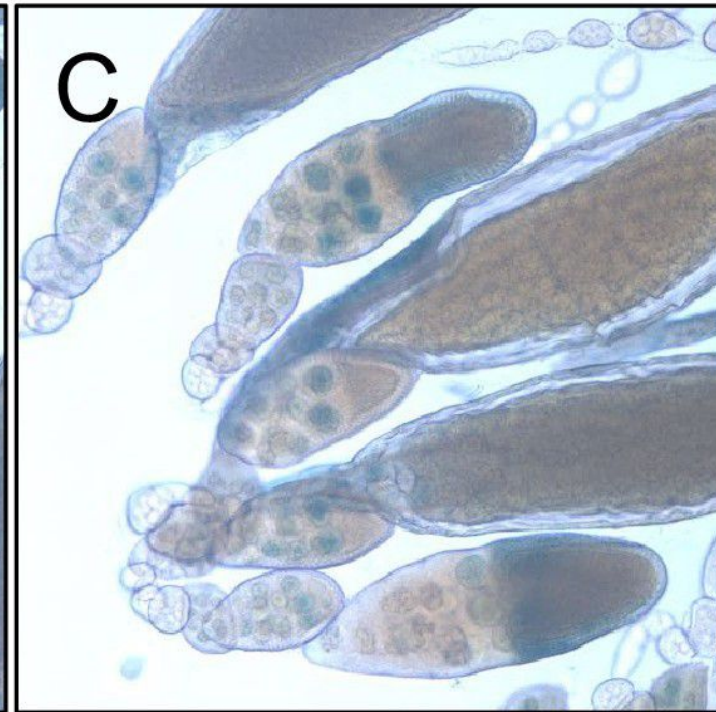
**Figure 3 – Supplementary figure 2. No *BX2* conversion at 29°C with a transgene that is not expressed in the germline.** This experiment is similar to the one in [Figure 3](#), except that  $P(TARGET)^{GS}$  has been replaced by  $P(TARGET)^S$  expressed only in the somatic cells surrounding the germ cells. (A) The  $BX2^{OFF}$  state of G0 females was confirmed by staining the progeny of a cross between females from this strain and  $P(TARGET)^{GS}$  males. G0 females maintained at 25°C lay eggs at 25°C during three days. (B) Eggs were then transferred at 29°C until emergence of the next generation. 163 G1 females were mated *en masse* with males bearing  $P(TARGET)^{GS}$  and left to lay for three days at 25°C. It was not possible to know the epigenetic state of *BX2* in these females because of the absence of a germinally-expressed reporter transgene. (C) Analysis of the progeny of G1 females reveals that no repression was observed among 784 analyzed flies ( $\approx 94080$  estimated egg chambers).



*BX2<sup>ON</sup> nos>w<sup>KD</sup> PTARGET)<sup>G</sup>*

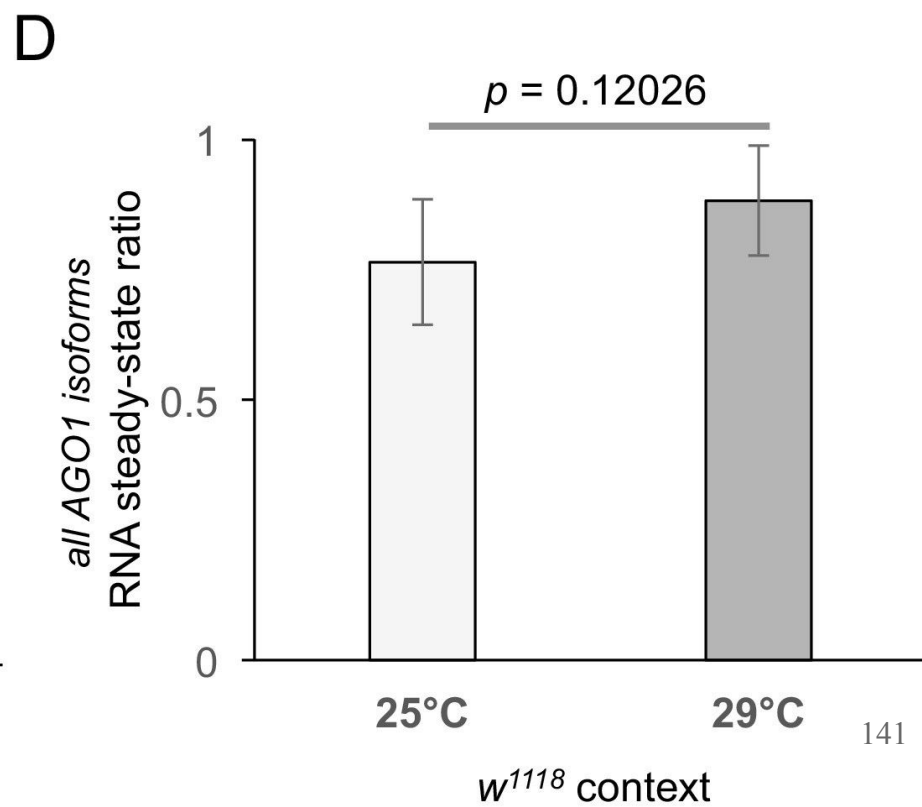
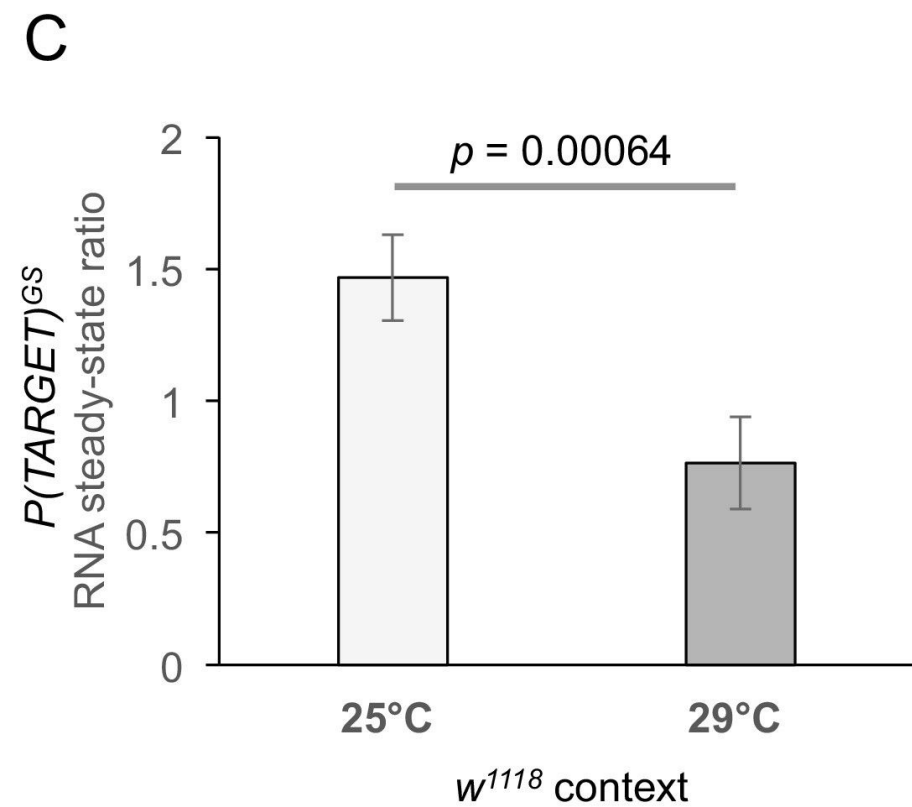
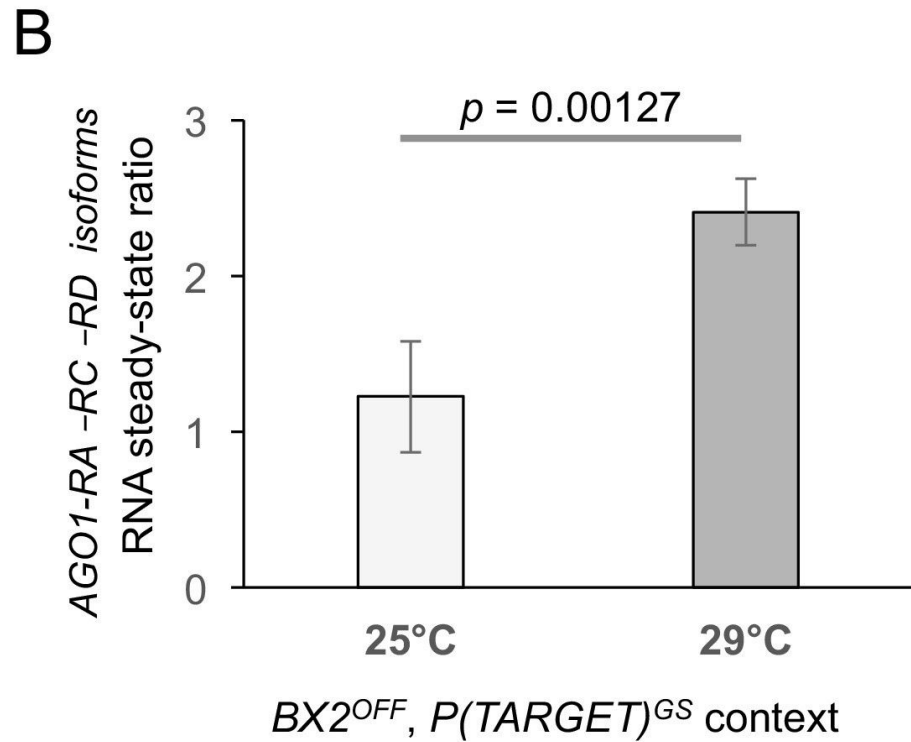
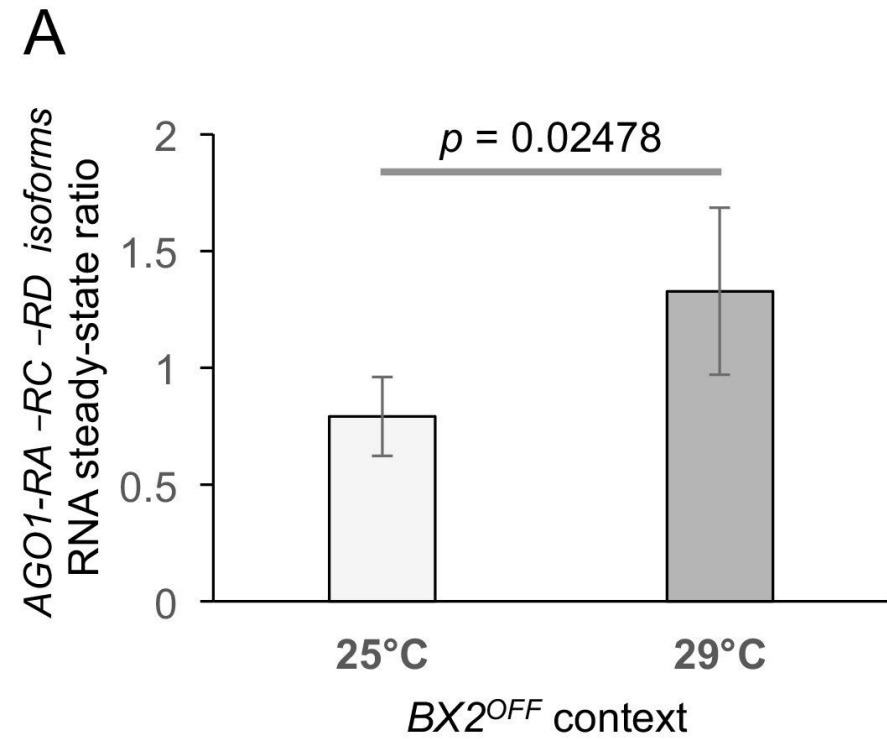


*BX2<sup>ON</sup> nos>moon<sup>PA61</sup> PTARGET)<sup>G</sup>*



*BX2<sup>ON</sup> nos>moon<sup>PA62</sup> PTARGET)<sup>G</sup>*

**Figure 3 – Supplementary figure 3.  $BX2^{ON}$  silencing depends on *moonshiner*.**  $\beta$ -Galactosidase staining of ovaries from females harboring  $BX2^{ON}$ ,  $P(TARGET)^G$ , the *nosGAL4* driver and a RNAi against the *white* gene as a control (**A**) or a RNAi against *moonshiner* (*moon<sup>PA61</sup>* and *moon<sup>PA62</sup>*, **B** and **C**, respectively). Both of the *moon* knockdown lines affect the silencing induced by  $BX2^{ON}$  showing that *moonshiner* is required for the  $BX2^{ON}$  state.



**Figure 4 – Supplementary figure 1. Complementary RNA steady-state measurements by quantitative RT-PCR.** (A, B) Measurement of *AGO1* RNA at 25°C and 29°C in different genetic contexts. (A) *BX2<sup>OFF</sup>* flies at 25°C (n = 6) and 29°C (n = 4). As in the *w<sup>1118</sup>* background experiment (shown in [Figure 4D](#)), the amount of *AGO1* transcripts is increased at 29°C. (B) Same as (A) except that RNA was extracted from flies carrying *BX2<sup>OFF</sup>* and *P(TARGET)<sup>GS</sup>*. Again, the amount of *AGO1* transcripts is increased at 29°C (n = 4) compared to 25°C (n = 4). (C) The amount of *P(TARGET)<sup>GS</sup>* transcripts is decreased at 29°C compared to 25°C. RT-qPCR experiments performed on total RNA extracted from ovaries of flies containing only *P(TARGET)<sup>GS</sup>* transgene in *w<sup>1118</sup>* context revealed that the steady-state level of *lacZ* RNA from *P(TARGET)<sup>GS</sup>* is less abundant at 29°C (n = 6) compared to 25°C (n = 3). (D) Total steady-state *AGO1* RNA was measured in *w<sup>1118</sup>* context at 25°C (n = 6) and 29°C (n = 5) using a set of primers located downstream the *BX2* insertion point (represented as red arrows on [Figure 4C](#)). A slight but non significant difference is observed. *p*-value are given (bilateral Student's *t*-test).

Lines	25°C				29°C									
	<i>BX2<sup>OFF</sup></i>		<i>BX2<sup>ON</sup></i>		<i>BX2<sup>OFF</sup></i>								<i>BX2<sup>ON</sup></i>	
	1	2	1	2	1A	1B	1C	1D	2A	2B	2C	2D	1	2
G1	0/7	0/8	8/8	8/8	0/14	0/14	0/14	0/15	0/15	0/15	0/15	0/15	9/9	10/10
G2	0/11	0/12	12/12	12/12	0/9	0/12	0/10	0/9	0/13	0/7	0/16	1/28	12/12	9/9
G3	0/12	0/3	7/7	8/8	1/17	0/14	0/16	2/14	1/18	1/23	2/23	0/23	7/7	10/10
G4	0/15	0/15	10/10	10/10	0/7	1/7	1/16	2/11	0/17	1/20	4/23	1/23	10/10	8/8
G5	0/12	0/12	12/12	12/12	2/20	6/31	3/39	11/27	0/38	0/33	0/3	0/24	8/8	6/6
G6	0/6	0/8	8/8	8/8	0/2	4/20	1/8	1/18	0/9	0/15	1/14	0/4	1/1	1/1
G7	0/8	0/8	8/8	8/8	6/19	7/21	2/36	2/21	2/38	1/20	3/17	2/37	7/7	8/8
G8	0/10	0/8	6/6	8/8	3/33	11/33	0/49	8/31	5/22	14/35	5/33	6/50	8/8	8/8
G9	nt	nt	nt	nt	10/23	13/21	1/22	5/28	3/26	17/44	3/36	4/38	8/8	8/8
G10	0/8	0/8	8/8	8/8	8/23	16/21	1/22	4/23	11/21	7/25	7/35	5/24	8/8	8/8
G11	0/8	0/8	8/8	8/8	7/14	16/18	1/15	3/15	1/16	8/16	2/16	4/16	8/8	8/8
G12	0/8	0/8	8/8	8/8	5/14	7/8	1/5	0/8	0/6	2/17	3/18	5/21	8/8	5/5
G13	0/10	0/2	6/6	5/5	7/11	11/11	0/37	1/9	nt	6/19	1/25	5/18	9/9	8/8
G14	0/6	0/10	5/5	5/5	28/41	25/25	0/24	5/41	nt	6/11	0/24	7/24	6/6	5/5
G15	0/12	0/4	8/8	8/8	14/22	16/16	3/21	1/23	nt	12/23	1/24	2/20	8/8	8/8
G16	0/12	0/11	8/8	5/5	3/11	17/17	1/17	2/17	nt	11/20	2/20	0/19	8/8	8/8
G17	0/8	0/12	5/5	7/7	5/9	22/22	4/18	0/6	nt	9/16	0/13	0/19	8/8	3/3
G18	0/11	0/12	8/8	8/8	2/8	6/6	1/10	0/5	nt	13/21	3/8	0/14	8/8	8/8
G19	0/11	0/12	6/6	8/8	6/42	21/21	4/38	6/17	nt	16/32	5/18	0/44	8/8	8/8
G20	0/12	0/3	6/6	6/6	4/13	21/21	4/17	10/18	nt	10/20	12/33	0/36	8/8	8/8
G21	0/12	0/12	8/8	8/8	5/16	12/12	8/28	9/23	nt	16/30	7/27	0/31	8/8	8/8
G22	0/12	0/12	8/8	8/8	2/9	14/14	5/13	12/21	nt	10/14	0/13	0/23	8/8	7/7
G23	0/4	0/12	5/5	5/5	3/18	25/25	9/20	16/19	nt	13/17	0/20	0/25	8/8	8/8
G24	nt	nt	nt	nt	3/12	17/17	7/13	7/9	nt	7/8	0/16	0/30	8/8	8/8
G25	nt	nt	nt	nt	5/24	25/25	4/25	12/25	nt	16/17	1/20	0/18	6/6	6/6
Total	0/215	0/200	168/168	171/171	129/431	313/452	61/533	119/453	23/297	196/518	62/510	42/624	195/195	182/182
%R	0% n=415		100% n=339		24.7% n=3818								100% n=377	

**Supplementary file 1. Silencing capacities of *BX2<sup>ON</sup>* and *BX2<sup>OFF</sup>* lines throughout generations at 25°C and at 29°C. *BX2<sup>OFF</sup>* and *BX2<sup>ON</sup>* are recombined lines carrying the *P(TARGET)<sup>GS</sup>* and the *BX2* locus transgenes on the same chromosome. Numbers show the**



fraction of females harboring complete germline repression of  $P(TARGET)^{GS}$  at each generation. Complete stability of the initial epigenetic state was observed at 25°C for  $BX2^{OFF}$  and  $BX2^{ON}$  lines, 0% repression (n = 415) and 100% repression (n = 339), respectively. At 29°C, all  $BX2^{OFF}$  lines showed emergence of silencing capacities, 24.7% (n = 3812).  $BX2^{ON}$  lines maintained their silencing capacities over generations at 29°C, 100% (n = 377). nt: not tested.

<b>Lines</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
G1	4/4	3/3	4/4	3/3	4/4
G6	2/2	4/4	4/4	4/4	5/5
G7	7/7	8/8	8/8	6/6	8/8
G10	nt	3/3	nt	5/5	5/5
G11	6/6	6/6	6/6	6/6	6/6
G14	2/2	8/8	8/8	1/1	3/3
G16	5/5	5/5	6/6	6/6	6/6
G25	3/3	6/6	6/6	8/8	nt
G26	10/10	5/5	10/10	6/6	5/5
G34	8/8	8/8	8/8	8/8	8/8
G35	12/12	12/12	12/12	12/12	12/12
G43	8/8	7/7	8/8	8/8	8/8
G50	8/8	8/8	8/8	8/8	8/8
<b>Total</b>	<b>75/75</b>	<b>83/83</b>	<b>88/88</b>	<b>81/81</b>	<b>78/78</b>

**Supplementary file 2. Stability of  $BX2^O$  lines.** Five  $BX2$ ,  $P(TARGET)^{GS}$  lines showing full repression capacities after 23 generations kept at 29°C were transferred at 25°C and tested for their silencing capacities throughout generations. Numbers show females presenting full  $P(TARGET)^{GS}$  repression and the total number of tested flies. In all cases, the  $BX2^{ON}$  epiallele induced by high temperature remains completely stable during 50 additional generations at 25°C.

Lines	<i>BX2<sup>θ</sup></i>						<i>BX2*</i>					
	MI-1	MI-2	MI-3	PI-1	PI-2	PI-3	MI-1	MI-2	MI-3	PI-1	PI-2	PI-3
G1	7/7	8/8	8/8	0/8	0/8	0/8	3/3	5/5	8/8	0/5	0/5	0/6
G2	8/8	8/8	8/8	0/8	0/8	0/8	8/8	8/8	8/8	0/8	0/8	0/12
G3	8/8	7/7	8/8	0/8	0/8	0/8	8/8	8/8	8/8	0/8	nt	0/8
G5	8/8	8/8	8/8	0/7	0/8	0/8	8/8	8/8	8/8	0/8	0/8	0/8
G10	8/8	8/8	8/8	0/7	0/8	0/8	8/8	8/8	8/8	0/7	0/8	0/8
G15	6/6	8/8	8/8	0/8	0/7	0/6	8/8	8/8	8/8	0/8	0/8	0/8
G20	5/5	4/4	2/2	0/6	0/6	0/6	7/7	8/8	8/8	0/8	0/8	0/8
Subtotal	50/50	49/49	53/53	0/50	0/54	0/52	50/50	53/53	56/56	0/52	0/45	0/58
Total	152/152			0/156			159/159			0/155		

**Supplementary file 3. Maternal effect of *BX2<sup>ON</sup>* lines.** Reciprocal crosses were performed at 25°C between *BX2<sup>ON</sup>*, *P(TARGET)<sup>GS</sup>* (*BX2<sup>θ</sup>* or *BX2\**) individuals and flies carrying a balancer of the second chromosome (*Cy*). For both the maternal and paternal inheritances (named MI and PI, respectively), lines were established and maintained at 25°C by crossing G1 individuals carrying the *BX2<sup>θ</sup>*, *P(TARGET)<sup>GS</sup>* chromosome over *Cy* ([Figure 1—figure supplement 2A](#)). Silencing capacities of these lines was tested over generations by intra-strain ovarian β-Galactosidase staining. Numbers represent the fraction of females showing complete repression of *P(TARGET)<sup>GS</sup>*. In all cases, maternal transmission of the *BX2<sup>θ</sup>* cluster results in progeny showing complete repression capacities which are stable over generations whereas paternal transmission of the *BX2<sup>θ</sup>* cluster results in the definitive loss of *BX2* silencing capacities similarly to the *BX2\** epigenetic state.

Lines	<i>BX2*</i>			<i>BX2<sup>θ</sup></i>		
	1	2	3	1	2	3
G1	8/8	8/8	10/10	10/10	9/9	9/9
G2	10/10	10/10	10/10	10/10	10/10	10/10
G3	8/8	8/8	8/8	8/8	8/8	7/7
G5	6/6	7/7	7/7	8/8	5/5	7/7
G10	6/6	3/3	6/6	6/6	6/6	6/6
G15	8/8	8/8	7/7	7/7	8/8	6/6
G20	4/4	5/5	5/5	7/7	8/8	4/4
Subtotal	50/50	49/49	53/53	56/56	54/54	49/49
Total	152/152			159/159		

**Supplementary file 4. Paramutagenic effect of *BX2<sup>ON</sup>* lines.** The capacity of the cytoplasm of *BX2<sup>ON</sup>* females to activate a *BX2<sup>OFF</sup>* cluster was tested as shown in the mating scheme. *BX2<sup>ON</sup>* females (either *BX2<sup>θ</sup>* or *BX2\**) were crossed with *BX2<sup>OFF</sup>* males, incubated at 25°C. Lines were established with G1 individuals, which have maternally inherited piRNAs and paternally inheritance of the *BX2* cluster ([Figure 1—figure supplement 2B](#)). These lines were maintained at 25°C and their silencing capacities were tested over generations by crossing females with *P(TARGET)<sup>GS</sup>* males. Numbers represent the fraction of females showing complete repression of *P(TARGET)<sup>GS</sup>*. All of the derived lines showed complete silencing capacities over generations, revealing that the cytoplasm of *BX2<sup>ON</sup>* females, either *BX2<sup>θ</sup>* or *BX2\**, can fully activate a *BX2<sup>OFF</sup>* cluster.

<b>Library ID</b>	<b>GRH77</b>	<b>GRH78</b>	<b>GRH49</b>
Genotype	<i>BX2<sup>OFF</sup>, P(TARGET)<sup>GS</sup></i>	<i>BX2<sup>θ</sup>, P(TARGET)<sup>GS</sup></i>	<i>BX2*</i>
D. melanogaster Jan. 2013 (Flybase r5.49)	1,692,3421	18,545,573	16,242,473
dme_miR_r20	3,621,703	3,951,462	3,871,488
Dmel_all-tRNA	450,601	370,751	411,604
Dmel_all-miscRNA	2,004,915	2,823,568	992,800
Dmel_transposon_set_BDGP_v941	6,161,407	6,580,495	6,396,904
D. melanogaster transcripts Jan. 2013 (Flybase r5.49)	1,747,937	1,734,672	1,521,258
<b>Normalization factor miRNA rpm</b>	<b>0.276</b>	<b>0.253</b>	<b>0.258</b>
Normalization factor Dmel rpm	0.059	0.054	0.062

**Supplementary file 5. Annotation of small RNA libraries.** Small RNAs were prepared from ovaries of females of the indicated genotype. Values for the different categories of sequences are the total number of sequence reads that matched reference libraries. For comparisons, libraries were normalized (normalization factor) to 1 million miRNA (miRNA rpm) or to 1 million Dmel reads (Dmel rpm).

Lines	25°C				29°C									
	<i>OFF</i> BX2		<i>ON</i> BX2		<i>OFF</i> BX2								<i>ON</i> BX2	
	1	2	1	2	1A	1B	1C	1D	2A	2B	2C	2D	1	2
G1	0/840	0/980	960/960	960/960	52/1680	18/1680	33/1680	47/1800	26/1800	31/1800	36/1800	60/1800	1080/1080	1200/1200
G2	0/1320	0/1440	1440/1440	1440/1440	33/1080	11/1440	8/1200	102/1080	17/1560	4/840	112/1920	158/3360	1440/1440	1080/1080
G3	0/1440	0/360	840/840	960/960	145/2040	41/1680	42/1920	268/1680	155/2160	175/2760	444/2760	67/2760	840/840	1200/1200
G4	0/1800	0/1800	1200/1200	1200/1200	15/840	145/840	157/1920	251/1320	43/2040	156/2400	527/2760	195/2760	1200/1200	960/960
G5	0/1440	0/1440	1440/1440	1440/1440	278/2400	808/3720	405/4680	1362/240	67/4560	36/3960	72/360	39/2880	960/960	720/720
G6	0/720	0/960	960/960	960/960	75/240	524/2400	128/960	134/2160	67/1080	70/1800	214/1680	64/480	120/120	120/120
G7	0/960	0/960	960/960	960/960	750/2280	840/2520	318/4320	256/2520	249/4560	130/2400	422/2040	250/4440	840/840	960/960
G8	0/1200	0/960	720/720	960/960	425/3960	1352/3960	37/5880	1008/3720	614/2640	1684/4200	643/3960	743/6000	960/960	960/960
G9	nt	nt	nt	nt	1204/2760	1565/2520	243/2640	596/3360	361/3120	2093/5280	375/4320	514/4560	960/960	960/960
G10	0/960	0/960	960/960	960/960	970/2760	1920/2520	264/2640	492/2760	1324/2520	861/3000	854/4200	626/2880	960/960	960/960
G11	0/960	0/960	960/960	960/960	846/1680	1920/2040	214/1800	443/1800	120/1920	987/1920	352/1920	550/1920	960/960	960/960
G12	0/960	0/960	960/960	960/960	600/1680	840/840	130/600	9/960	9/720	244/2040	460/2160	641/2520	960/960	600/600
G13	0/1200	0/240	720/720	600/600	840/1320	1320/1320	132/4440	131/1080	nt	732/2280	161/3000	603/2160	1080/1080	960/960
G14	0/720	0/1200	600/600	600/600	3395/4920	3000/3000	20/2880	690/4920	nt	720/1320	33/2880	882/2880	720/720	600/600
G15	0/1440	0/480	960/960	960/960	1687/2640	1920/1920	367/2520	156/2760	nt	1440/2760	148/2880	244/2400	960/960	960/960
G16	0/1440	0/1320	960/960	600/600	360/1320	2040/2040	130/2040	265/2040	nt	1362/2400	263/2400	21/2280	960/960	960/960
G17	0/960	0/1440	600/600	840/840	480/1080	2640/2640	480/2160	101/720	nt	1128/1920	40/1560	21/2280	960/960	360/360
G18	0/1320	0/1440	960/960	960/960	249/960	720/720	124/1200	6/600	nt	1569/2520	365/960	1/1680	960/960	960/960
G19	0/1320	0/1440	720/720	960/960	750/5040	2520/2520	556/4560	727/2040	nt	1942/3840	616/2160	18/5280	960/960	960/960
G20	0/1440	0/360	720/720	360/360	484/1560	2520/2520	520/2040	1207/2160	nt	1366/2400	1460/3960	4/4320	960/960	960/960
G21	0/1440	0/1440	960/960	960/960	608/1920	1440/1440	1011/3360	1102/2760	nt	1922/3600	907/3240	6/3720	960/960	960/960
G22	0/1440	0/1440	960/960	960/960	259/1080	1680/1680	513/1560	1560/2520	nt	1200/1680	16/1560	11/2760	960/960	840/840
G23	0/480	0/1440	600/600	480/480	372/3160	3000/3000	1110/2400	1920/2280	nt	1571/2040	38/2400	15/3000	960/960	960/960
G24	nt	nt	nt	nt	361/1440	2040/2040	844/1560	842/1080	nt	844/960	44/1920	4/3600	960/960	960/960
G25	nt	nt	nt	nt	622/2880	3000/3000	525/3000	1448/3000	nt	1920/2040	163/2400	11/2160	720/720	720/720
Total	0/25800	0/24000	20160/20160	20040/20040	15860/51720	37824/54000	8311/63960	15123/54360	3052/28680	24187/62160	8765/61200	5748/74880	24600/24600	23040/23040

**Supplementary file 6. Silencing capacities of  $BX2^{ON}$  and  $BX2^{OFF}$  lines across generations cultured at 25°C and at 29°C.** Same as [Supplementary file 1](#) except that egg chambers were monitored for  $P(TARGET)^{GS}$  repression instead of whole ovaries. Numbers show the fraction of repressed egg chamber per generation.

Library ID	GRH160	GRH161	GRH162	GRH163	GRH164	GRH165
Genotype	<i>BX2<sup>OFF</sup></i>					
Temperature	25°C			29°C		
D. melanogaster Jan. 2013 (Flybase)	3,288,5973	3,198,7242	3,135,9878	3,147,6903	364,50841	269,32299
dme_miR_r20	8,098,219	7,458,369	7,225,648	6,804,381	8,472,003	6,374,911
Dmel_all-tRNA	549,423	676,558	587,138	475,249	793,839	410,800
Dmel_all-miscRNA	4,328,239	6,119,121	4,001,654	6,605,018	7,361,768	5,374,259
Dmel_transposon_set_BDGP_v941	11,688,163	10,268,266	11,511,292	9,968,778	11,046,992	8,462,052
D. melanogaster transcripts Jan. 2013	4,656,410	4,732,913	4,796,480	4,427,408	5,930,087	3,777,340
Normalization factor miRNA rpm	0.123	0.134	0.138	0.147	0.118	0.157
<b>Normalization factor Dmel rpm</b>	<b>0.030</b>	<b>0.031</b>	<b>0.032</b>	<b>0.032</b>	<b>0.027</b>	<b>0.037</b>

**Supplementary file 7. Annotation of small RNA libraries from *BX2<sup>OFF</sup>* raised at 25°C or 29°C.** Small RNAs were prepared from ovaries of females of the indicated genotype. Values for the different categories of sequences are the total number of sequence reads that matched reference libraries. For comparisons, libraries were normalized (normalization factor) to 1 million miRNA (miRNA rpm) or to 1 million Dmel reads (Dmel rpm).



Repression occurrence	Number of G2 females		
	none	partial	complete
with $P(TARGET)^{GS}$	1789	24	17
without $P(TARGET)^{GS}$	1136	1	0

**Supplementary file 8.  $P(TARGET)^{GS}$  requirement in the *BX2* conversion process.** Comparison of the conversion frequency in one generation between *BX2*,  $P(TARGET)^{GS}$  (Figure 3) and *BX2* (Figure 3—figure supplement 1) genotypes. The difference between the presence and absence of the  $P(TARGET)^{GS}$  transgene is highly significant ( $p=8.5 \times 10^{-6}$ , homogeneity  $\chi^2 = 23.35$  with 2 degrees of freedom).

<b><i>BX2<sup>OFF</sup></i> recombined in a <i>P(TARGET)<sup>GS</sup></i> genetic background</b>									
<b>Lines</b>	<b>#2</b>	<b>#5</b>	<b>#8</b>	<b>#12</b>	<b>#13</b>	<b>#16</b>	<b>#17</b>	<b>#19</b>	<b>Total</b>
<b>G1</b>	0/20	0/68	0/78	0/21	0/30	0/57	0/59	0/56	0/389
<b>G5</b>	0/82	0/50	0/115	0/86	0/107	0/86	0/60	0/87	0/673
<b>G13</b>	0/128	0/137	0/165	0/108	0/140	0/151	0/103	0/43	0/975
<b>G30</b>	0/27	0/37	0/25	0/24	0/22	0/51	0/33	nt	0/219
<b>Total</b>	0/257	0/292	0/383	0/239	0/299	0/345	0/255	0/186	0/2256

**Supplementary file 9. Silencing capacities of *BX2<sup>OFF</sup>* lines recombined in a *P(TARGET)<sup>GS</sup>* background throughout generations developed at 29°C.** *BX2<sup>OFF</sup>* was initially recombined with a line carrying the *P(TARGET)<sup>GS</sup>* transgene to obtain the *BX2<sup>OFF</sup>*, *P(TARGET)<sup>GS</sup>* lines. From these crosses, eight independent recombinants without the *P(TARGET)<sup>GS</sup>* transgene were recovered and were further cultured at 29°C. To test if some of them acquired silencing capacities, females were crossed with males harboring the *P(TARGET)<sup>GS</sup>* transgene and their progeny was stained for  $\beta$ -Galactosidase expression. Numbers show the fraction of females harboring complete germline repression of *P(TARGET)<sup>GS</sup>* at each generation. A complete stability of the initial epigenetic OFF state was observed for all recombinant lines.

Lines	25°C		29°C							
	1	2	V4	V6	n6	n7	n8	n9	n10	
G1			0/92	0/16	0/51	0/42	0/128	5/58	0/108	
G2	0/22	0/32	0/106	0/32	0/221	0/117	3/261	30/153	0/269	
G3			0/86	0/21	0/150	1/167	5/185	61/160	0/208	
G4			0/45	13/41	0/138	0/146	19/161	43/88	0/142	
G5			1/37	25/56	0/106	0/101	16/83	40/83	0/82	
G6			1/17	49/73	0/111	0/108	16/89	65/95	0/74	
G7			2/9	111/151	0/85	0/75	0/92	76/130	0/29	
G8	0/22	0/15	3/24							
G9			9/12							
G10			0/23	38/63						
G11			1/3	69/106						
G12	0/20	0/23	2/25	44/67						
G13			8/75	139/200						
G14			9/116							
G15				141/193						
G16										
G17			43/130							
G18			83/249							
G19			119/284							
G20	0/14	0/14	49/118							
G21										
G22										
G23										
G24	0/7	0/13								
Total	0/92	0/97	330/1451	629/1019	0/862	1/756	59/999	320/767	0/912	
%R	0% n=189		19.79% n=6766							

**Supplementary file 10. Silencing capacities of  $BX2^{OFF}$ ;  $P(TARGET)^G$  lines throughout generations at 25°C and at 29°C.** Numbers show the fraction of females harboring complete germline repression of  $P(TARGET)^G$  at each generation. Complete stability of the initial epigenetic state was observed at 25°C for  $BX2^{OFF}$ , 0% repression (n = 189). At 29°C,  $BX2^{OFF}$ ;  $P(TARGET)^G$  lines showed emergence of silencing capacities, 19.79% (n = 6766).

	29°C	Control 25°C	+ Heat Shock	+ NaCl
<i>BX2<sup>OFF</sup>, P(TARGET)<sup>GS</sup></i>	586/≈21720	0/≈1680	0/≈3840	0/≈4200

**Supplementary file 11. Heat shock and saline stresses do not induce conversion of *BX2<sup>OFF</sup>*.** *BX2<sup>OFF</sup>, P(TARGET)<sup>GS</sup>* flies were raised during one generation either on classical cornmeal medium (control) at 25°C or were heat shocked for 1 hr at 37°C at the 0–2 hr embryo stage or were cultured on medium supplemented with 150 mM NaCl. The female progeny were then stained for β-Galactosidase expression and egg chambers were individually monitored in order to detect any possible conversion event. No repressed egg chambers were observed (0/total number of egg chambers). Compared to results obtained at 29°C (from data observed in G1 in [Figure 3](#)), differences are significant: for heat shock experiment,  $p=7.2\times 10^{-25}$ , homogeneity  $\chi^2 = 106.03$  with 2 degrees of freedom and for NaCl experiment,  $p=4.9\times 10^{-27}$ , homogeneity  $\chi^2 = 115.93$  with 2 degrees of freedom.

## 2. Effet du froid

La régulation de l'élément *P* est sensible à la température. A une température élevée (26,5°C), ses capacités régulatrices augmentent (Ronsseray *et al*, 1984). De surcroît, à une température élevée (29°C), le cluster *BX2<sup>OFF</sup>* peut être activé pour la production de piRNA dès la première génération (Casier *et al*, 2019/Article I). Ces travaux suggèrent qu'un système de régulation des ET dépendant des piRNA peut se mettre en place à une température élevée (26,5°C ou 29°C). A l'inverse, les capacités régulatrices de l'élément *P* n'émergent pas (Ronsseray *et al*, 1984) à une température plus basse (18°C). En conséquence, on peut se demander si *BX2<sup>ON</sup>* peut être inactivé à une température plus basse.

Pour tester l'effet d'une température plus basse sur l'état épigénétique de *BX2*, nous avons incubé des lignées portant le cluster *BX2<sup>OFF</sup>* ou *BX2<sup>ON</sup>* à 18°C pendant 20 générations puis nous avons compté le nombre de chambres ovariennes qui présentent de la répression du transgène *PlacZ* cible (=P(*TARGET*)) à différentes générations. Ces résultats sont présentés dans la Table 1.

A 25°C, l'état épigénétique de *BX2* ne change pas au cours des générations ; en effet lorsque *BX2* est inactif (*BX2<sup>OFF</sup>*), on ne détecte pas de répression du transgène cible (0%, n=81960; voir Annexe 1) et lorsque *BX2* est actif (*BX2<sup>ON</sup>*), la répression du transgène cible est totale (100%, n=106320 ; voir Annexe 1). A 18°C, *BX2<sup>OFF</sup>* ne change pas d'état épigénétique puisqu'on n'observe pas de répression du transgène cible (0%, n=15755). En revanche, dans le cas de *BX2<sup>ON</sup>* on observe que la répression du transgène cible n'est pas toujours totale au cours des générations (98.26%, n=99375), et la différence de pourcentage de répression entre 18°C (Table 1, Annexe 2) et 25°C (Annexe 1) est significative (p-value=0.005, test du  $\chi^2$ ). Néanmoins, la diminution du pourcentage de répression du transgène cible pour ces sous-lignées est très faible ; on estime que le taux de conversion (ou d'inactivation) de *BX2<sup>ON</sup>* vers *BX2<sup>OFF</sup>* en une génération est inférieur à 0.027% (n=3705, effectifs disponibles dans l'Annexe 2). On peut également constater que le pourcentage de répression ne diminue pas de manière progressive au cours des générations, par exemple pour la sous-lignée 1 le pourcentage de répression passe de 100 à 75% entre les générations 3 et 5 puis remonte à 93.3% à la génération 6 avant d'atteindre de nouveau 100% à la génération 10.

Sous-lignées	<i>BX2<sup>ON</sup></i>								<i>BX2<sup>OFF</sup></i>	
	1	2	3	4	5	6a	6b	7	1	2
<b>G1</b>	100%	100%	100%	100%	100%	100%	100%	100%	0%	0%
<b>G2</b>	99.8%	100%	100%	99.3%	98.4%	100%	99.9%	100%	0%	0%
<b>G3</b>	100%	100%	100%	100%	90%	98.9%	100%	99.5%	0%	0%
<b>G5</b>	75%	100%	100%	99.8%	100%	97.6%	100%	100%	0%	0%
<b>G6</b>	93.3%	100%	100%	100%	92.7%	100%	100%	100%	0%	0%
<b>G8</b>	99.8%	100%	100%	99.9%	100%	100%	100%	100%	0%	nt
<b>G10</b>	100%	100%	100%	100%	93.3%	100%	100%	100%	0%	0%
<b>G15</b>	99.6%	100%	100%	100%	99.9%	100%	100%	100%	0%	0%
<b>G20</b>	99.9%	100%	100%	100%	79%	98.7%	100%	100%	0%	0%
<b>Total</b>	98.26% n=99375								0% n=15755	

**Table 1 :** Pourcentages de répression de sous-lignées *BX2<sup>OFF</sup> P(TARGET)* et *BX2<sup>ON</sup> P(TARGET)* à 18°C au cours des générations. A 25°C, les sous-lignées *BX2<sup>OFF</sup>* ne présentent pas de répression du transgène cible *P(TARGET)* (0%, n=81960 ; voir Annexe 1) tandis que les sous-lignées *BX2<sup>ON</sup>* présentent une répression complète au cours des générations (100%, n=106320 ; voir Annexe 1). A 18°C, les lignées *BX2<sup>OFF</sup>* ne présentent pas de répression du transgène cible (0%, n=15755) et certaines lignées *BX2<sup>ON</sup>* ne présentent pas une répression complète au cours des générations (98.26%, n=99375). Pour plus de clarté, seuls les pourcentages de répression sont représentés dans cette table. Ces pourcentages de répression sont calculés à partir du nombre de chambres ovariennes réprimées comptées à chaque génération ; les effectifs sont détaillés dans l'Annexe 2. nt=non testé.

Le faible pourcentage d'inactivation en une génération pourrait expliquer pourquoi certaines sous-lignées (les sous-lignées 2 et 3) restent à 100% tandis que les autres ne présentent pas un

pourcentage de répression stable à 100%. A ce faible pourcentage d'inactivation s'ajoute un biais d'échantillonnage à chaque génération. En effet, pour une génération, tous les individus ne participent pas à la génération suivante ; soit un individu ne donne pas de descendance viable, soit les descendants ne survivent pas jusqu'au moment de la dissection et ne font ainsi partie de l'analyse. Par conséquent, l'émergence d'une sous-lignée complètement inactive (0% de répression du transgène cible) est difficile à 18°C. Parallèlement, Valérie Delmarre maintient des stocks de lignées *BX2<sup>ON</sup>* à 18°C et elle a pu constater que certaines de ces lignées ne se maintiennent pas toujours à 100% de répression, voire elles atteignent les 0% de répression après une quarantaine de générations (Communication personnelle). Ces résultats suggèrent que la diminution de la température peut déstabiliser l'état épigénétique du cluster *BX2<sup>ON</sup>* mais n'a pas d'effet sur le cluster lorsqu'il est déjà inactif (*BX2<sup>OFF</sup>*).

On peut imaginer qu'une température inférieure à 18°C pourrait déstabiliser davantage l'état épigénétique du cluster *BX2<sup>ON</sup>*. Pour le tester, nous avons incubé des lignées *BX2* à 10°C pendant 8 semaines (= génération 0, Table 2). A 10°C, les femelles ont une longévité plus importante qu'à 25°C (entre 4 et 6 mois à 10°C contre environ 1 mois à 25°C) mais ne pondent pas, on ne peut donc pas suivre l'état épigénétique de *BX2* pendant plusieurs générations en les maintenant à cette température. Après le traitement de 8 semaines à 10°C, les femelles sont incubées à 25°C pendant 5 jours avant d'être disséquées. Cette étape permet aux femelles de pondre et ainsi d'estimer l'effet du traitement sur la génération suivante (= génération 1, Table 2). Les résultats obtenus sont présentés dans la Table 2.

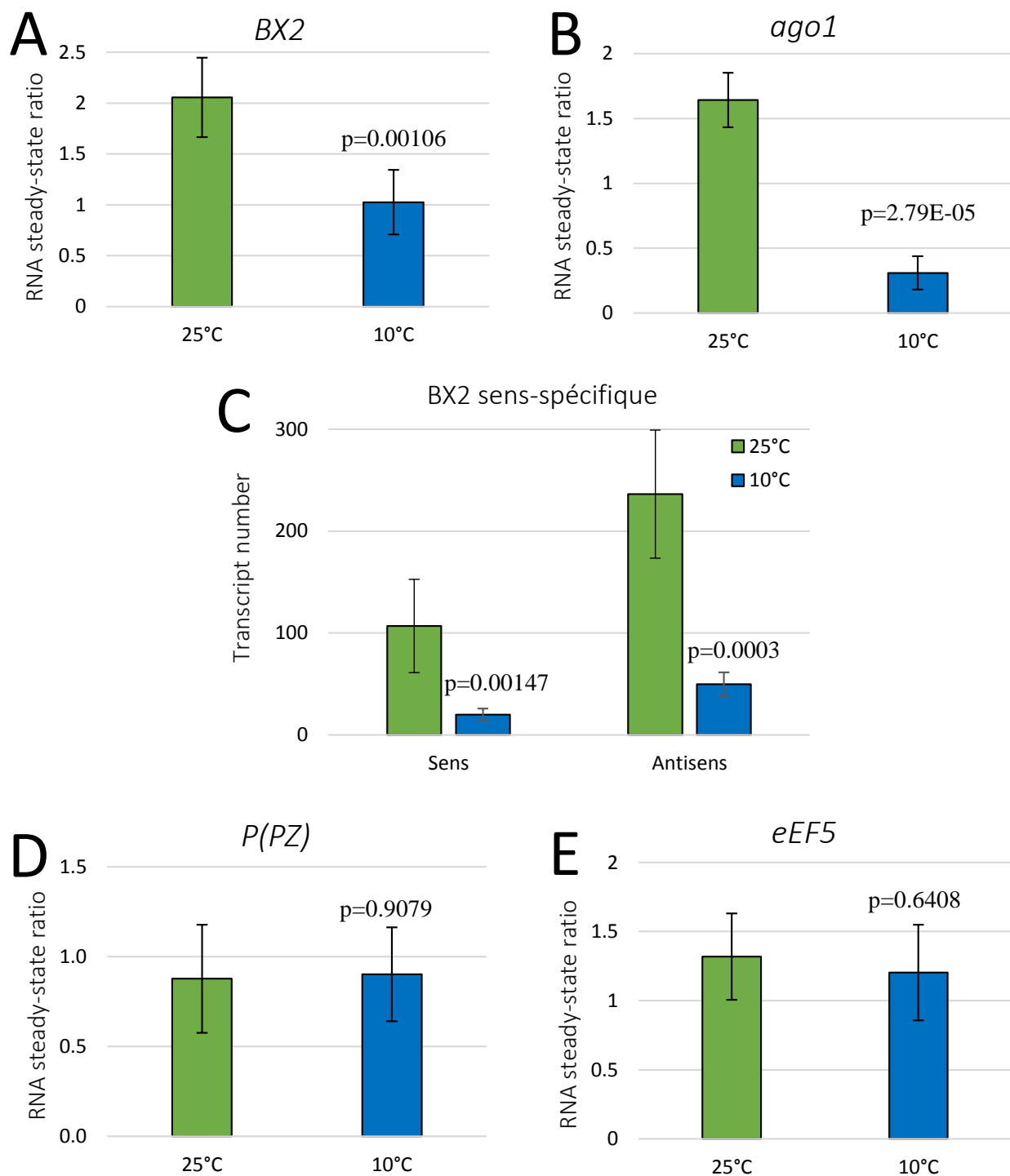
Après 8 semaines à 10°C, on constate que le cluster *BX2<sup>OFF</sup>* ne change pas d'état épigénétique puisqu'on n'observe pas de répression du transgène cible (0%, n=10510). De manière cohérente on n'observe pas non plus de changement d'état épigénétique dans la descendance qui s'est développée à 25°C (0% de répression, n=2880). Dans le cas du cluster *BX2<sup>ON</sup>*, on peut observer que le pourcentage de répression du transgène cible n'est pas total après 8 semaines à 10°C (99.12%, n=34920), cependant, par rapport à des lignées *BX2<sup>ON</sup>* à 25°C (100%, n=3480, Annexe 1) ces résultats ne sont pas statistiquement différents (p-val=0.72, test du  $\chi^2$ ). Dans la descendance des individus incubés à 10°C, on observe que le pourcentage de répression n'est pas total (93.28%, n=5460 ; p-val=0.02, test du  $\chi^2$ ). De même qu'à 18°C, à 10°C l'état épigénétique du cluster *BX2<sup>ON</sup>* est affecté. Ce changement d'état épigénétique peut ensuite être maintenu à la génération suivante. En revanche, cette température n'a pas d'effet sur l'état épigénétique du cluster *BX2<sup>OFF</sup>*.

		Génération 0		Génération 1	
		C-O réprimées / C-O totales	%R	C-O réprimées / C-O totales	%R
<b><i>BX2<sup>OFF</sup></i></b>	SI 1	0/1800	0%	0/960	0%
	SI 2	0/5040	0%	0/960	0%
	SI 3	0/3720	0%	0/960	0%
	<b>Total</b>	<b>0/10510</b>	<b>0%</b>	<b>0/2880</b>	<b>0%</b>
<b><i>BX2<sup>ON</sup></i></b>	SI 1	7399/7440	99.5%	nt	nt
	SI 2	13563/13680	99.1%	1440/1400	100%
	SI 3	2038/2040	99.9%	nt	nt
	SI 4	2385/2400	99.4%	1433/1440	99.51%
	SI 5	1320/1320	100%	nt	nt
	SI 6	3750/3840	97.7%	1260/1260	100%
	SI 7	4160/4200	99.1%	960/1320	72.73%
	<b>Total</b>	<b>34615/34920</b>	<b>99.12%</b>	<b>5093/5460</b>	<b>93.28%</b>

**Table 2 :** Pourcentages de répression de sous-lignées *BX2<sup>OFF</sup>* *P(TARGET)* et *BX2<sup>ON</sup>* *P(TARGET)* après un traitement de 8 semaines à 10°C (=Génération 0) et de leur descendance, développée à 25°C (=Génération 1). Pour chaque sous-lignée, le nombre de chambres présentant de la répression sur le nombre total de chambres est indiqué ; le pourcentage de répression est calculé à partir de ces valeurs. A 25°C, les sous-lignées *BX2<sup>OFF</sup>* ne présentent pas de répression du transgène cible *PlacZ* (0%, n=81960 ; Voir Annexe 1) tandis que les sous-lignées *BX2<sup>ON</sup>* présentent une répression complète au cours des générations (100%, n=106320 ; Voir Annexe 1). Après 8 semaines à 10°C, les sous-lignées *BX2<sup>OFF</sup>* ne présentent pas de répression du transgène cible (0%, n=10510) et certaines sous-lignées *BX2<sup>ON</sup>* ne présentent pas une répression complète (99.12%, n=34920). Pour la descendance développée à 25°C (= génération 1), les sous-lignées *BX2<sup>OFF</sup>* ne présentent pas de répression du transgène cible *PlacZ* (0%, n=2880) et les sous-lignées *BX2<sup>ON</sup>* ne présentent pas une répression complète (93.28%, n=5460). SI : sous-lignée ; C-O : chambres ovariennes ; nt : non testé ; %R : pourcentage de répression.

Nous avons cherché à savoir si le changement d'état épigénétique de *BX2<sup>ON</sup>* est associé à des changements au niveau de la quantité de transcrits de *BX2* ainsi que du transgène *PlacZ* cible (*P(PZ)*), mais également des gènes dans lesquels ces transgènes sont insérés – respectivement *ago1* et *eEF5*. Pour ce faire, Chloé Laurin, une étudiante en L3 que j'ai encadrée lors de son stage, a réalisé des expériences de RT-qPCR sur des ovaires de femelles incubées 8 semaines à 10°C ou maintenues à 25°C. Ces résultats sont présentés dans la figure 28.





**Figure 28** : Mesure des quantités relatives de transcrits par RT-qPCR dans les ovaires de différents acteurs moléculaires à 10°C (en bleu) et à 25°C (en vert). (A, B) A 10°C, on détecte moins de transcrits de *BX2* et d'*ago1* – gène dans lequel *BX2* est inséré – comparé à 25°C. (C) A 10°C, les transcrits sens et antisens de *BX2* sont moins abondants qu'à 25°C. (D, E) La quantité relative de transcrits du transgène cible *lacZ* (*P(PZ)*) et de *eEF5* – le gène dans lequel le transgène cible *P(PZ)* est inséré – est la même à 10°C et à 25°C. Pour ces expériences, les valeurs de Sq du gène/transgène testé sont normalisées sur les valeurs de Sq du gène *hfp* dont l'expression ne varie pas en fonction de la température et les *p*-values sont calculées grâce à un test de student bilatéral.

On peut constater qu'il y a moins de transcrits de *BX2* et d'*ago1* à 10°C comparé à 25°C (Figure 28A et 28B). De plus, les deux types de précurseurs de *BX2* sont impactés à 10°C puisqu'il y a moins de précurseurs sens et antisens (Figure 28C). En revanche, pour le transgène cible *P(PZ)* et le gène *eEF5*, la quantité relative de transcrits ne change pas entre 10°C et 25°C (Figure 28D et 28E). La diminution du pourcentage de répression qu'on observe après 8 semaines à 10°C est peut-être une conséquence de la diminution de la quantité de transcrits précurseurs – aussi bien sens et antisens - de *BX2* (Figure 28A et 28C). Par ailleurs, *ago1* est transcrit en antisens comparé à l'orientation du cluster *BX2*, ainsi même si *BX2* possède des promoteurs internes, il n'est pas exclu que des précurseurs chimères *ago1-BX2* soient transcrits. Si c'est le cas, la diminution de la quantité de transcrits d'*ago1* à 10°C (Figure 28B) pourrait être responsable en partie de la diminution de la quantité de transcrits antisens de *BX2*. Pour autant, la diminution des deux types de précurseurs, sens et antisens (Figure 28C) suggère que la transcription du cluster *BX2* est sensible à la température indépendamment d'*ago1*.

De manière étonnante, ces résultats sont à l'inverse de ceux obtenus à une température élevée. En effet, dans l'article I nous avons montré qu'à 29°C il y a plus de transcrits précurseurs antisens de *BX2* qu'à 25°C, cette augmentation étant associée à une augmentation de la quantité relative de transcrits d'*ago1* (Casier *et al*, 2019/Article I). Ainsi, la température module la quantité de transcrits précurseurs ainsi que l'état épigénétique de *BX2* : *BX2<sup>OFF</sup>* peut être activé pour la production de piRNA à 29°C mais n'est pas affecté par des températures plus faibles (10°C ou 18°C); à l'inverse *BX2<sup>ON</sup>* ne change pas d'état épigénétique à 29°C mais peut être inactivé pour la production de piRNA lorsque la température est plus faible (10°C ou 18°C). Pour autant, ces résultats ne permettent pas d'établir nettement un lien de cause à effet entre l'augmentation/la diminution de la quantité de transcrits et le changement d'état épigénétique de *BX2*. En effet, la quantité de transcrits n'est pas toujours corrélée à la quantité de piRNA produits et nous savons qu'il n'y a pas de différence de quantité de transcrits entre *BX2<sup>OFF</sup>* et *BX2<sup>ON</sup>* (de Vanssay *et al*, 2012; Robine *et al*, 2009).

### 3. Autres stress

#### 3.1. Sel (NaCl)

En 2011, il a été montré chez la drosophile que l'exposition à des stress thermique et osmotique (chocs thermiques et forte concentration en NaCl) était capable d'induire la phosphorylation de la protéine dAFT2, provoquant une désorganisation de l'hétérochromatine. Ces modifications de l'état chromatinien se traduisent par une augmentation de la variegation dans l'œil des drosophiles (Seong *et al*, 2011). Les auteurs ont également montré que cet effet peut être maintenu à la génération suivante. Dans le « supplementary file 11 » de l'article I, nous avons montré que les chocs thermiques (37°C) et une forte concentration en NaCl n'ont pas d'effet sur l'état épigénétique de *BX2<sup>OFF</sup>*. Nous avons voulu savoir si une forte concentration en NaCl était capable d'induire un changement d'état épigénétique de *BX2<sup>ON</sup>*; en effet, la désorganisation de l'hétérochromatine au niveau du cluster pourrait impacter sa capacité à produire des piRNA. Nous avons donc exposé des drosophiles à un milieu riche en NaCl (0.15 M ou 0.3 M) pendant tout leur temps de développement, de l'embryon jusqu'à l'âge adulte. Les résultats obtenus sont présentés dans la table 3 ci-dessous.

	<b>Pourcentage de répression</b>
Milieu normal	100% (n=1920)
Milieu normal + NaCl 0.15 M	100% (n=7200)
Milieu normal + NaCl 0.3 M	100% (n=1200)

**Table 3 : Effet d'une exposition au NaCl au cours du développement.** Des mouches portant le cluster *BX2<sup>ON</sup>* et un transgène cible *PlacZ* se sont développées sur un milieu normal ou sur un milieu contenant en plus du NaCl (0.15 M ou 0.3 M). A 0.3M de NaCl dans le milieu on observe un retard de croissance et une létalité très importante des individus. La présence de NaCl dans le milieu n'affecte pas l'état épigénétique du cluster *BX2<sup>ON</sup>* qui garde une capacité répressive totale (100%). Les effectifs correspondent au nombre de chambres ovariennes comptées.

On peut constater que le pourcentage de répression est toujours total dans un milieu riche en NaCl (à 0.15 M ou à 0.3 M). Ainsi, l'exposition au NaCl ne modifie pas l'état épigénétique de *BX2<sup>ON</sup>*.

### 3.2. Cuivre (CuSO<sub>4</sub>)

Dans l'article I, nous avons montré qu'il y a plus de transcrits antisens de *BX2* à 29°C qu'à 25°C. De plus, le gène *ago1* dans lequel *BX2* est inséré est plus exprimé à 29°C qu'à 25°C. Dans l'article, nous avons émis l'hypothèse qu'à 29°C, le gène *ago1* étant plus exprimé, des transcrits chimères *ago1-BX2* sont produits en plus grande quantité. Sur le site Flybase (<https://flybase.org/>), les données d'expression de gènes après différents traitements sont répertoriées et on peut y voir qu'une exposition au cuivre (CuSO<sub>4</sub>, 15 mM) pendant 48 heures active la transcription d'*ago1*. Cependant le ou les tissu(s) étudié(s) ne sont pas indiqué(s). Pour savoir si le cuivre peut modifier l'état épigénétique de *BX2*, nous avons exposé des femelles à un milieu riche en cuivre (3 mM ou 15 mM) pendant 12 jours suivi de 2 à 3 jours sur un milieu normal pour obtenir la génération suivante (G1). Les G1 se développent ensuite sur milieu normal. Les résultats sont présentés dans la table 4.

Cluster	Milieu normal	CuSO <sub>4</sub> 3mM (G0)	CuSO <sub>4</sub> 3mM (G1)	CuSO <sub>4</sub> 15mM (G0)
<i>BX2<sup>OFF</sup></i>	0% (n=960)	0% (n=1320)	0% (n=4800)	0% (n=4920)
<i>BX2<sup>ON</sup></i>	100% (n=480)	100% (n=360)	100% (n=5160)	100% (n=240)

**Table 4 : Effet d'une exposition au CuSO<sub>4</sub> au cours du développement.** Des mouches portant le cluster *BX2<sup>OFF</sup>* ou *BX2<sup>ON</sup>* et un transgène cible se sont développées sur un milieu normal ou riche en CuSO<sub>4</sub> (3 mM ou 15 mM) pendant 12 jours, suivi de trois jours sur un milieu normal pour permettre aux femelles (G0) de donner la génération suivante (G1) qui se développe sur un milieu normal. A 3 mM, les ovaires sont atrophiés mais les femelles survivent et pondent ; à 15 mM, les ovaires sont très atrophiés également et les femelles pondent très peu, les individus qui se développent sur milieu normal meurent au stade L1. L'exposition au cuivre à 3 mM ou à 15 mM n'affecte pas le pourcentage de répression de *BX2<sup>OFF</sup>* ou *BX2<sup>ON</sup>*. Les effectifs correspondent au nombre de chambres ovariennes comptées.

L'exposition au cuivre n'a pas d'effet sur le pourcentage de répression des clusters  $BX2^{OFF}$  et  $BX2^{ON}$ . Ce stress ne semble donc pas capable de modifier l'état épigénétique de  $BX2$ .

### 3.3. Bisphénol A

Dans une étude parue en 2018, il a été montré que l'exposition de nématodes au bisphénol A provoque une diminution de la quantité de H3K9me3 au niveau d'un transgène rapporteur GFP, normalement réprimé dans les cellules germinales. De manière plus globale, le traitement au bisphénol A modifie la répartition des marques H3K9me3 et H3K27me3 sur l'ensemble du génome, ce qui se traduit par une modification de l'expression de nombreux gènes. Ces effets génomiques sont ensuite hérités jusqu'à 4 générations après le traitement grâce à l'activité de déméthylases d'histones à domaine jumonji (*JMJD-2* et *JMJD-3*) (Camacho *et al*, 2018). Nous avons voulu tester l'effet du bisphénol A sur l'état épigénétique de  $BX2$  car nous avons découvert que l'inactivation d'une déméthylase d'histone à domaine jumonji, *JHDM2*, chez la drosophile entraîne l'activation de  $BX2$  (Partie II des résultats). Nous avons exposé des lignées à un milieu contrôle contenant du DMSO ou un milieu contenant du DMSO et du bisphénol A (100  $\mu$ M) pendant 8 jours suivi de trois jours sur un milieu normal pour obtenir la génération suivante. Les résultats sont présentés dans la table 5.

Cluster	Génération 0		Génération 1	
	DMSO	DMSO + BPA	DMSO	DMSO + BPA
$BX2^{OFF}$	0% (n=1800)	0% (n=2640)	0% (n=2400)	0% (n=3960)
$BX2^{ON}$	100% (n=1800)	100% (n=3000)	100% (n=2400)	100% (n=4320)

**Table 5:** Effet d'une exposition au bisphénol A (BPA) au cours du développement. Des mouches portant le cluster  $BX2^{OFF}$  ou  $BX2^{ON}$  se sont développées sur un milieu contrôle contenant du DMSO ou un milieu contenant DMSO + BPA (100  $\mu$ M) pendant 8 jours, suivi de trois jours sur un milieu normal pour permettre aux femelles (G0) de donner la génération suivante (G1) qui se développe sur un milieu normal. L'exposition au BPA n'affecte pas le pourcentage de répression de  $BX2^{OFF}$  ou  $BX2^{ON}$ . Les effectifs correspondent au nombre de chambres ovariennes comptées.

L'exposition au bisphénol A n'a pas d'effet sur le pourcentage de répression des clusters *BX2<sup>OFF</sup>* et *BX2<sup>ON</sup>*. Il est possible que le traitement au bisphénol A affecte également l'activité de certaines protéines à domaine jumonji chez la drosophile mais pas l'activité de *JHDM2*. Le bisphénol A ne modifie pas l'état épigénétique de *BX2*.

## II. Effet de facteurs génétiques dans l'activation de *BX2*

En 2013, trois cribles perte de fonction ont permis d'identifier une centaine de facteurs impliqués dans la biogénèse des piRNA dans les cellules folliculaires et germinales de la drosophile (Czech *et al*, 2013; Handler *et al*, 2013; Muerdter *et al*, 2013). En 2014, à l'inverse de ces cribles, Antoine Boivin a initié un crible de perte de fonction pour identifier des facteurs dont l'absence permettrait l'activation de la production de piRNA par *BX2<sup>OFF</sup>*, et ce en l'absence d'héritage de piRNA maternels. Si la perte de fonction d'un facteur est capable d'activer la production de piRNA, ceci implique qu'un tel facteur joue un rôle actif dans le maintien du cluster *BX2* à l'état inactif, c'est-à-dire dans un état de non production de piRNA. Grâce à ce crible génétique, il a pu identifier deux gènes donc la perte de fonction entraîne une activation de *BX2<sup>OFF</sup>* : *hfp* et *JHDM2*. J'ai participé à ce projet pour lequel j'ai testé un peu plus de 80 gènes, établi le lien moléculaire entre *hfp* et *JHDM2* ainsi que leur implication dans l'activation de *BX2<sup>OFF</sup>*, étudié les effets pan-génomiques en contexte *JHDM2 GLKD* et les origines des phénotypes observés chez les descendants des individus *JHDM2 GLKD*.

### 1. Perte de fonction de *hfp* et activation de *BX2*

#### 1.1. Description du crible génétique

Le crible génétique 'perte de fonction' mis au point par Antoine est basé sur la technique d'ARN interférence (RNAi) permettant l'inactivation de gènes d'intérêts. Le 'Bloomington Drosophila Stock Center' (BDSC) de l'université d'Indiana entretient de nombreuses lignées, dont des lignées RNAi. Ces lignées portent un transgène (appelé *UAS-gene<sup>KD</sup>*) contenant des séquences UAS permettant l'expression tissu-spécifique d'un short hairpin RNA (shRNA) capable d'induire une réponse RNAi conduisant à la diminution de l'expression du gène ciblé par le shRNA (Plus de détails dans le matériel et méthodes). Les femelles de ces lignées sont croisées par des mâles portant le cluster *BX2<sup>OFF</sup>*, un transgène *nos-gal4* et un transgène cible *PlacZ (P(TARGET))*. En présence des transgènes *nos-gal4* et *UAS-gene<sup>KD</sup>*, la protéine GAL4 dont la production est sous le contrôle du promoteur du gène *nanos (nos)*, est produite spécifiquement dans les ovaires. Ainsi, le shRNA est exprimé spécifiquement dans les

cellules germinales, où la protéine GAL4 est produite et reconnaît les séquences UAS. Dans la descendance, les femelles portant le cluster *BX2*<sup>OFF</sup>, le transgène *P(TARGET)*, le transgène *nos-gal4* et le transgène *UAS-gene*<sup>KD</sup> sont sélectionnées, les ovaires disséqués et l'état épigénétique du cluster *BX2* est analysé par coloration β-galactosidase (Figure 29A-D). Si les ovaires sont colorés (coloration bleue des cellules germinales), cela signifie que le *BX2* est à l'état inactif et que la perte de fonction du gène ciblé n'affecte pas l'état épigénétique du cluster. A l'inverse, si les ovaires ne sont pas colorés (pas de coloration bleue des cellules germinales), cela signifie que *BX2* est actif et produit des piRNA capables de réprimer l'expression du transgène cible *P(TARGET)* ; la perte de fonction du gène ciblé affecte alors l'état épigénétique de *BX2*<sup>OFF</sup>.

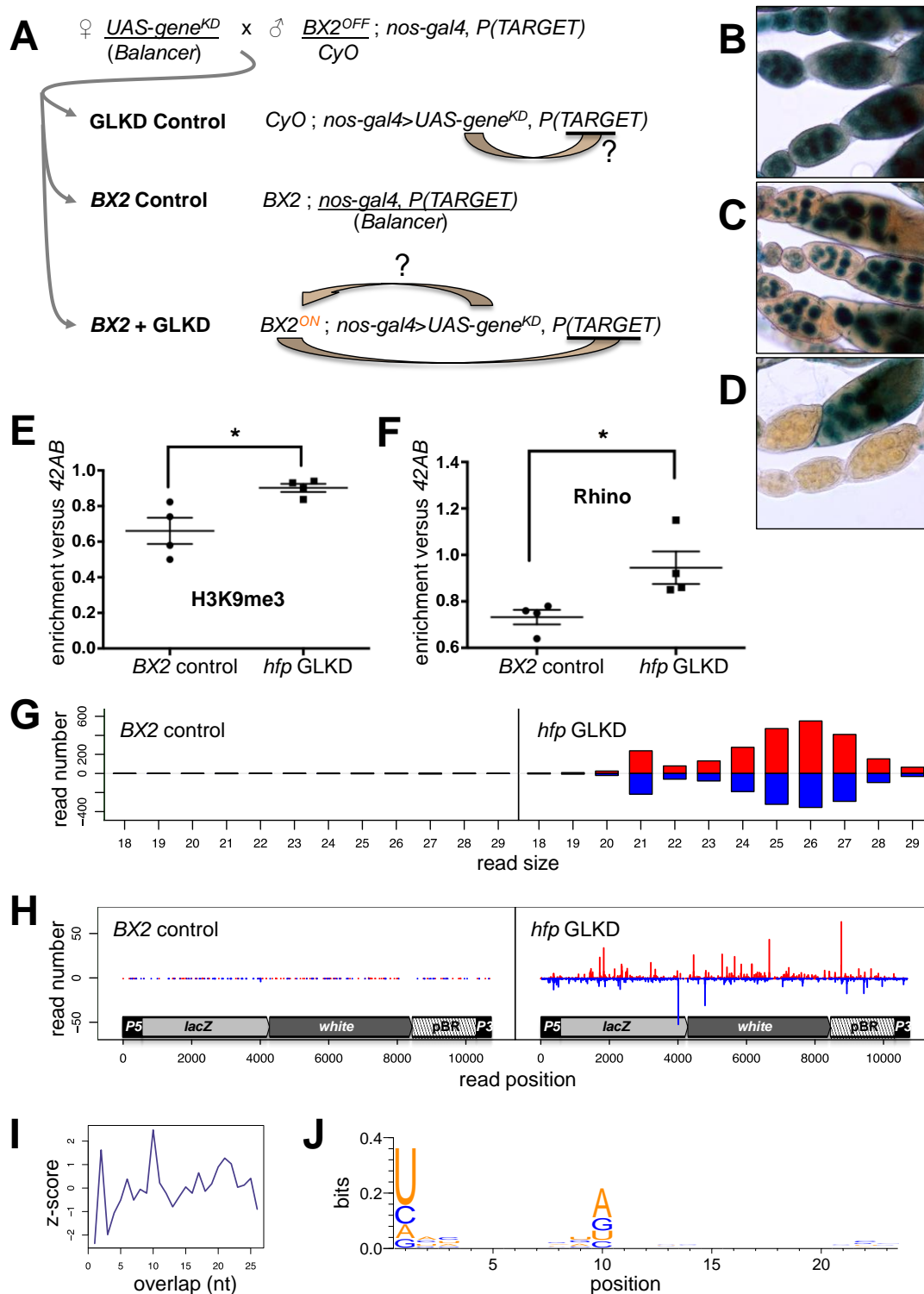
Avec ce crible génétique, nous avons testé 272 gènes impliqués dans des fonctions biologiques différentes : gènes impliqués dans l'organisation de la chromatine (incluant notamment les gènes des groupes Polycomb et Trithorax), dans la réparation de l'ADN, dans la réponse à la température ou encore des facteurs d'épissage (Liste complète des gènes et shRNA testés dans l'Annexe 3). Parmi ces gènes, nous avons identifié, au départ, un seul gène, *half-pint* (*hfp*), dont la perte de fonction induit une activation du cluster *BX2*<sup>OFF</sup>.

## 1.2. Analyse de la perte de fonction de *hfp*

*hfp* code un facteur impliqué dans la transcription et l'épissage. La diminution d'expression de *hfp* dans les ovaires (appelée *hfp GLKD*, GLKD pour Germ Line Knock Down), entraîne la conversion de *BX2*<sup>OFF</sup> dans la plupart des chambres ovariennes (87% de répression du transgène cible *P(TARGET)*, figure 29D). Les contrôles confirment que le transgène *UAS-gene*<sup>KD</sup> seul n'affecte pas l'expression du transgène cible *P(TARGET)* (Figure 29B) et que le cluster *BX2* ne peut pas être converti en l'absence du transgène *UAS-gene*<sup>KD</sup> (Figure 29C). Les analyses moléculaires indiquent qu'en contexte *hfp GLKD*, le cluster *BX2* produit des piRNA sens et antisens (Figures 29G-H). Ces piRNA possèdent un biais 1U et 10A et arborent une signature Ping-Pong, indiquant qu'il y a des piRNA primaires et secondaires (Figure 29I-J). Les expériences de ChIP-qPCR indiquent que *BX2* est enrichi en marques H3K9me3 (Figure 29E) et Rhino (Figure 29F) en contexte *hfp GLKD*. Ainsi, le cluster *BX2*, lorsqu'il est activé par la perte de fonction de *hfp*, possède toutes les caractéristiques moléculaires d'un piRNA cluster double-brin. Pour exclure la possibilité d'un 'off-target effect' du transgène *UAS-hfp*<sup>KD</sup>



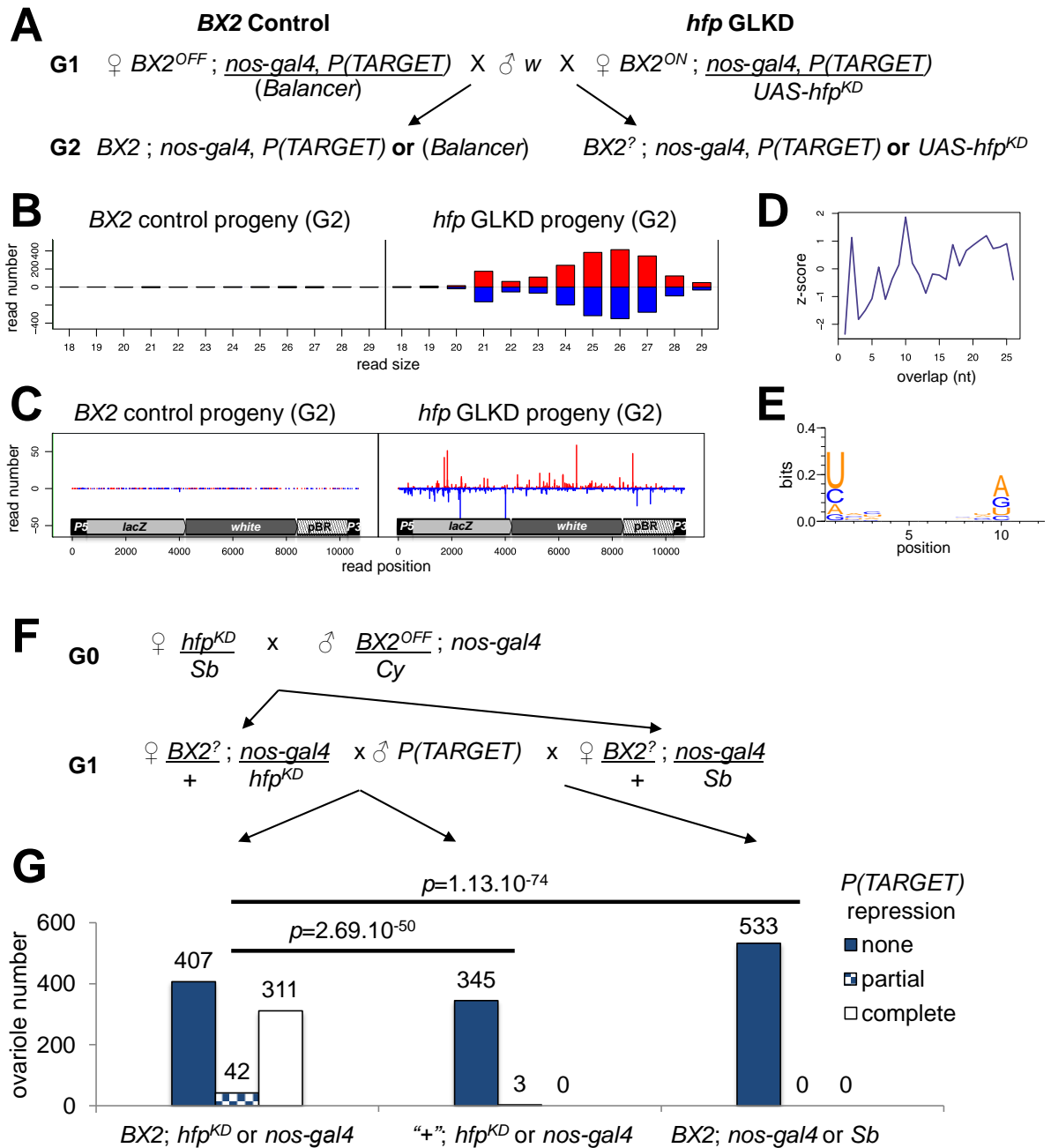
utilisé, l'activation de *BX2* en contexte de perte de fonction pour *hfp* a été confirmée à l'aide de différentes combinaisons de vrais mutants (Annexe 4).



**Figure 29** : La perte de fonction de *hfp* entraîne l'activation de *BX2*<sup>OFF</sup> pour la production de piRNA. (A) Principe du crible génétique : Des femelles portant un transgène *UAS-gene*<sup>KD</sup> sont croisées par des mâles portant le cluster *BX2*<sup>OFF</sup>, un transgène *nos-gal4* et un transgène cible *P(TARGET)*. Dans la descendance, les femelles portant tous ces éléments

génétiqes (= BX2+GLKD) sont sélectionnées et leurs ovaires sont analysés par coloration  $\beta$ -galactosidase. Si les ovaires sont colorés, c'est que la perte de fonction du gène ciblé n'affecte pas l'état épigénétique de  $BX2^{OFF}$  ; en revanche lorsque les ovaires ne sont pas colorés, alors la perte de fonction du gène affecte l'état épigénétique de  $BX2^{OFF}$ . Dans le cas où  $BX2$  change d'état épigénétique, deux autres contrôles sont réalisés : le premier avec des femelles ne portant pas le cluster  $BX2$  (= GLKD control) et le second avec des femelles ne portant pas le transgène  $UAS-gene^{KD}$  (=  $BX2$  control). Ces contrôles permettent de confirmer respectivement que le transgène  $UAS-gene^{KD}$  seul n'affecte pas l'expression du transgène cible  $P(TARGET)$  et que  $BX2^{OFF}$  n'est pas converti même en l'absence du transgène  $UAS-gene^{KD}$ . GLKD : Germ Line Knock Down. (B) Le transgène  $UAS-gene^{KD}$  seul n'a pas d'effet sur l'expression du transgène cible  $P(TARGET)$ . Dans toutes les chambres ovariennes observées, les noyaux des cellules germinales sont bleus après coloration  $\beta$ -galactosidase (0% de répression, n=447 ovarioles) (C)  $BX2$  reste inactif même en l'absence du transgène  $UAS-gene^{KD}$  (0% de répression, n=578 ovarioles) (D) La perte de fonction de  $hfp$  dans la lignée germinale ( $hfp$  GLKD) entraîne l'activation de  $BX2^{OFF}$  pour la production de piRNA, le transgène  $P(TARGET)$  est réprimé dans la plupart des chambres ovariennes (87% d'ovarioles présentant de la répression, n=974 ovarioles) (E, F) Les expériences d'immunoprécipitation de chromatine suivies de qPCR (ChIP-qPCR) montrent que  $BX2$  est enrichi en marques H3K9me3 (E) et Rhino (F) en condition  $hfp$  GLKD par rapport au contrôle ( $BX2$  control), \* pvalue < 0.05 (Test Mann Whitney, n=4) (G, H) Les expériences de small RNA-seq indiquent que des piRNA (23 à 29 nt) sont produits par le cluster  $BX2$  en contexte  $hfp$  GLKD alors qu'ils ne sont pas produits dans le contrôle. En rouge sont représentés les petits ARN sens et en bleu les petits ARN antisens (G) Répartitions par taille des petits ARN entre 18 et 29 nucléotides. Les ARN de 21 nucléotides correspondent aux siRNA tandis que les ARN de 23 à 29 nucléotides correspondent aux piRNA. (H) L'alignement des petits ARN de 23 à 29 nucléotides (unique mappers) montre une répartition uniforme des piRNA produits le long du transgène  $P[lacW]$ . (I) La fréquence relative (z-score) de chevauchement entre les petits ARN sens et antisens de 23 à 29 nucléotides en contexte  $hfp$  GLKD montre un enrichissement de l'appariement en position 10, caractéristique de la signature Ping-Pong. (J) En contexte  $hfp$  GLKD, on observe un enrichissement de U en position 1 et de A en position 10, indiquant que des piRNA primaires et secondaires sont produits.

L'analyse moléculaire des descendants de  $hfp$  GLKD révèle que le cluster  $BX2$  produit des piRNA à la génération suivante et ce même lorsque l'expression de  $hfp$  n'est plus diminuée (Figure 30A-E). Ces résultats indiquent que même lorsqu'on rétablit l'expression de  $hfp$ , l'état épigénétique de  $BX2$  est maintenu. Par ailleurs, à l'inverse de la conversion de  $BX2^{OFF}$  à 29°C, la conversion de  $BX2^{OFF}$  en contexte  $hfp$  GLKD ne nécessite pas la présence du transgène cible  $P(TARGET)$ , suggérant des mécanismes d'activation distincts (Figure 30F-G).



**Figure 30** : La conversion du cluster *BX2<sup>OFF</sup>* en contexte *hfp* GLKD est stable à la génération suivante et ne requiert pas la présence du transgène cible *P(TARGET)*. (A) Croisements permettant d’analyser la descendance (G2) des contrôles (*BX2 control*) et des *hfp GLKD*. Aucune distinction n’est faite parmi les descendants obtenus en G2 et les individus sont traités en masse. (B, C) Les expériences de small RNA-seq indiquent que des piRNA (23 à 29 nt) sont produits par le cluster *BX2* chez les descendants des individus *hfp GLKD* tandis qu’ils ne sont pas produits par les descendants du contrôle (*BX2 control*). En rouge sont représentés les petits ARN sens et en bleu les petits ARN antisens (B) Répartitions par taille des petits ARN entre 18 et 29 nucléotides. Les ARN de 21 nt correspondent aux siRNA tandis que les ARN de 23 à 29 nucléotides correspondent aux piRNA. (C) L’alignement des petits ARN de 23 à 29 nucléotides (unique mappers) montre une répartition uniforme des piRNA produits le long du transgène *P[lacW]*. (D) La fréquence relative (z-score) de chevauchement entre les petits ARN sens et antisens de 23 à 29 nucléotides chez les descendants des *hfp*

*GLKD* montre un enrichissement de l'appariement en position 10, caractéristique de la signature Ping-Pong. (E) Pour les descendants des *hfp GLKD*, on observe un enrichissement de U en position 1 et de A en position 10, indiquant que des piRNA primaires et secondaires sont produits. (F) Schéma de croisements permettant de tester le rôle du transgène cible *P(TARGET)* dans la conversion de *BX2* en contexte *hfp GLKD*. *P(TARGET)* est absent dans les croisements à la génération 0 (G0), contrairement aux expériences précédentes. En conséquence, on ne connaît pas l'état épigénétique de *BX2* à la génération 1 (G1). Le croisement de femelles G1 par des mâles portant le transgène cible *P(TARGET)* (à gauche) permet de connaître l'état épigénétique de *BX2* grâce aux femelles de la génération 2 (G2). Un croisement contrôle est réalisé en parallèle entre des femelles qui ne portent pas le *GLKD* de *hfp* et des mâles portant *P(TARGET)* (à droite). (G) Histogramme montrant le nombre d'ovarioles présentant une répression complète, partielle ou pas de répression de *P(TARGET)*, indiquant la conversion de *BX2*. En G2, seules les femelles *BX2 hfp GLKD* ou *BX2 nos-gal4* présentent de la répression de *P(TARGET)* (à gauche) comparées aux contrôles (au centre et à droite). Les différences avec les contrôles sont par ailleurs très significatives (Test du  $\chi^2$ : les p-values sont indiquées et les  $\chi^2_{\text{obs}}$  sont respectivement de 228.27 et 340.53 avec 2 degrés de liberté).

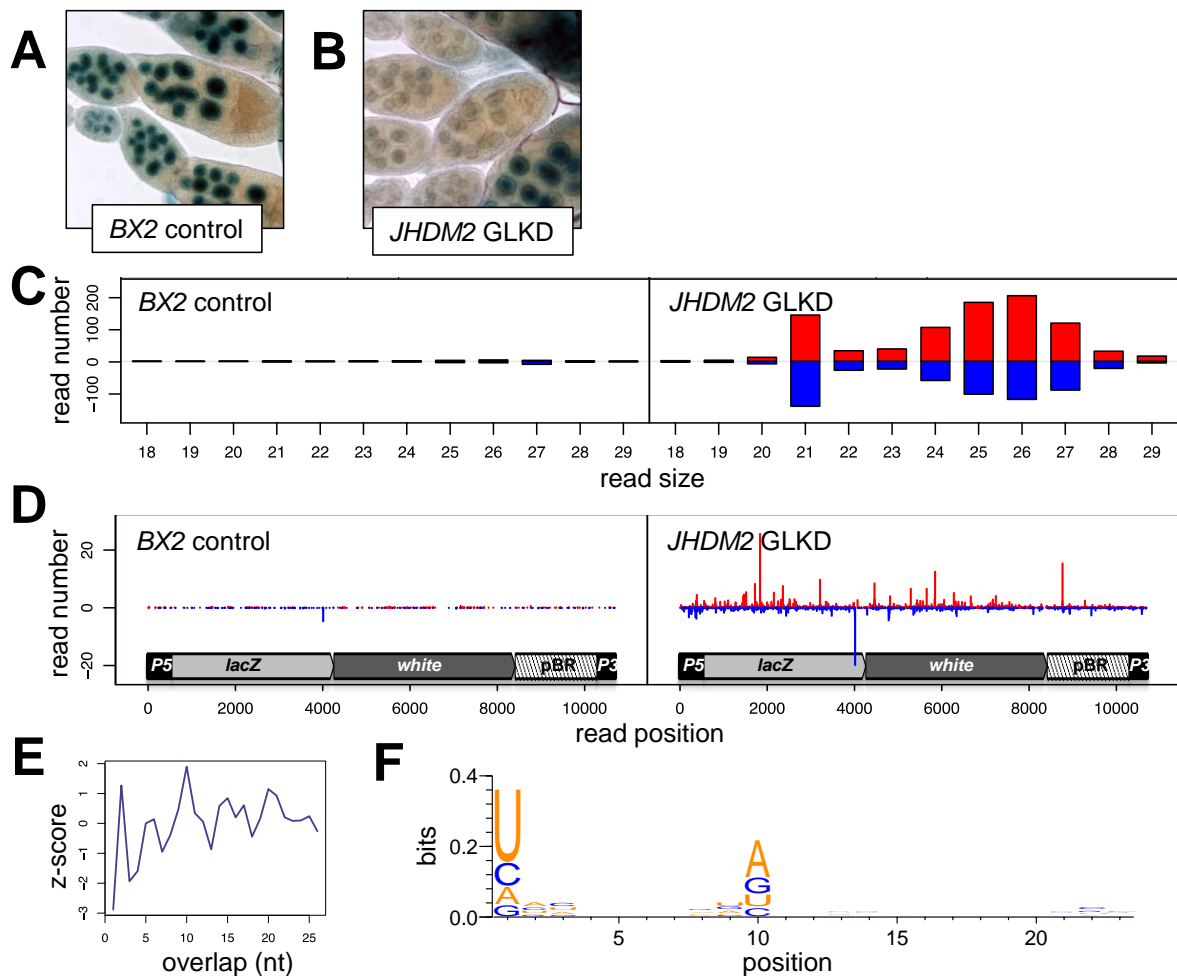
Pour préciser par quels mécanismes moléculaires la perte de fonction de *hfp* entraîne la conversion de *BX2<sup>OFF</sup>*, nous avons réalisé un RNA-seq dans les ovaires pour comparer l'expression des gènes entre le contexte *hfp GLKD* et le contrôle (*BX2 control*). A partir de la liste des gènes dérégulés, nous avons établi une liste de gènes candidats parmi les gènes moins exprimés en contexte *hfp GLKD* par rapport au contrôle (fold change > 2, p-value ajustée < 0.05 et baseMean > 400) que nous avons testé grâce au criblage génétique 'perte de fonction' décrit précédemment. Parmi tous les gènes que nous avons pu tester (dont les lignées RNAi sont disponibles ; liste complète des gènes et shRNA testés dans l'Annexe 5), un seul gène permet la conversion de *BX2* en contexte *GLKD*. Il s'agit du gène *JmjC domain-containing histone demethylase 2 (JHDM2)* aussi connu sous le nom de *KDM3B* chez les mammifères.

## 2. Perte de fonction de *JHDM2* et activation de *BX2*

### 2.1. Identification et régulation de *JHDM2*

*JHDM2* code une déméthylase d'histone contenant un domaine Jumonji-C spécifique de la déméthylation des résidus H3K9. Il a été montré *in vitro* et *in vivo*, dans des cellules HeLa mais aussi chez la drosophile que *JHDM2* participe à la déméthylation des marques H3K9me1 et H3K9me2 (Herz *et al.*, 2014; Yamane *et al.*, 2006). De la même manière que la

perte de fonction de *hfp*, la perte de fonction de *JHDM2* dans les ovaires entraîne une répression du transgène cible *P(TARGET)*, indiquant le changement d'état épigénétique de *BX2* (Figure 31A-B). La conversion de *BX2* a été vérifiée avec des vrais mutants perte de fonction de *JHDM2* (Shalaby *et al*, 2017) et indique qu'il n'y a pas d'effet 'off-target'. Les données de small RNA-seq confirme la production de piRNA, primaires et secondaires, par le cluster *BX2* en contexte *JHDM2 GLKD* (Figure 31C-F).

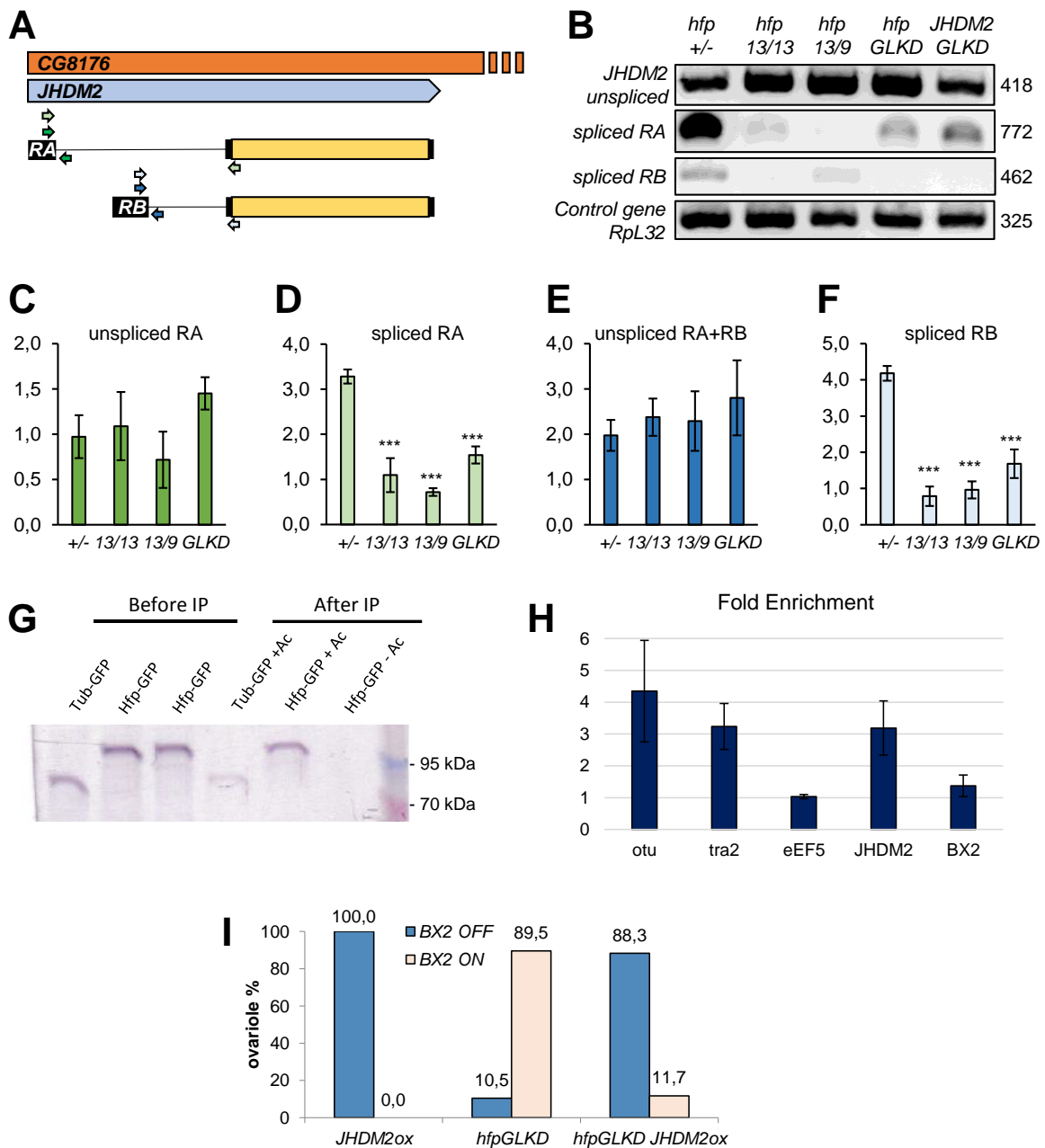


**Figure 31:** La perte de fonction de *JHDM2* entraîne l'activation de *BX2<sup>OFF</sup>* pour la production de piRNA (A) Dans le contrôle (*BX2 control*), il n'y a pas de répression du transgène *P(TARGET)* dans les ovaires, le cluster *BX2* est à l'état inactif (0% de répression, n=587 ovarioles). (B) En contexte *JHDM2 GLKD*, le transgène *P(TARGET)* présente de la répression, indiquant que le cluster *BX2* est activé dans les cellules germinales (~91% d'ovarioles présentant de la répression, n=595 ovarioles). (C, D) Les expériences de small RNA-seq indiquent que des piRNA (23 à 29 nt) sont produits par le cluster *BX2* en contexte *JHDM2 GLKD* alors qu'ils ne sont pas produits dans le contrôle. En rouge sont représentés les petits ARN sens et en bleu les petits ARN antisens (C) Répartitions par taille des petits ARN entre 18 et 29 nucléotides. Les ARN de 21 nucléotides correspondent aux siRNA tandis

que les ARN de 23 à 29 nucléotides correspondent aux piRNA. (D) L'alignement des petits ARN de 23 à 29 nucléotides (unique mappers) montre une répartition uniforme des piRNA produits le long du transgène *P[lacW]*. (E) La fréquence relative (z-score) de chevauchement entre les petits ARN sens et antisens de 23 à 29 nucléotides en contexte *JHDM2 GLKD* montre un enrichissement de l'appariement en position 10, caractéristique de la signature Ping-Pong. (F) En contexte *JHDM2 GLKD*, on observe un enrichissement de U en position 1 et de A en position 10, indiquant que des piRNA primaires et secondaires sont produits.

*JHDM2* est localisé sur le chromosome 3R et inséré dans un autre gène, *CG8176*. De plus, *JHDM2* possède trois exons et deux promoteurs alternatifs permettant la transcription de deux isoformes transcriptionnelles différentes : *JHDM2-RA* et *JHDM2-RB* (Figure 32A). Les isoformes *JHDM2-RA* et *JHDM2-RB* sont transcrites à partir de promoteurs situés respectivement dans le premier et le deuxième exon. La séquence codante, située dans le troisième exon, est commune aux deux isoformes entraînant la formation d'une seule protéine. L'insertion de *JHDM2* dans un autre gène et l'existence de promoteurs alternatifs suggèrent une régulation par épissage alternatif. Des études ont montré que *hfp* est impliqué dans la régulation post-transcriptionnelle des gènes *ovarian tumor (otu)* (Van Buskirk & Schüpbach, 2002) et *transformer 2 (tra2)* (Wang *et al*, 2013). Grâce à des expériences de RT-PCR et de RT-qPCR, nous avons analysé la production des transcrits épissés et non épissés de *JHDM2* chez différents mutants *hfp* ou en contexte *hfp GLKD* (Figure 32B-F). Ces analyses indiquent que les quantités des transcrits non épissés de *JHDM2*, aussi bien pour les isoformes RA que RB, ne sont pas affectées chez les mutants *hfp* ou en contexte *hfp GLKD* (Figure 32C, E). En revanche, les quantités des transcrits épissés de *JHDM2*, pour les isoformes RA et RB, sont significativement plus faibles chez les mutants *hfp* et en contexte *hfp GLKD* comparé au contrôle (Figure 32D, F). Ces résultats indiquent un contrôle de l'épissage de *JHDM2* par *hfp*. Pour déterminer si cette régulation s'effectue de manière directe ou indirecte, nous avons réalisé des expériences d'immunoprécipitation d'ARN (RNA-IP) afin d'identifier les ARN associés à une protéine de fusion Hfp-GFP. Nous avons confirmé par western blot que l'utilisation d'un anticorps anti-GFP permet l'immunoprécipitation de la protéine de fusion Hfp-GFP (Figure 32G). Comme attendu, les transcrits *otu* et de *tra2* sont enrichis indiquant qu'ils sont tous deux associés à la protéine Hfp, tandis que les transcrits du gène *eEF5* ne sont pas enrichis. On peut constater que les transcrits de *JHDM2* sont également enrichis alors que les transcrits de *BX2* ne le sont pas (Figure 32H). Ces résultats indiquent que *hfp* régule l'épissage de *JHDM2* en interagissant physiquement avec ses transcrits. En revanche, les expériences de RNA-IP indiquent que Hfp n'interagit pas physiquement avec les transcrits de

*BX2*, suggérant un mécanisme indirect pour l'activation de *BX2*. En effet, en contexte *hfp* *GLKD* la surexpression d'un cDNA de *JHDM2*, dont la régulation est indépendante de *hfp*, ne permet pas l'activation de *BX2*. Ce sauvetage fonctionnel indique que l'activation de *BX2* en contexte *hfp* *GLKD* dépend de l'activité – disons plutôt de la perte d'activité – de *JHDM2* (Figure 32I). Ainsi, lorsque *hfp* est muté, les transcrits de *JHDM2* (RA et RB) ne sont plus épissés et la protéine n'est pas produite. Notons toutefois qu'il n'existe pas d'anticorps anti-*JHDM2* qui permettraient de confirmer par western blot que la protéine *JHDM2* n'est plus produite en contexte mutant ou *GLKD* pour *hfp*.



**Figure 32:** La conversion de *BX2*<sup>OFF</sup> dépend de l'épissage de *JHDM2* par *hfp*. (A) Les deux isoformes transcriptionnelles de *JHDM2* (RA et RB) partagent leur 5'UTR avec le gène

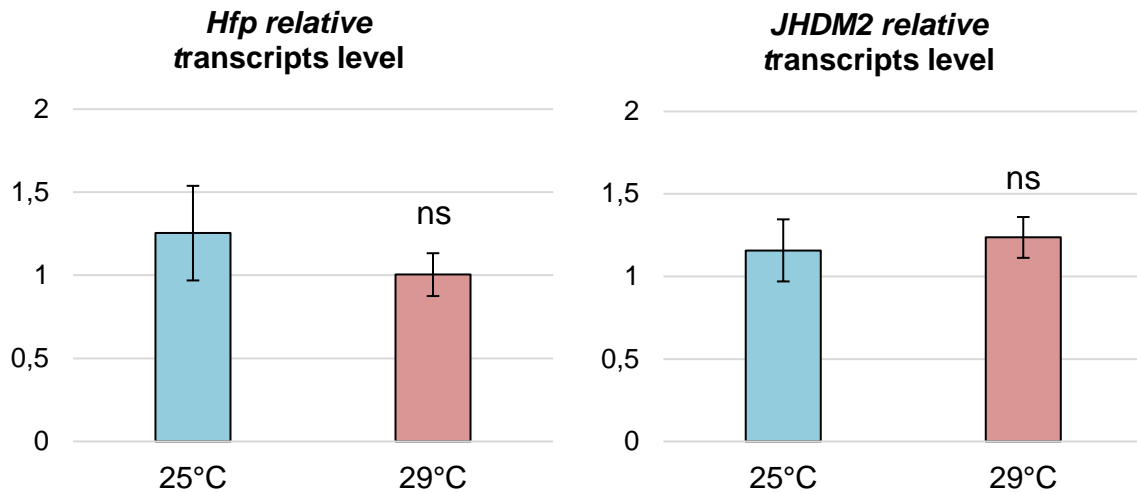
CG8176 suggérant une régulation par épissage alternatif. Les flèches indiquent la position des amorces utilisées pour les expériences de RTqPCR présentées de C à F, les couleurs représentées correspondent à celles utilisées pour les histogrammes. (B) Des expériences de RT-PCR indiquent que les formes non épissées de *JHDM2* (*JHDM2* unspliced) ne sont pas affectées alors que les formes épissées (spliced RA et spliced RB) sont affectées dans différents mutants *hfp* ou en contexte *hfp GLKD*. Les transcrits du gène contrôle *RpL32* ne sont pas affectés dans les mutants *hfp* et *JHDM2*. Des mutants hétéroalléliques *hfp<sup>13</sup>/hfp<sup>9</sup>* et homoalléliques *hfp<sup>13</sup>/hfp<sup>13</sup>* (Plus de détails dans le matériel et méthodes) et *hfp GLKD* sont testés. Les mutants *hfp<sup>+/-</sup>* et *JHDM2 GLKD* servent de contrôles négatif et positif, respectivement. Les tailles des fragments attendus après migration dans un gel d'agarose sont indiquées à droite. La localisation des amorces utilisées est indiquée en A. Notons qu'on ne trouve pas de couple d'amorces permettant d'amplifier les transcrits *JHDM2-RB* non épissés seuls. (G) Western blot sur des échantillons protéiques avant et après immunoprécipitation (IP) de deux protéines de fusion ; Hfp-GFP et Tubulin-GFP (Tub-GFP). Tub-GFP sert de contrôle négatif, permet d'estimer le bruit de fond et ainsi d'estimer l'enrichissement pour les expériences d'immunoprécipitation d'ARN (RNA-IP) représentées en H. Avant IP et après IP avec un anticorps anti-GFP, les protéines de fusion sont détectées aux tailles attendues, 77 kDa pour Tub-GFP et environ 95 kDa pour Hfp-GFP. Dans la dernière piste, nous confirmons qu'il n'y a pas de signal en l'absence d'anticorps. (H) L'histogramme représente les enrichissements (Fold Enrichment) de différents transcrits, *otu* et *tra2* sont utilisés comme contrôles positifs car connus comme cibles de *hfp*, tandis qu'*eEF5* sert de contrôle négatif. Pour plus de clarté les p-values calculées (test de student) ne sont pas représentées sur l'histogramme. Les expériences de RNA-IP indiquent que les transcrits de *JHDM2* sont enrichis de la même manière que les transcrits d'*otu* (p-val=0,2472) et de *tra2* (p-val=0,929) par rapport à *eEF5* (p-val=0,0023). On peut également constater les transcrits de *BX2* (testés grâce à des amorces reconnaissant *lacZ*) ne sont pas enrichis : l'enrichissement n'est pas significativement différents de celui d'*eEF5* (p-val=0,098), mais différents d'*otu* (p-val=0,011) et de *tra2* (p-val=0.003). Ces résultats suggèrent un contrôle direct de l'épissage de *JHDM2* par *hfp*. (I) Histogramme montrant la proportion (en %) d'ovarioles ne présentant pas de répression (*BX2 OFF*) ou présentant une répression partielle (quelques chambres ovariennes) ou complète (toutes les chambres ovariennes) (*BX2 ON*). La surexpression de *JHDM2* (*JHDM2ox*) n'affecte pas l'état épigénétique de *BX2* qui reste inactif (n=406) tandis que la perte de fonction d'*hfp* (*hfp GLKD*) entraîne majoritairement l'activation de *BX2* (n=478). La surexpression de *JHDM2* en contexte *hfp GLKD* (*hfp GLKD, JHDM2ox*) bloque majoritairement l'activation de *BX2* (n=393) indiquant que la fonction de *hfp* intervient en amont et dépend de *JHDM2*.

## 2.2. *hfp* et *JHDM2* dans l'activation de *BX2*

Si les pertes de fonction de *hfp* et de *JHDM2* entraînent l'activation de *BX2*, alors ces deux gènes sont impliqués dans le maintien de *BX2* à l'état inactif. Autrement dit, l'état non producteur de piRNA de *BX2<sup>OFF</sup>* n'est pas un état par défaut puisqu'il semble activement maintenu dans cet état par des mécanismes moléculaires spécifiques. La compréhension de ces mécanismes nécessite de préciser les processus mis en jeu dans la conversion de *BX2* en contexte mutant ou *GLKD* pour *hfp* et *JHDM2*.



Pour vérifier que les mécanismes moléculaires impliqués ne sont pas les mêmes que ceux responsables de la conversion de  $BX2^{OFF}$  à 29°C, nous avons mesuré la quantité de transcrits de *hfp* et *JHDM2* à 25°C et à 29°C par RT-qPCR (Figure 33). Les résultats indiquent que *hfp* et *JHDM2* ne sont pas affectés à 29°C comparé à 25°C, indiquant que les mécanismes moléculaires de conversion de  $BX2^{OFF}$  ne sont pas les mêmes à 29°C que dans le cas d'une perte de fonction de *hfp* ou de *JHDM2*.



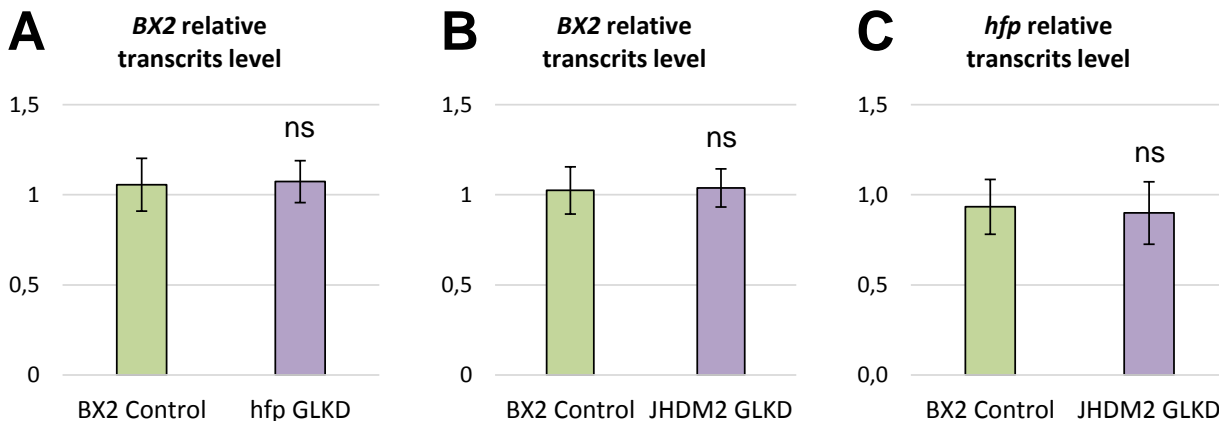
**Figure 33 :** Quantité de transcrits relative de *hfp* et *JHDM2* à 25°C et à 29°C. Les expériences RT-qPCR sur ARN extraits d'ovaires montrent qu'à 29°C, il n'y a pas de différence de quantité de transcrits par rapport à 25°C pour *hfp* (à gauche) et *JHDM2* (à droite). ns pval > 0.05 (Test de student bilatéral).

Les pertes de fonction de *hfp* et de *JHDM2* entraînent une conversion du cluster  $BX2^{OFF}$  dans la majorité des ovarioles, respectivement 87% et 91% des ovarioles présentant de la répression. Parmi celles-ci, on identifie des ovarioles présentant une répression complète (toutes les chambres ovariennes de l'ovariole présentent de la répression) et des ovarioles présentant une répression partielle (seule une partie des chambres ovariennes de l'ovariole présentent de la répression). Or, les proportions d'ovarioles présentant de la répression complète et partielle ne sont pas les mêmes entre *hfp GLKD* et *JHDM2 GLKD* (Table 6). En effet, en contexte *hfp GLKD* on observe plus d'ovarioles présentant de la répression complète (53.1%) que d'ovarioles présentant de la répression partielle (34.6%). A l'inverse, en contexte *JHDM2 GLKD* on observe plus d'ovarioles présentant de la répression partielle (57%) que d'ovarioles présentant de la répression complète (34.1%). Ces résultats indiquent qu'il y a une différence entre *hfp GLKD* et *JHDM2 GLKD* dans la conversion de  $BX2^{OFF}$ .

	Ovarioles 'ON'	Ovarioles 'VAR'	Ovarioles 'OFF'	Total Ovarioles
<i>hfp GLKD</i>	53,1% (n=517)	34,5% (n=336)	12,4% (n=121)	100% (n=974)
<i>JHDM2 GLKD</i>	34,1% (n=203)	57% (n=339)	8,9% (n=53)	100% (n=595)

**Table 6 :** Répartition de la répression dans les ovaires dans les contextes *hfp GLKD* et *JHDM2 GLKD*. Les pourcentages d'ovarioles présentant de la répression complète (ON), partielle (VAR, pour variégué) ou pas de répression (OFF) ne sont pas les mêmes entre les contextes *hfp GLKD* et *JHDM2 GLKD*. Les nombres d'ovarioles comptés sont indiqués entre parenthèses.

De la même manière qu'il n'y a pas de différence de quantité de transcrits entre  $BX2^{OFF}$  et  $BX2^{ON}$ , il n'y a pas non plus de différence de quantité de transcrits provenant du cluster *BX2* chez les individus *hfp GLKD* et *JHDM2 GLKD* par rapport aux contrôles (Figure 34A-B). Par ailleurs, *JHDM2 GLKD* n'affecte pas non plus la quantité de transcrits de *hfp* (Figure 34C), suggérant qu'il n'y a pas de rétrocontrôle de *JHDM2* sur la quantité de transcrits d'*hfp*. En conséquence, la différence de distribution de la répression observée dans les deux contextes, *hfp GLKD* et *JHDM2 GLKD*, ne s'explique ni par une différence dans la quantité de transcrits de *BX2*, ni par un rétrocontrôle de *JHDM2* sur *hfp*.



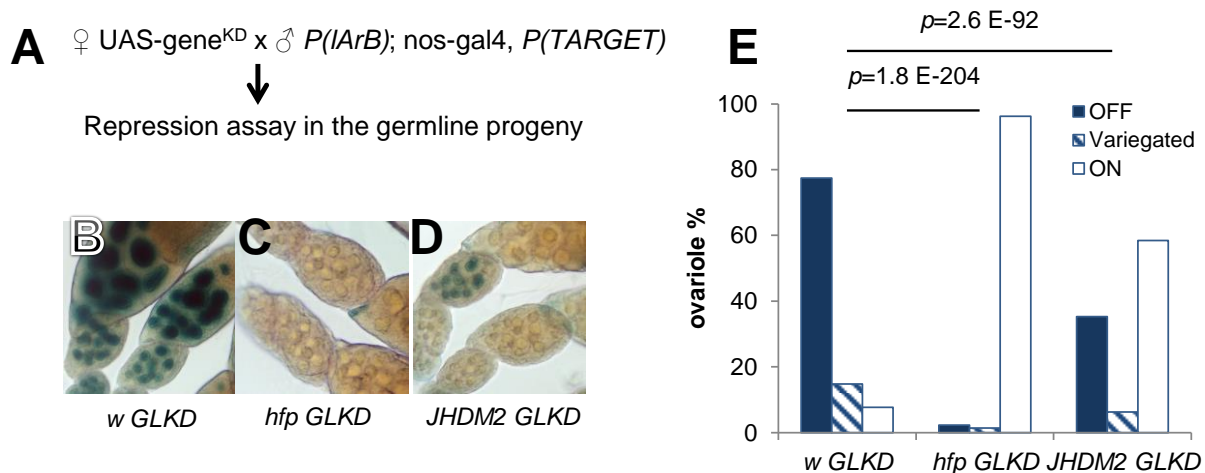
**Figure 34 :** Quantité relative de transcrits *BX2* et *hfp* dans les contextes *hfp GLKD* et *JHDM2 GLKD*. Les expériences RT-qPCR sur ARN extraits d'ovaires montrent que (A) En contexte *hfp GLKD*, il n'y a pas de différence de quantité de transcrits de *BX2* par rapport au contrôle (*BX2 control*) ; (B-C) En contexte *JHDM2 GLKD*, il n'y a pas non plus de différence de quantité de transcrits de *BX2* (B) et d'*hfp* (C) par rapport au contrôle (*BX2 control*). ns p-val > 0.05 (Test de student bilatéral).

Une hypothèse permettant d'expliquer cette différence de distribution est que *hfp*, en plus de la régulation de *JHDM2*, régule de nombreux autres gènes et doit probablement réguler indirectement une partie des cibles de *JHDM2*. Par ailleurs, la dérégulation des cibles de *hfp* pourrait avoir des effets synergiques ou antagonistes à l'activité de *JHDM2*. Une analyse comparative des données de RNA-seq issues des *hfp GLKD* et *JHDM2 GLKD* permettraient peut-être d'identifier de tels facteurs s'ils existent.

### 3. Etude de la perte de fonction de *JHDM2*

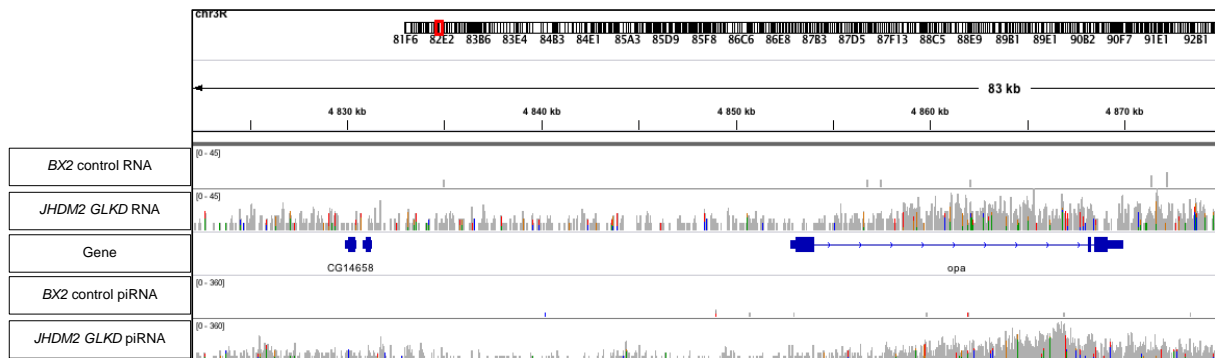
#### 3.1. *JHDM2 GLKD* et effets pan-génomiques

*hfp GLKD* et *JHDM2 GLKD* entraînent l'activation de la production de piRNA par *BX2<sup>OFF</sup>* dans la majorité des ovarioles. Pour déterminer si ces contextes de diminution d'expression peuvent, comme pour *BX2<sup>OFF</sup>*, activer la production de piRNA au niveau d'autres séquences spécifiques, nous avons analysé l'effet des contextes *hfp GLKD* et *JHDM2 GLKD* sur l'état épigénétique du transgène *P(lArB)*. *P(lArB)* est un transgène contenant notamment le gène *lacZ* inséré dans un piRNA cluster (cluster *IA*) au niveau du subtélomère du chromosome X. Lorsque le transgène *P(lArB)* est hérité maternellement, l'expression du transgène cible *P(TARGET)* est réprimée dans la majorité des chambres ovariennes à la première génération ; en revanche lorsque le transgène *P(lArB)* est hérité paternellement, le transgène cible *P(lArB)* est exprimé dans la majorité des chambres ovariennes à la première génération. Aux générations suivantes, le nombre de chambres ovariennes présentant de la répression reste constant lorsque le transgène *P(lArB)* est hérité maternellement en G1 mais il augmente lorsque *P(lArB)* est hérité paternellement en G1 (Josse *et al*, 2007). Lorsque *P(lArB)* est hérité paternellement, on observe que le pourcentage d'ovarioles présentant de la répression (totale ou partielle) est significativement plus élevé dans les contextes *hfp GLKD* et *JHDM2 GLKD* comparé au contrôle *white GLKD* (Figure 35). Ces résultats indiquent que *hfp GLKD* et *JHDM2 GLKD* n'affectent pas spécifiquement l'état épigénétique de *BX2* mais qu'ils peuvent aussi affecter d'autres régions du génome.



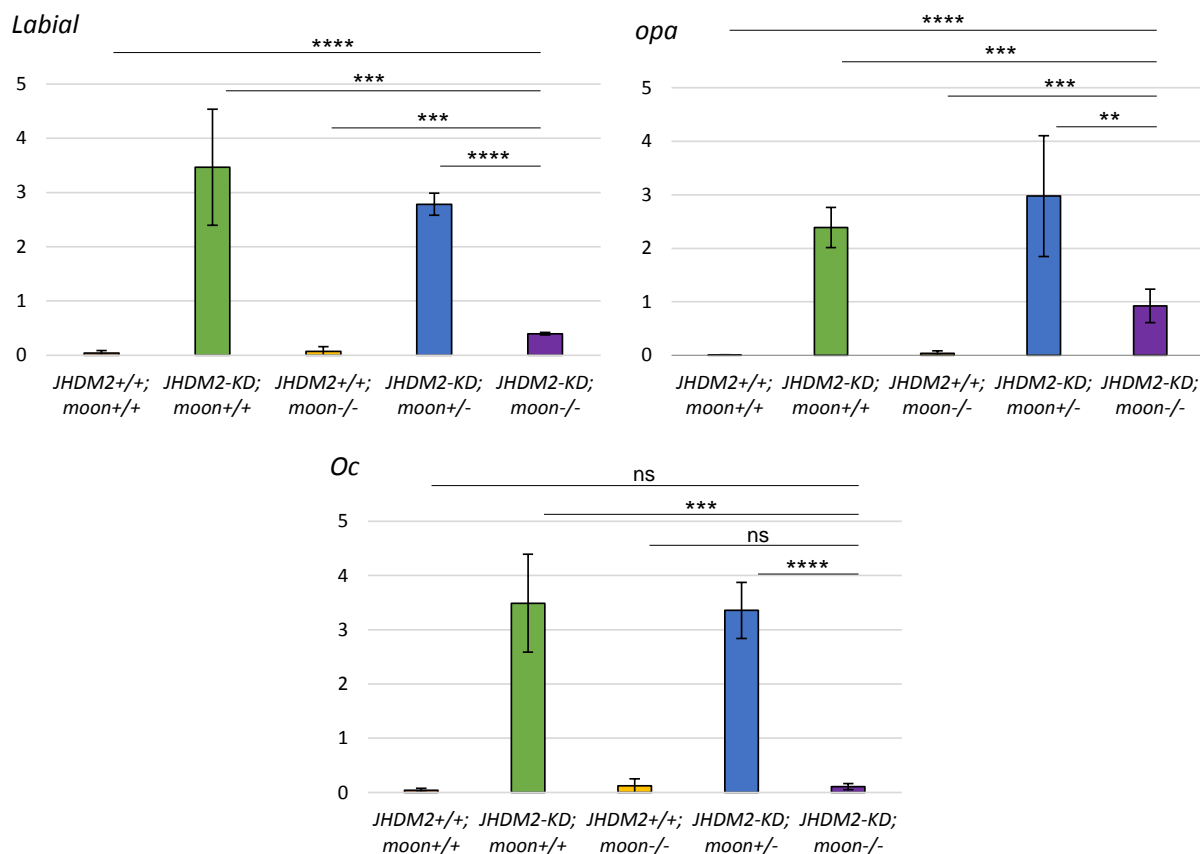
**Figure 35** : Effet des contextes *hfp GLKD* et *JHDM2 GLKD* sur l'état épigénétique de *P(lArB)*. (A) Schéma de croisement pour analyser l'effet des contextes *hfp GLKD* et *JHDM2 GLKD* sur l'état épigénétique de *P(lArB)*. Des femelles portant le transgène *UAS-gene<sup>KD</sup>*, dirigé contre *white* (contrôle), *hfp* ou *JHDM2*, sont croisées par des mâles portant les transgènes *P(lArB)*, *nos-gal4*, permettant l'expression du transgène *UAS-gene<sup>KD</sup>* dans les ovaires, et un transgène cible *P(TARGET)* ; et les descendants analysés par coloration  $\beta$ -galactosidase. Les résultats sont présentés en B, C et D. (B) Le contrôle *white GLKD* confirme que le transgène *P(lArB)*, lorsqu'il est hérité paternellement, est exprimé dans la majorité des chambres ovariennes. (C-D) Dans les contextes *hfp GLKD* (C) et *JHDM2 GLKD* (D), on observe que la majorité des chambres ovariennes présentent de la répression du transgène *P(TARGET)*. (E) Comptage du nombre d'ovarioles présentant une répression complète (ON), une répression partielle (VAR) ou pas de répression (OFF) du transgène *P(TARGET)* chez les *white GLKD* (contrôle, à gauche), *hfp GLKD* (au milieu) et *JHDM2 GLKD* (à droite). On observe de la répression (ON+VAR) en contexte *hfp GLKD* et *JHDM2 GLKD* tandis qu'il y en a peu chez le contrôle *white GLKD*. Les p-values indiquées sont calculées avec un test  $\chi^2$  avec 3 degrés de liberté.

Si *hfp GLKD* et *JHDM2 GLKD* peuvent moduler la production de piRNA pour le cluster *BX2<sup>OFF</sup>* et le transgène *P(lArB)*, alors ils pourraient également affecter la production de piRNA pour d'autres régions génomiques. Pour le déterminer, nous avons analysé des données de small RNA-seq issues des deux contextes de diminution d'expression. L'analyse des données de small RNA-seq issues de *hfp GLKD* n'a pas mis en évidence de différences importantes dans la production de piRNA au niveau du génome entier. En revanche, l'analyse des données de small RNA-seq pour *JHDM2 GLKD* a révélé l'existence de régions génomiques particulières. En effet, ces régions produisent des piRNA en contexte *JHDM2 GLKD* alors qu'elles n'en produisent pas chez le contrôle (*BX2 control*). L'analyse des données de RNA-seq indique également que ces régions sont transcrites chez *JHDM2 GLKD* alors qu'elles ne le sont pas chez le contrôle (*BX2 control*) (Figure 36, Annexes 6 et 7).



**Figure 36 :** Identification de régions génomiques pour lesquelles le contexte *JHDM2 GLKD* induit une transcription et une production de piRNA aberrantes. Les analyses de RNA-seq et de small RNA-seq ont permis d'identifier huit régions génomiques particulières: ces régions ne sont pas (ou peu) transcrites et ne produisent pas (ou peu) de piRNA dans le contrôle (*BX2* control RNA et *BX2* control piRNA) mais elles sont transcrites et produisent des piRNA en contexte *JHDM2 GLKD* (*JHDM2 GLKD* RNA et *JHDM2 GLKD* piRNA). Pour des raisons de lisibilité, seule une région (en partie tronquée) a été représentée ici comme exemple. Les huit régions identifiées sont représentées dans l'Annexe 6 et des informations complémentaires sur les régions sont apportées dans l'Annexe 7.

Ces huit différentes régions sont localisées sur les différents chromosomes, mesurent de 30 à 130 kb et contiennent beaucoup de gènes codant des facteurs de transcription impliqués dans le développement qui ne sont pas (ou sont très peu) exprimés dans les ovaires (Annexe 7). Nous avons généré des mutants *moon-JHDM2 GLKD* pour déterminer si la transcription de ces régions dépend de la machinerie de transcription des piRNA clusters double-brin (Andersen *et al*, 2017). Les résultats de RT-qPCR confirment que les trois gènes testés, *labial*, *opa* et *Ocelliless (Oc)*, ne sont pas exprimés dans les ovaires en contexte sauvage pour *JHDM2* mais sont exprimés fortement en contexte *JHDM2 GLKD*. On peut également constater que *labial*, *opa* et *Oc* ne s'expriment pas dans le mutant homozygote *Moon* seul. En contexte *JHDM2 GLKD ; Moon<sup>+/-</sup>*, les trois gènes s'expriment autant qu'en contexte *JHDM2 GLKD*. En revanche, en contexte *JHDM2 GLKD ; Moon<sup>-/-</sup>*, les trois gènes sont beaucoup moins exprimés comparé à *JHDM2 GLKD* (Figure 37). Ces données suggèrent fortement que *Moon* est requis pour la transcription d'au moins trois des huit régions identifiées. Ces données suggèrent également que ces régions deviennent des piRNA clusters chez *JHDM2 GLKD*.



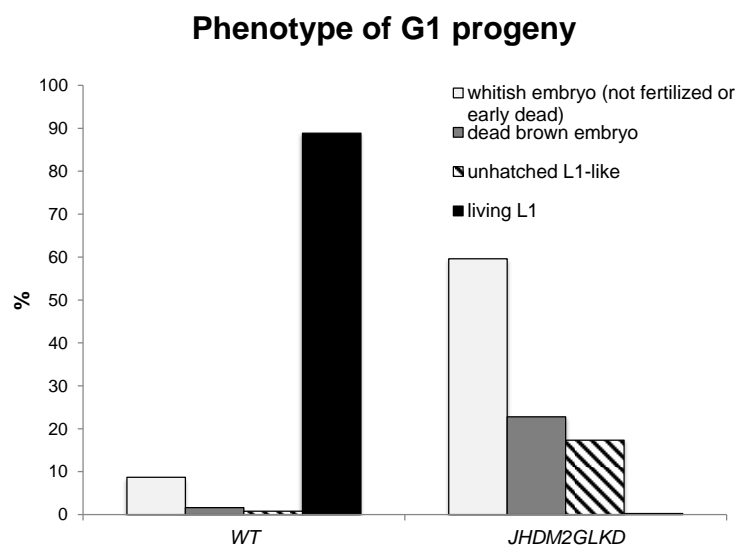
**Figure 37** : *moon* est impliqué dans la transcription de *labial*, *opa* et *Oc*, localisés dans trois des huit régions identifiées. Les RTqPCR montrent que les ovaires, en contexte sauvage pour *JHDM2* et *moon* (en rouge) de même qu'en contexte mutant *JHDM2*<sup>+/+</sup> ; *moon*<sup>-/-</sup> (en jaune), *labial* (en haut à gauche), *opa* (en haut à droite) et *Oc* (en bas) ne sont pas exprimés. Lorsque *JHDM2* est inactivé (en vert), *labial*, *opa* et *Oc* sont fortement exprimés. En contexte *JHDM2* *GLKD* ; *moon*<sup>+/-</sup> (en bleu), *labial*, *opa* et *Oc* sont toujours fortement exprimés alors qu'en contexte *JHDM2* *GLKD* ; *moon*<sup>-/-</sup> (en violet), *labial*, *opa* et *Oc* sont peu (voire pas) exprimés. ns p-val > 0.05, \*\* p-val < 0.01 ; \*\*\* p-val < 0.001 ; \*\*\*\* p-val < 0.0001 (Test de student bilatéral). Pour plus de clarté, seules les p-values des tests permettant de comparer le contexte *JHDM2* *GLKD* ; *Moon*<sup>-/-</sup> aux autres génotypes ont été représentées. Pour information : il n'y a pas de différences significatives entre les contextes *JHDM2*<sup>+/+</sup> ; *moon*<sup>+/+</sup> et *JHDM2*<sup>+/+</sup> ; *moon*<sup>-/-</sup> (p-val<sup>labial</sup> = 0,488, p-val<sup>opa</sup> = 0,081 et p-val<sup>Oc</sup> = 0,186) montrant que le mutant *moon* seul n'induit pas la transcription de ces gènes; il n'y a pas non plus de différences significatives entre les contextes *JHDM2* *GLKD* ; *moon*<sup>+/+</sup> et *JHDM2* *GLKD* ; *moon*<sup>+/-</sup> (p-val<sup>labial</sup> = 0,366 , p-val<sup>opa</sup> = 0,306 et p-val<sup>Oc</sup> = 0,828), montrant qu'il n'y a pas d'effet dominant du mutant *moon*.

De même qu'il y a des régions qui produisent plus de piRNA, il y a aussi des régions qui produisent moins de piRNA en contexte *JHDM2* *GLKD* par rapport au contrôle (*BX2* control). En effet, l'analyse des données de small RNA-seq a également permis d'identifier 155 régions qui produisent moins de piRNA en contexte *JHDM2* *GLKD*. Parmi ces régions on retrouve en partie les piRNA clusters définis par Julius Brennecke et ses collaborateurs

(Brennecke *et al*, 2007). L'ensemble des résultats suggèrent que la production de piRNA au niveau du génome entier est réorganisée en contexte *JHDM2 GLKD*. Ainsi, *JHDM2* jouerait un rôle pour maintenir certaines régions génomiques dans un état non-producteur de piRNA et, indirectement peut-être, délimiter les frontières des piRNA clusters. Des expériences de ChIP-seq sur les marques H3K9me2, H3K9me3 et la protéine Rhino sont en cours, en collaboration avec le laboratoire de Chantal Vaury et Emilie Brasset (GReD, Clermont-Ferrand). Ces expériences permettront de déterminer si *JHDM2 GLKD* induit une réorganisation de la chromatine au niveau du génome entier et, au quel cas, d'identifier les régions affectées.

### 3.2. La perte de fonction de *JHDM2* affecte-t-elle la génération suivante ?

Les mutants *JHDM2*, de même que les individus *JHDM2 GLKD*, sont stériles : les femelles pondent des œufs mais ceux-ci ne dépassent jamais le stade L1 (Figure 38). Il apparaît que tous les descendants des individus *JHDM2 GLKD* sont stériles, incluant ceux qui n'héritent pas du transgène *UAS-gene<sup>KD</sup>* capable d'induire la perte de fonction de *JHDM2*. Cet effet maternel suggère un mécanisme épigénétique responsable de la stérilité des descendants des individus *JHDM2 GLKD*.

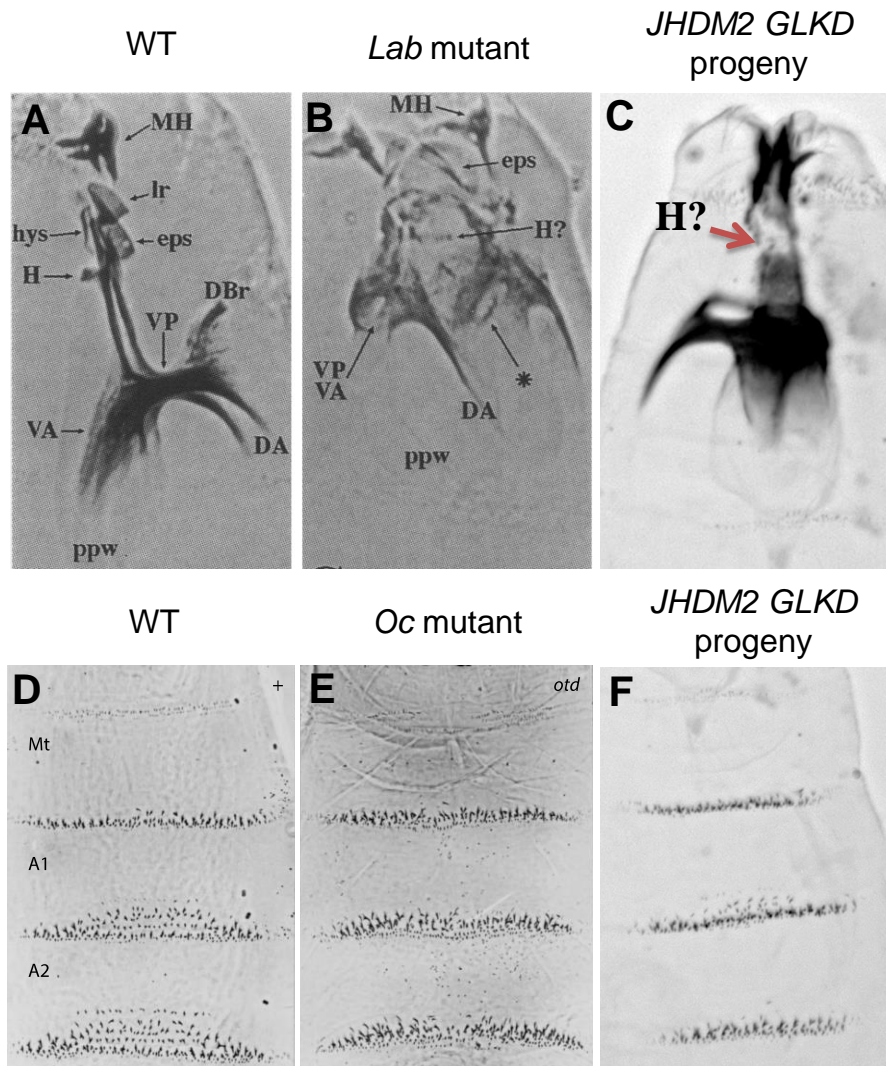


**Figure 38** : Les descendants des *JHDM2 GLKD* ne dépassent pas le stade L1. En contexte sauvage (à gauche, n=368) près de 90% des descendants dépassent le stade L1 (en noir) tandis qu'un peu plus de 10% des individus ne survivent pas. Chez les descendants des *JHDM2*

*GLKD* (à droite, n=369), les individus ne dépassent pas le stade L1. On observe que parmi les embryons, certains sont vides/ne sont pas fécondés (près de 60%, en gris clair), d'autres sont morts à des stades précoces (un peu plus de 20%, en gris foncé) et d'autres encore arrivent jusqu'au stade L1 mais n'éclosent pas, appelé 'L1-like' (un peu moins de 20%, en hachuré). La répartition des phénotypes entre les deux conditions est statistiquement différente p-val = 5,422<sup>-127</sup> (Test  $\chi^2$  d'homogénéité avec 3 degrés de liberté).

Parmi les descendants des *JHDM2 GLKD* nous avons identifié des embryons de différentes natures : une majorité des œufs est totalement vide, une autre partie correspond à des embryons morts et une minorité d'individus, qu'on appelle 'L1-like' arrivent jusqu'en L1 mais n'éclosent pas. Nous avons cherché s'il existe des phénotypes embryonnaires associés aux descendants L1-like des *JHDM2 GLKD* – les autres types d'embryons ne pouvant pas être analysés. Les descendants L1-like présentent des phénotypes qui ressemblent à ceux que l'on observe chez des mutants perte de fonction de *labial* et *Oc* (Figure 39). En effet, au niveau de la tête des larves les descendants des *JHDM2 GLKD* ont un phénotype qui ressemble à un mutant perte de fonction *labial* où la pièce H semble absente (Figure 39A-C, Lutz et al., 1996). Au niveau de la région abdominale, les descendants des individus *JHDM2 GLKD* ont un phénotype qui ressemble à un mutant perte de fonction *Oc* où les denticules antérieurs de chaque segment sont absents (Figure 39C-F, Wieschaus and Nüsslein-Volhard, 2016). Ces observations suggèrent que l'expression des gènes *labial* et *Oc* est affectée chez les descendants des *JHDM2 GLKD*.

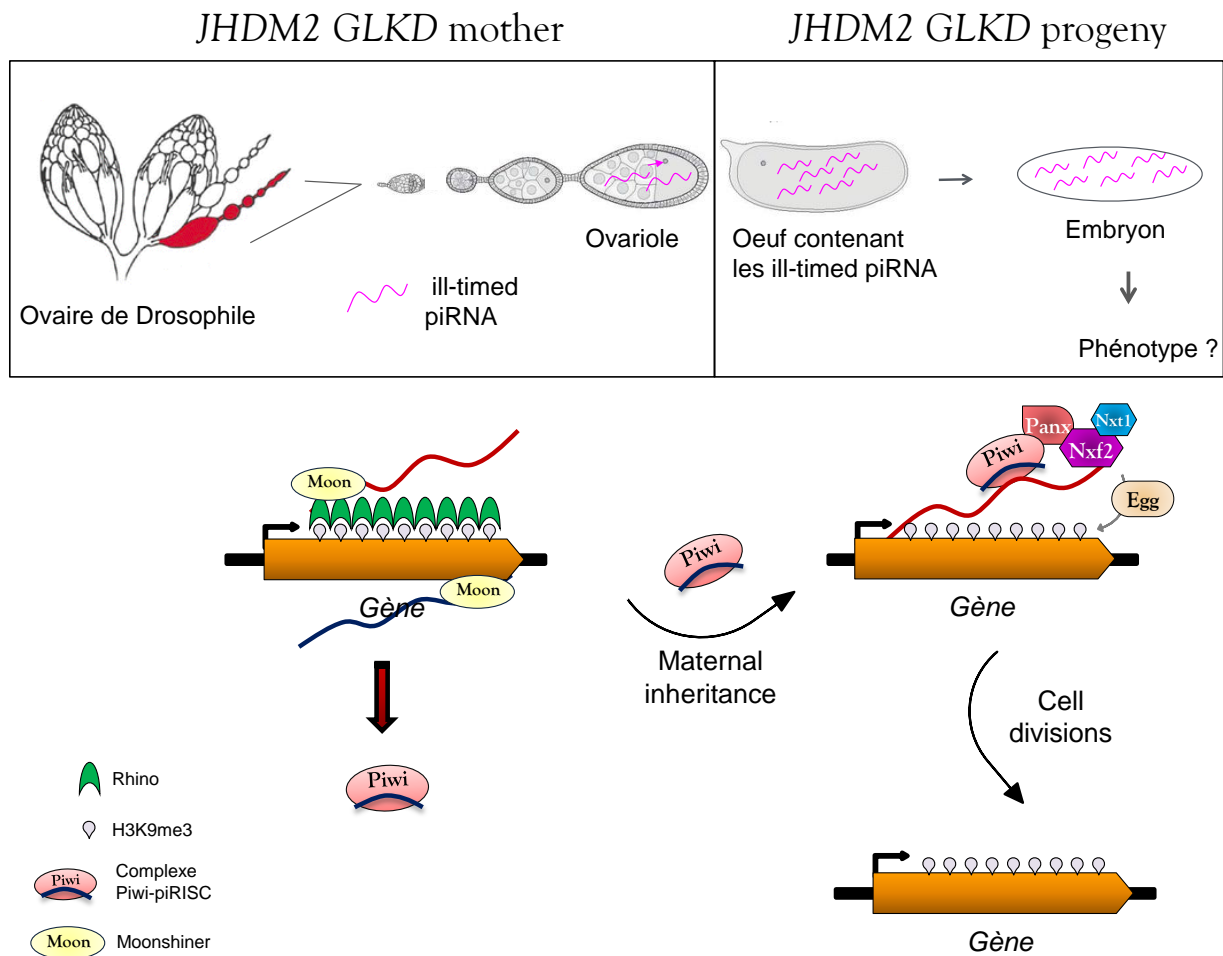




**Figure 39 : Phénotypes des mutants *labial* et *Oc* de larves de stade 1 et des descendants des *JHDM2 GLKD* chez les L1-like.** (A-C) Région antérieure (tête) de la larve L1 chez le sauvage (A), chez le mutant *labial* (B) et chez les descendants des *JHDM2 GLKD* (C). Chez le sauvage, on identifie différentes structures buccales: Mouth-hook (MH, mandibule), Labrum (lr), epistomale sclerite (eps), hypostomale sclerite (hps), Dorsal bridge (DBr), Ventral plate (VP), Ventral arm (VA), Dorsal arm (DA) et H-piece (H). Chez le mutant *labial*, on observe une disparition de la pièce H (Lutz *et al*, 1996). Les descendants des *JHDM2 GLKD* ont un phénotype qui ressemble à un mutant *labial* où la pièce H semble absente. (D-F) Région abdominale de la larve L1 chez le sauvage (D), chez le mutant *Oc* (E) et chez les descendants des *JHDM2 GLKD* (F). Chez le sauvage, on observe des denticules au niveau des segments abdominaux. Chez le mutant *Oc*, on observe une diminution de la taille voire une absence des denticules. Mt: Metathorax, A1: Segment abdominal 1, A2: Segment abdominal 2 (Wieschaus & Nüsslein-Volhard, 2016). Les descendants des *JHDM2 GLKD* ont un phénotype qui ressemble à un mutant *Oc*. Les photos A et B sont tirées de Lutz *et al*, 1996 et les photos D et E proviennent de Wieschaus & Nüsslein-Volhard, 2016.

En contexte *JHDM2 GLKD*, les gènes *labial* et *Oc* sont transcrits et produisent des piRNA dans les ovaires tandis que chez les descendants des *JHDM2 GLKD*, *labial* et *Oc* semblent moins exprimés voire inactivés. Ces données suggèrent que, en contexte *JHDM2 GLKD*, les piRNA produits anormalement par les gènes *labial* et *Oc* – et potentiellement les piRNA produits par les autres gènes présents dans les huit régions – sont hérités par l’embryon et entraînent la perte de fonction des gènes *labial* et *Oc*. Notre hypothèse est que ces piRNA ‘anormaux’, chargés par Piwi, reconnaissent les transcrits naissant des gènes *labial* et *Oc* et induisent une hétérochromatinisation locale au tout début du développement de l’embryon par le même mécanisme que celui impliqué dans la répression transcriptionnelle des ET (Figure 40).

Pour valider ce modèle, il nous faut réaliser plusieurs expériences. En contexte *JHDM2 GLKD*, des expériences de ChIP-seq en cours sur les marques H3K9me2, H3K9me3 et la protéine Rhino nous permettront de confirmer si ces régions deviennent des piRNA clusters ou non dans les ovaires. Pour montrer que ce sont bien les piRNA produits anormalement en contexte *JHDM2 GLKD* puis hérités qui sont responsables de la létalité embryonnaire et des phénotypes observés, nous devons générer des lignées mutantes ou GLKD pour *JHDM2* et pour la voie des piRNA (mutant *Piwi*, *Rhino*...) afin de sauver fonctionnellement les embryons des phénotypes (létalité, denticules ou appendices buccaux). Cependant, la plupart des mutants homozygotes pour les gènes de la voie des piRNA sont stériles (ovaires très atrophiés voire absents) et empêchent le développement précoce de l’embryon. Il s’agit donc d’identifier un mutant qui affecte suffisamment la voie des piRNA sans affecter la viabilité trop fortement et permettre la survie des embryons. Nous réaliserons également des expériences de RT-qPCR sur embryons pour déterminer si les gènes potentiellement impactés sont différenciellement exprimés. En parallèle, nous devons effectuer des expériences de ChIP-seq sur les descendants des *JHDM2 GLKD* pour déterminer s’il y a un enrichissement en marques H3K9me3 au niveau des régions identifiées.



**Figure 40 : Modèle permettant d'expliquer les phénotypes des L1-like descendants des *JHDM2 GLKD*.** En contexte *JHDM2 GLKD*, les gènes localisés dans huit régions semblent transformés en piRNA clusters (enrichissement en H3K9me3 et Rhino à confirmer par ChIP-seq), exprimés anormalement grâce à la machinerie de transcription des piRNA clusters double-brin (impliquant l'activité de *moon*) et produisent des piRNA dans les ovaires. Ces piRNA 'anormaux' chargés par Piwi sont hérités maternellement. Dans l'embryon précoce, les complexes Piwi-piRISC, qui interagissent avec les co-facteurs Panoramax (Panx), Nxf2 et Nxt, reconnaissent les transcrits naissant des gènes dérégulés à la génération précédente (*labial*, *Oc...*) et recrutent la méthyltransférase Eggless qui induit l'hétérochromatinisation locale (enrichissement des marques H3K9me3). Les marques H3K9me3 sont ensuite maintenues au cours des divisions cellulaires successives lors de l'embryogenèse. Ainsi, l'hétérochromatinisation de la région empêche l'expression des gènes qui y sont présents et engendre des phénotypes de perte de fonction pour certains des gènes.

### III. La dynamique des piRNA clusters

#### 1. Présentation de l'article II

Lorsqu'un nouvel ET apparaît dans un génome par transfert horizontal, la régulation de sa mobilité et de celle des autres copies de l'ET se met en place lorsqu'une copie s'insère dans un piRNA cluster. Des travaux montrent qu'après son insertion dans un piRNA cluster, des transgènes dérivés de l'élément *P* (mimant l'insertion d'un ET) acquièrent les propriétés moléculaires du piRNA cluster dans lequel ils sont insérés, à savoir capacités répressives, production de piRNA et enrichissement en marques H3K9me3 et Rhino (Khurana *et al*, 2011; Muerdter *et al*, 2012; Olovnikov *et al*, 2013; Radion *et al*, 2019, 2018). Cependant, la dynamique de conversion d'un ET après son insertion dans un piRNA cluster n'est pas clairement établie. Dans une étude parue en 2007, le laboratoire a montré que lorsque le transgène *P(lArB)* est hérité maternellement, ses capacités répressives sont élevées à la génération suivante et sont maintenues par la suite. En revanche, lorsque *P(lArB)* est hérité paternellement, ses capacités répressives ne sont pas totales dès la première génération et nécessitent plusieurs générations avant d'atteindre les mêmes capacités répressives que pour un héritage maternel du transgène (Josse *et al*, 2007). Ces travaux indiquent que la dynamique de conversion n'est pas la même selon qu'un transgène est hérité maternellement ou paternellement. Amna Asif-Laidin, ancienne étudiante en thèse, et Laure Teysset ont initié une étude afin de comprendre cette différence de dynamique. Par ailleurs, l'étude de la dynamique d'un transgène hérité paternellement représente également un moyen d'étudier l'insertion d'un nouvel ET, apparu par transfert horizontal, dans un piRNA cluster.

Pour cette étude en cours de finalisation, j'ai participé à la préparation d'échantillons et à la réalisation des expériences de RT-qPCR et de ChIP-qPCR (incluant aussi la mise au point du protocole). J'ai également participé à l'écriture du manuscrit (relecture, correction, écriture du matériel et méthodes concernant les expériences réalisées et figures).

**2. Article II: Modelling early germline immunization by transposable element cooption reveals internal piRNA cluster heterogeneity**

1 **Modelling early germline immunization by transposable element cooption reveals**  
2 **internal piRNA cluster heterogeneity**

3  
4  
5  
6  
7  
8  
9

Amna Asif-Laidin<sup>1#</sup>, Karine Casier<sup>1#</sup>, Antoine Boivin<sup>1</sup>, Elise Viodé<sup>1</sup>, Valérie Delmarre<sup>1</sup>, Stéphane  
Ronsseray<sup>1</sup>, Clément Carré<sup>1</sup> and Laure Teyssset<sup>\*1</sup>

10 <sup>1</sup>Sorbonne Université, CNRS, Institut de Biologie Paris-Seine, Laboratoire Biologie du  
11 Développement, UMR7622, "Transgenerational Epigenetics & small RNA Biology," F-75005, Paris,  
12 France. # These authors contributed equally

13 **Email:** laure.teyssset@sorbonne-universite.fr

14 **Running title:** Internal molecular heterogeneity of germline piRNA cluster

15  
16

17 **Key words:** transposons, piRNA cluster, Subtelomeres, Drosophila, Horizontal Transfer,  
18 transgenerational inheritance

19 **Abstract**

20 All genomes comprise a fraction of Transposable Elements (TE), which mobility needs to be  
21 carefully controlled. In gonads, TEs are repressed by PIWI-interacting RNAs (piRNAs) a class of  
22 small RNAs synthesized by heterochromatic loci called piRNA clusters enriched by fragments of  
23 TEs. Maintenance of active piRNA clusters across generations is secured by maternal inheritance  
24 of piRNAs that accounts for the memory of TE to repress. In some occasions, genomes are facing  
25 Horizontal Transfer (HT) of TEs that endanger the host genome integrity. It is assumed that in  
26 absence of specific maternal piRNAs, naïve genomes will eventually be able to produce new  
27 piRNAs, but the timing of their production remains elusive. Using a set of transgenes consistent  
28 with the size of autonomous TEs inserted in different piRNA clusters, we have been able to model  
29 such events. Using functional assays, we have found that the frequency of complete cooption of  
30 these transgenes by germline piRNA cluster depends on the size of the sequence. Our results lead  
31 us to hypothesize that fragments of TEs, instead of full length copies, residing in piRNA clusters  
32 could be a prerequisite for efficient activation of piRNA cluster alleles at each generation. We also  
33 found evidence that dependency for piRNA factors may vary within piRNA cluster revealing an  
34 unexpected level of complexity shedding light on a new magnitude of piRNA cluster flexibility  
35 fundamental for the maintenance of genome integrity.

36

## 37 **Introduction**

38 Transposable Elements (TE) are mobile endogenous genetic elements predominantly  
39 transmitted through vertical transfer like any other genes. In most metazoan germline, their  
40 transposition is repressed during gametogenesis thanks to specialized small RNAs associated  
41 with the PIWI proteins (Brennecke et al., 2007; Gunawardane et al., 2007). These small RNAs have  
42 been named PIWI-interacting RNAs or piRNAs (23-29nt). In *Drosophila melanogaster* ovaries,  
43 piRNAs are produced from about 140 piRNA heterochromatic loci, called piRNA clusters. They  
44 mainly contain TE fragments coming from ancient insertions functioning as a library of mobile  
45 copies to repress. In *Drosophila* due to piRNAs maternal inheritance, H3K9me3 deposition by Piwi  
46 has been shown to be critical in the determination of germline piRNA clusters as it recruits the  
47 HP1 homolog Rhino (Rhi) (Akkouche et al., 2017). On the one hand, Rhi interacts with Deadlock  
48 and Cutoff, associated with repression of splicing and transcription termination and, on the other  
49 hand, with a specialized RNA polymerase II initiation complex dependent on Moonshiner (Moon)  
50 (Andersen et al., 2017; Chen et al., 2016; Le Thomas et al., 2014; Mohn et al., 2014; Rangan et al.,  
51 2011; Rozhkov et al., 2013; Zhang et al., 2014). The major *42AB* cluster, the most prominent dual-  
52 strand piRNA cluster in *Drosophila*, has been intensively used to identify all these factors.  
53 However, germline piRNA cluster such as the *38C1* is a Rhi dependent, Moon independent  
54 cluster, while the *20A* is a Rhi and Moon independent uni-strand piRNA cluster (Andersen et al.,  
55 2017; Klattenhoff et al., 2009; Mohn et al., 2014). Finally, by sequence complementarity, piRNAs  
56 loaded PIWI proteins produced from the piRNA precursors will silence TEs through  
57 Transcriptional and/or Posttranscriptional Gene Silencing (Czech and Hannon, 2016). These  
58 loaded piRNAs deposited in oocytes constitute the memory of mobile sequences to repress at each  
59 generation thanks to maternal inheritance. Although our understanding of molecular events  
60 involved in the maintenance of active piRNA clusters through generations has deepened  
61 substantially during the discovery of piRNAs, major gaps still exist in our understanding of the  
62 establishment of active piRNA clusters.

63 TEs can be also transmitted through Horizontal Transfer (HT) corresponding to DNA  
64 transmission between unrelated individuals. It has been emphasized that HTs are more frequent



65 than originally thought (Gilbert and Feschotte, 2018; Hill and Betancourt, 2018; Loreto et al., 2008)  
66 raising the question of how and with which dynamics new piRNAs can be produced by naïve  
67 genome in absence of the maternal inheritance. One of the best documented HT is the case of the  
68 *P* element that successfully invaded the genome of natural populations of *D. melanogaster* during  
69 the 20th century and of *D. simulans* since the beginning of the 21st century (Anxolabéhère et al.,  
70 1988; Kofler et al., 2015). In *D. melanogaster* genome, *P* has a hot spot of insertion into the  
71 subtelomere of the *X* chromosome (cytological site *1A*) (Ronsseray et al., 1989). This locus (also  
72 known as Telomeric Associated Sequences (TAS) and hereafter called "*1A* cluster") is one of the  
73 *Drosophila* piRNA clusters (Brennecke et al., 2007) that has been shown to be dispensable in  
74 laboratory environment (Asif-Laidin et al., 2017). *Drosophila* can live without apparent defect in  
75 its absence. In addition, it contains repeats sharing homologies with the autosomic subtelomeric  
76 piRNA clusters of the *2R* and the *3R* chromosomes (*60F* and *100F* clusters). Finally, the *1A* cluster  
77 contains a region of 0.9 kb, called *T3*, not present elsewhere in the genome, a unique feature among  
78 all known piRNA clusters (Asif-Laidin et al., 2017). From a *P* copy inserted in the *1A* cluster, *P*  
79 piRNAs are produced in female germline capable of repressing hundreds of euchromatic active *P*  
80 elements (Brennecke et al., 2008; Reiss et al., 2004; Ronsseray et al., 1989; Srivastav et al., 2019). As  
81 for *P* elements, *lacZ* encoding *P* transgenes inserted in the *1A* cluster (for instance, the *P(lArB)*  
82 transgene inserted in the *P-1152* strain) have been shown to silence female germline expression of  
83 another *P-lacZ* transgene located in euchromatin (Roche and Rio, 1998; Ronsseray et al., 1998).  
84 This euchromatic counterpart (referred to the "target") represents a reliable reporter system for  
85 studying functional piRNA biology as their silencing depends on piRNA biogenesis actors (Josse  
86 et al., 2007; Le Thomas et al., 2014; Marie et al., 2017; Muerdter et al., 2012; Poyhonen et al., 2012;  
87 Todeschini et al., 2010; de Vanssay et al., 2012). The phenotype of this ovarian repression is taking  
88 place according to an ON/OFF mode, where egg chambers show strong (ON) or no (OFF) silencing  
89 (Josse et al., 2007). These targets also provide an easy access to the dynamics and function of  
90 piRNAs throughout generations. Previously, we have identified that when the subtelomeric  
91 *P(lArB)* transgene is paternally inherited, the number of fully repressed egg chambers for the *lacZ*  
92 target is low in the first generation and increases progressively to reach a maximum after five

93 generations on the ON/OFF mode (Josse et al., 2007). We also uncovered a case of paramutation,  
94 that fully converts a naïve locus made of seven *P(lacW)* transgenes, called *BX2*, into a stable piRNA  
95 cluster in one generation when maternal piRNAs matching the whole length of the transgene are  
96 inherited (de Vanssay et al., 2012). The switch of a naïve locus to an active one able to produce  
97 piRNAs will be hereafter referred as "conversion". This conversion is correlated with H3K9me3  
98 and Rhino enrichment on chromatin (Casier et al., 2019; Le Thomas et al., 2014), a feature of dual-  
99 strand germline piRNA cluster. Such conversion can occur when the locus producing maternal  
100 piRNAs is located on different chromosomes and is partially homologous to *P(lacW)* (Hermant et  
101 al., 2015). In this study, ovarian small RNAs analyses revealed that the conversion of the whole  
102 locus can be complete when tested at the third generation. These results along with others suggest  
103 that the piRNA machinery is eventually able to coopt a sequence unknown from the maternal  
104 piRNAs repertoire and trapped into a piRNA cluster to produce *de novo* piRNAs (Duc et al., 2019;  
105 Kofler, 2019; Zanni et al., 2013). Such circumstances can be encountered when the genome is facing  
106 HT by new TE that inserts into a piRNA cluster. The rareness of such events and the repetitive  
107 nature of piRNA clusters make it strenuous to directly address the precise latency and subsequent  
108 molecular events.

109 Using a combination of *P* derived transgenes inserted into piRNA cluster, strains with or without  
110 subtelomeric *1A* cluster, we have been able to copycat TE *neo* insertion into a piRNA cluster within  
111 a naïve genome. It allowed us to question the kinetics of newly produced piRNAs from the very  
112 first generation and for various sequences. Using reporter transgenes and in absence of  
113 complementary maternal piRNAs, we were able to correlate the quantity of new piRNAs with the  
114 rate of egg chamber conversion measured by functional repression assays. We also report  
115 evidence that the conversion efficiency of a sequence to be coopted by a piRNA cluster depends  
116 on its size. Indeed, sequences shorter than 1.8 kb not targeted by the maternal piRNA pool, have  
117 a higher probability to be converted from the first generation than a sequence of several kb,  
118 consistent with the size of an autonomous active TE. Consistently with previous results on  
119 euchromatic transgenes, this conversion is limited to sequences embedded into piRNA clusters  
120 (Olovnikov et al., 2013). This result suggests the existence of an active mechanism preventing

121 uncontrolled propagation to the flanking regions of a piRNA cluster. Unexpectedly for a piRNA  
122 cluster, our study also identifies the *1A* cluster as a heterogeneous piRNA cluster regarding  
123 transcription. Indeed, it is a dual-stranded cluster with regions that are Rhi and Moon dependent  
124 while other regions are Rhi and Moon independent.

125 Altogether, this study brings new insights on the biology and the dynamics of piRNA clusters.  
126 The fact that piRNA clusters are enriched in fragments of TEs rather than full length copies might  
127 indicate a functional role that was underestimated. First on the conversion of paternal piRNA  
128 cluster alleles induced at each generation by maternal piRNAs even if fragments are absent from  
129 this maternal piRNA pool. Second, it underlines new considerations concerning the rule of  
130 successful HT that might depend on the capacity to get inserted into a piRNA cluster, but also on  
131 the size of the inserted copy, favoring in some cases incomplete and non autonomous copies.

132

## 133 **Materials and Methods**

### 134 **Drosophila strains**

135 *P-1152* (FBti0005700) carries two enhancer trap *P(lArB)* transgenes inserted in the 1A X  
136 subtelomeric piRNA cluster. These *P(lArB)* transgenes are inserted in separate solo LTR  
137 subrepeats of *INV-4* in the same orientation and at the same nucleotide position (**Fig. 1A and B**)  
138 (Josse et al., 2008). They contain an in-frame translational fusion of the *E. coli lacZ* gene to the  
139 second exon of the *P* transposase gene (O’Kane and Gehring, 1987). This strain contains  
140 subtelomeric repeats on the X (*1A*) and 3R (*100F*) (Asif-Laidin et al., 2017) and is mutant for the  
141 *white* gene (*w<sup>c</sup>* allele). The *Canton* strain is devoid of *P* element as well as of the 1A piRNA cluster  
142 (Asif-Laidin et al., 2017). The *Canton* line used here is marked by a *yellow* allele affecting color  
143 pigmentation and is derived from the classical laboratory strain. The *BX2* strain contains seven  
144 *P(lacW)* (FBtp0000204) inserted in tandem at cytological site 50C on the second chromosome  
145 (Dorer and Henikoff, 1994). *P(RS3)* is a *P-FRT-white* transgene (FBtp0001534) inserted in the  
146 subtelomere of the 3R chromosomal arm (cytological site 100E3) (Dufourt et al., 2014). The *P-lacZ-*  
147 *target* (*BQ16*, FBtp0000154) carries a *P-lacZ* fusion enhancer trap transgene located at 64C in  
148 euchromatin strongly expressing *lacZ* in germline cells. When *P-1152* females are crossed by *P-*  
149 *lacZ* target males, *lacZ* is repressed in female germline (Josse et al., 2007). *OregonRC* strain is a  
150 classical laboratory strain that possess the X, 2R and 3R subtelomeric piRNA clusters.

151 Transgenic strains, called *pRFP-T3*, that carry insertions of the *RFP* reporter genes  
152 transcriptionally fused to the *T3* domain were obtained after transformation into the *w<sup>1118</sup>* strain.  
153 This transgene was first obtained by cloning the PCR-amplified *RFP* gene (using the primers: 5’-  
154 AGGTACCATGCCCAAGAAGAAGCGCAAGGTGGCCTCCTCCGAGGACGTCATCAAG-3’  
155 and 5’-ATCTAGATTAGGCGCCGGTGGAGTGGC-3’) into the pUASp plasmid (Rørth, 1998)  
156 digested by *KpnI* and *XbaI* and then by cloning in the *XbaI* site, the PCR amplified *T3* sequence  
157 (using the primers: 5’ AATCTAGACCCAGCAAATTTATGGATAAAC 3’ and 5’  
158 ATTCTAGACCTAATTTTTGGCAAAGTTGTAC 3’). The female germline expression of *pRFP-T3*  
159 is driven by the PBac[w[+mW.hs]=GreenEye.nosGAL4] FBtp0056000 transgene. RFP expression is

160 intense and reproducible at the ovarian stages 9-10. Germline specific gene knockdowns were  
161 performed by crossing short hairpin (shRNA) transgene strains with the maternal driver  
162 FBtp0056000 transgene. The **Table S1** summarizes the characteristics of each strains used in this  
163 study.

164

### 165 **Experimental Conditions**

166 All crosses were performed at 25°C and involved five to ten couples. The *y* and *w* marker genes  
167 present on the *Canton* and *P-1152* strains respectively were used to control the telomeric region in  
168 successive generations as well as to identify rare recombination events. We generated hemizygous  
169 females for the subtelomeric *P(lArB)* transgene with a maternal (MI) or paternal (PI) inheritance  
170 (G1 generation) that were then crossed with *Canton* males in successive generations. In parallel,  
171 repression capacities were measured using a germline functional assay at each generation for 4  
172 independent sublines by crossing hemizygous females at each generation with males containing  
173 a *P-lacZ* target transgene (*BQ16*). The repression was quantified as previously described by  
174 determining the percentage of egg chambers with no lacZ expression in the germline (Josse et al.,  
175 2007).

176 The repression capacity of subtelomeric *T3* piRNAs was quantified by determining the percentage  
177 of repressed egg chambers encoding a *pRFP-T3* reporter transgene. The quantification was  
178 restricted to stages 9-10 egg chambers where RFP expression was intense and reproducible.  
179 Images were acquired with an Axio-ApoTome (Zeiss) and ZEN2 software.

### 180 **Ovarian Small RNA sequencing and analysis**

181 Total ARN extraction, small RNA libraries preparation and sequencing were performed as  
182 described in Hermant et al (Hermant et al., 2015). Sequencing analyses were carried out according  
183 to (Asif-Laidin et al., 2017). In brief, after trimming from the adapter sequences, sequence reads  
184 were matched the reference sequences using Bowtie (Langmead et al., 2009). The 23-29nt reads  
185 (piRNA) matching the reference sequences with 0 mismatch were retained for subsequent

186 analysis. Annotation of small RNA libraries is described in **Table S2**. Sequence length  
187 distributions, small RNA mapping were run from a Galaxy instance publicly available at  
188 <http://mississippi.fr>. For library comparisons, read counts were normalized (normalization factor)  
189 relatively to the number of sequence reads aligning to the *D. melanogaster* genome but not to  
190 miscRNAs (including rRNA and snoRNA) or tRNAs. Similar results were obtained, when tested,  
191 by normalizing the read counts to one million miRNA reads.

## 192 **RT-qPCR experiments**

193 For each sample, 2 µg of total RNA was treated with DNase (Fermentas). One µg of DNase-treated  
194 RNA was used for reverse transcription using random hexamer primers (Fermentas). Real-time  
195 qPCR was performed on triplicates of each sample using primers referred in **Table S3**. *Rp49* and  
196 *otu* genes were used as reference. The same series of dilution of a mix of different RT preparations  
197 was used to normalize the quantity of transcripts in all RT preparations leading to standard  
198 quantity (Sq) values. Variations between technical triplicates were very low when compared to  
199 variations between biological replicates. The mean of the three technical replicates was then  
200 systematically used (meanSq). For each biological samples, we calculated the ratio  
201  $\text{meanSq}(\text{gene})/\text{meanSq}(\text{Reference Gene})$  to normalize the transcript quantity. Then, the mean of each  
202 sample ratio was used to compare the two conditions.

## 203 **ChIP-qPCR experiments and analyses**

204 For each sample, 50 pairs of ovaries from well-fed females (2 or 3 days) were manually dissected  
205 in cold 1XPBS, crossed with 1.8% formaldehyde, quenched with glycine, washed with 1XPBS,  
206 collected by centrifugation and pellets were flash-frozen in liquid nitrogen. Ovaries were grinded  
207 with pestle and lysed in ChIP lysis buffer (NaCl 100 mM, Tris pH8 50 mM, EDTA 5 mM, 1% SDS).  
208 Lysates were sonicated using Bioruptor (Bioruptor Standard Diagenode) three times during 15  
209 minutes (30 s ON, 30 s OFF) and cleared by centrifugation. 5% of cleared lysate was set aside to  
210 serve as input samples and the remainder was divided in two equal portions and incubated at  
211 4°C with anti-H3K9me3 (Active Motif Cat#39161) or negative IgG control (From Diagenode Low  
212 Cell ChIP-kit Cat#803-015) antibodies overnight under gentle rotation. Magnetic beads coupled to

213 G protein were washed two times in low salt wash buffer (0.1% SDS, 1% Triton X-100, EDTA 2  
214 mM, Tris pH8 20 mM, NaCl 150 mM), transferred into immunoprecipitated lysate and incubated at  
215 4°C during 1 hour under gentle rotation. The beads were washed two times in low salt wash  
216 buffer, one time in high salt wash buffer (0.1% SDS, 1% Triton X-100, EDTA 2 mM, Tris pH8 20  
217 mM, NaCl 500 mM), two times in LiCl wash buffer (LiCl 0.25M, 1% NP-40, 1% Na deoxycholate,  
218 EDTA 1 mM, Tris pH8 10 mM) and two times in TE buffer (Tris pH8 10 mM, EDTA 1 mM).  
219 Decrosslinking, elution and DNA purification are made by using IPure kit (Diagenode Cat#  
220 C03010015) as described in kit manual  
221 ([https://www.diagenode.com/files/products/kits/ipure\\_kit\\_v2\\_manual.pdf](https://www.diagenode.com/files/products/kits/ipure_kit_v2_manual.pdf)).

222 Real-time PCR was performed on duplicates for each biological sample leading to Cycle threshold  
223 (Ct) values. Variations between technical duplicates were very low compared to variations  
224 between biological replicates. The mean of the two technical replicates was then systematically  
225 used (meanCt). For each sample, IP fraction is normalized beside input to take account of sample  
226 preparation difference as follows:  $\Delta Ct$  [normalized ChIP] = (meanCt [ChIP] – (meanCt [Input] –  
227  $\text{Log}_2$  (Input Dilution Factor))) where meanCt [ChIP] is the Ct value measure for  
228 immunoprecipitated samples, meanCt [Input] is the Ct value measure for input and Input  
229 Dilution Factor corresponds to the chromatin fraction set aside for input (in this experiment, 5%  
230 of chromatin fraction was set aside, thus Input Dilution Factor is 20). From  $\Delta Ct$  values Fold  
231 Enrichment is calculated for each sample as follows: Fold Enrichment =  $2^{-(\Delta Ct$  [normalized  
232 NSI]). For each samples, fold enrichment is normalized on 42AB region which is known to be  
233 enriched in H3K9me3 mark.

#### 234 **Data availability**

235 Small RNA sequencing data will be available through the NCBI GEO database.

## 236 Results

### 237 Four generations are required to trigger full egg chamber conversion of paternally inherited 238 subtelomeric transgenes

239 Efficient silencing of an euchromatic *P-lacZ* target transgene can be induced by subtelomeric  
240 counterpart *P(lArB)* transgene (Roche and Rio, 1998). Previously, we have shown that this  
241 repression is female germline specific, has a maternal effect, shows a variegated ON/OFF *lacZ* egg  
242 chambers repression and depends on the piRNA biogenesis pathway (Josse et al., 2007, 2008;  
243 Muerdter et al., 2012; Todeschini et al., 2010; de Vanssay et al., 2012). On the contrary, paternally  
244 inherited subtelomeric *P(lArB)* transgene induces *lacZ* silencing in a restrictive number of ovarian  
245 egg chambers at the first generation. This number increases in subsequent generations to reach a  
246 maximum after five generations (Josse et al., 2007). The progressive increasing number of  
247 repressed egg chambers supports the model that the number of ovarian *lacZ* piRNAs produced  
248 by the subtelomeric *P(lArB)* is proportionally increasing across generations. To test this model,  
249 we performed reciprocal crosses using the *P-1152* and the *Canton* strains.

250 The *P-1152* strain contains two *P(lArB)* transgenes inserted in the *1A* cluster (**Fig. 1A, 1B**). The  
251 *Canton* strain lacks both *P*-derived transgenes and the *1A* cluster (**Table S1**). In these crosses, the  
252 subtelomeric *P(lArB)* transgenes were first paternally or maternally inherited and then maternally  
253 maintained at heterozygous state in the successive generations to established paternal or maternal  
254 lineages (*P-1152\_P(lArB)-PI* and *P-1152\_P(lArB)-MI* respectively) (**Fig. 1C**). Ovarian *lacZ*  
255 repression was assayed on *P-lacZ* target by crossing females at each generation with euchromatic  
256 *lacZ* target containing males for four independent lines in both lineages (**Fig. S1A**). In the *P(lArB)-*  
257 *PI*, few egg chambers showed repression in the first generation (G1; 9.25%), to reach the same  
258 magnitude of repression for the maternal lineage after four generations (>90%) (**Fig. S1B** and  
259 **Table S4**), reminiscent of previous results obtained with another euchromatic *lac-Z* target (Josse  
260 et al., 2007). To correlate the level of repression with the production of subtelomeric *lacZ* piRNAs,  
261 ovarian small RNAs from paternal and maternal lineages were sequenced and normalized 23-29  
262 nt *P(lArB)* small RNAs were analyzed. Ovarian libraries prepared from *P-1152* strain, carrying



263 homozygous *P(lArB)* transgenes, and from *Canton*, devoid of transgene, have been included as  
264 representative of G0 for maternal and paternal lineages (**Table S5**). The number of 23-29 nt small  
265 RNAs in the *P(lArB)-PI* increases progressively at each generation (**Fig. 1D** and **1E**) reaching a  
266 plateau at G4. Interestingly, the progressive increase of *P(lArB)* small RNAs is observed for each  
267 region of the transgene supporting the hypothesis that they are produced homogeneously,  
268 independently of their position within the transgene or their sequence origin (*Drosophila rosy* and  
269 *Adh* genes or *E. coli lacZ* gene (**Fig. S2**) and supporting an ON/OFF egg chamber state. The  
270 enrichment in uridine as first nucleotide for the 23-29 nt small RNAs (1U bias), a signature of  
271 piRNA biogenesis, reinforces the assumption that these small RNAs are piRNAs (**Fig. 1E**)  
272 consistent with previous mutant analyses (Josse et al., 2007; Le Thomas et al., 2014; Muerdter et  
273 al., 2012; Todeschini et al., 2010; de Vanssay et al., 2012). It is noteworthy, that *P(lArB)* piRNA  
274 distribution indicates that it is part of a dual strand piRNA cluster (**Fig. 1D**). The delay of  
275 conversion at G1 in *P(lArB)-PI* appears to be a general phenomenon, as it is observed in several  
276 genetic backgrounds (*Canton* and *w<sup>1118</sup>* **Fig. S3B**), for several transgenes (*P(lArB)* and *P(RS3)*,  
277 including a *P* derived transgene inserted in the *100F* piRNA cluster (**Fig. S4**)) and for several  
278 piRNA clusters (*1A* and *100F* clusters (**Fig. 1E, S3** and **S4**)).

279 As in the case of an HT, in G1 *P(lArB)-PI*, males contribute only to paternal transgenes without  
280 contributing to complementary piRNAs (**Fig. 1D**). Nonetheless, *P(lArB)* are inserted into the *1A*  
281 cluster that shares similarities with the other piRNA clusters and in particular with the autosomic  
282 *60F* and *100F* subtelomeric clusters (Asif-Laidin et al., 2017). Our results suggest that the  
283 redundancy of subtelomeric loci rarely converts stochastically *P(lArB)* infrequently in a very  
284 restricted number of egg chambers in the first generation, rather than an overall progressive  
285 increase of piRNAs produced by all egg chambers. The frequency of fully converted egg chambers  
286 is then increasing at each generation thanks to the maternal pool of piRNAs enriched in *P(lArB)*  
287 sequences.

288 Interestingly, in the *P(lArB)-MI*, the number of *P(lArB)* piRNAs is progressively decreasing across  
289 generations (**Fig. 1E** and **S2**). This result might suggest that there is a persistence of maternal

290 piRNAs at the adult stage explaining a progressive decay over four generations. This decay stops  
291 when the amount of piRNAs achieved a balance that corresponds to about half of the piRNAs  
292 produced by homozygous *P(lArB)* females (**Fig. 1E** and **Table S5**).

293 In conclusion, the transgenerational memory of maternal piRNAs inheritance can persist over few  
294 generations whether the lineage is paternal or maternal. How production of *P(lArB)* piRNAs is  
295 initiated in absence of complementary maternally inherited piRNAs in G1 is unclear. We  
296 hypothesized that maternally inherited piRNAs targeting the adjacent *1A cluster* might be the  
297 trigger inducing propagation to the whole *P(lArB)* transgene.

298

### 299 **Complete conversion of paternal *1A* cluster allele occurs within one generation**

300 The *1A* cluster is composed of repeats shared with at least the autosomic *60F* and *100F*  
301 subtelomeric piRNA clusters and of the 0.9 kb *T3* domain restricted to the *1A* cluster, used as a  
302 representative of the *1A* cluster (**Fig. 1A**) (Asif-Laidin et al., 2017).

303 Based on the progressive number of egg chambers showing paternal *P(lArB)* conversion (**Fig. 1E**  
304 and **S1B**), we hypothesized that egg chambers having a paternally inherited *T3* domain located  
305 in the same *1A* cluster should be converted with the same dynamics over four generations. To test  
306 this hypothesis, the small RNAs libraries were reanalyzed by aligning the 23-29 nt reads on the  
307 *T3* domain. Unexpectedly, the same number of normalized *T3* 23-29 nt small RNAs was observed  
308 in both paternal (*T3-PI*) and maternal (*T3-MI*) lineages from the very first generation (**Fig 1D, 1F**  
309 and **Table S5**). This one generation *T3* conversion was also observed in the *w<sup>1118</sup>* strain excluding  
310 a background effect (**Fig. S3C**). Consistent with previous work, the 23-29 nt *T3* small RNAs are  
311 mostly produced from one strand (**Fig. 1D** (Asif-Laidin et al., 2017)).

312 To test if the *T3* small RNAs were fully functional for repression from the first generation, we  
313 designed a repression sensor using the Red Fluorescent Protein (*RFP*) reporter gene  
314 transcriptionally fused to the *T3* domain and expressed under the control of the *UAS<sub>p</sub>* sequences  
315 (*pRFP-T3*). *RFP* expression was induced in female germline by the Gal4 protein expressed under

316 the control of the *nanos* promoter (**Fig. 2A**). Female germline expression of *pRFP-T3* was observed  
317 in the *w1118* genetic background devoid of *T3* encoding *1A* cluster (**Fig. 2B**). The *OregonRC* strain  
318 was used as a donor of *T3* piRNAs, because it contains the *1A* cluster without *P* derived sequences  
319 (*P* transgenes or *P* natural element) (Asif-Laidin et al., 2017) that interfere with the *pRFP-T3*  
320 transgene expression (**Fig. S5**). The ovarian *pRFP-T3* expression was repressed when the *1A*  
321 cluster was maternally inherited (98.12 % of RFP repressed egg chambers (n=373), **Fig. 2C**)  
322 suggesting that the small RNAs encoded by *T3* are functional germline repressors. In presence of  
323 a paternally inherited *T3* allele, *pRFP-T3* was also strongly repressed in the germline from the first  
324 generation (95.72% of RFP repressed egg chambers (n=467), **Fig. 2D**). This result suggests that the  
325 *T3* paternal allele is functionally converted in one generation consistent with the amount of *T3*  
326 small RNAs identified (**Fig. 1F**). To confirm that the *pRFP-T3* silencing was piRNA mediated,  
327 germline expression of Piwi was knocked down (Piwi KD) and RFP expression was monitored.  
328 **Figure 2E** shows that the transgene repression was affected by loss of Piwi supporting the idea  
329 that *pRFP-T3* is silenced by a piRNA mediated mechanism, as for the *P(lArB)* subtelomeric  
330 transgene (**Fig S5D**).

331 Thus, paternally inherited *P(lArB)* and *T3* have different rates of conversion and different piRNA  
332 distribution profiles (dual strand piRNA cluster for *P(lArB)* and asymmetrical dual strand piRNA  
333 cluster for *T3*) even though they are located in the *1A* cluster. Functionally, these results might  
334 indicate that piRNA cluster activation is not only dependent on maternal piRNAs inheritance at  
335 each generation that target the flanking subtelomeric regions, but could be also under the  
336 influence of environment as previously suggested (Komarov et al., 2020), leading to different  
337 mode of expression. The results strongly suggested that piRNA cluster loci are more intricate  
338 regions that originally though.

339  
340 **The *1A* cluster is a heterogeneous heterochromatic cluster where different piRNA transcription**  
341 **complexes can coexist**

342 To understand the puzzling difference of egg chambers conversion kinetics between the *P(lArB)*  
343 and the *T3* domain, we first analyzed chromatin states of both *P(lArB)* and *T3* sequences in ovaries  
344 of *MI* and *PI* lineages. Loss of H3K9me3 enrichment on *P(lArB)* and *T3* was observed when  
345 complementary maternal piRNA were absent (*PI* lineages) (**Fig. 3B**). Similar results on *P(lArB)*  
346 chromatin showing that high level of H3K9me3 requires maternally inherited piRNAs were  
347 previously observed (Le Thomas et al., 2014). Here, we showed that this is also true for the whole  
348 *1A* cluster, even though *T3* piRNA accumulation is not affected (**Fig. 1F**). Thus, H3K9me3  
349 enrichment appears to be not completely correlated with piRNA production.

350 We then measured ovarian RNA accumulation of both loci in both lineages by RT-qPCR, asking  
351 if differences of steady state transcription level might partly explain this discrepancy. At G1, RNA  
352 accumulation of *T3* and of different parts of *P(lArB)* was significantly higher when the *1A* cluster  
353 was paternally inherited (*i.e.* in absence of maternal complementary piRNAs) compare to the  
354 maternally inherited ones (see *PI* and *MI* control KD **Fig. 3D** and **3E**). After four generations, no  
355 more differences were observed between maternal and paternal lineages, reinforcing the fact that  
356 an equilibrium was reached at this generation (**Fig. S6A**). Despite the same RNA accumulation  
357 between the transgene and the *T3* domain in paternal lineage, *lacZ* and *pRFP-T3* repressions were  
358 different, as a limited number of egg chambers showed *lacZ* repression and almost all egg  
359 chambers showed *RFP* repression (**Fig. 3D** and **3E**). Thus, the low frequency of *P(lArB)* converted  
360 egg chambers at G1 in the paternal lineage cannot be explained by a simple defect of *P(lArB)*  
361 steady state transcription. One suggestion would be that *P(lArB)* and *T3* might be differently  
362 relying on piRNA transcription machinery during conversion leading to the different piRNA  
363 profiles observed in **Fig. 1D**.

364 To decipher such mechanism, germline knockdown of *Rhi* and *Monshiner* (*Moon*) were tested on  
365 transcript accumulation of maternally and paternally inherited loci. *Rhi* is a germline HP1  
366 homolog, a key factor of piRNA cluster chromatin that binds to H3K9me3 of dual-strand germline  
367 piRNA clusters (Klattenhoff et al., 2009; Mohn et al., 2014). *Moon*, a germline paralogue of TFIIA-  
368 L, is recruited at dual-strand germline piRNA cluster through its interaction with *Deadlock* that  
369 is bound to *Rhi* allowing the non canonical transcription of piRNA cluster heterochromatic loci

370 (Andersen et al., 2017). In the maternal lineage, germline knockdown of *Rhi* did not affect  
371 accumulation of *P(lArB)* nor *T3* RNAs (**Fig. 3D; S6B**). However, loss of *Rhi* was affecting *lacZ*  
372 repression but not *RFP-T3* repression (**Fig. 3D, S6D**). Lack of RNA accumulation of dual strand  
373 germline piRNA clusters has been previously observed in the context of embryonic Piwi-KD  
374 resulting in H3K9me3 and *Rhi* depletion (Akkouche et al., 2017). Using *lacZ* and *RFP-T3*  
375 functional assays, we also deciphered that loss of *Rhi* lead to different consequences as a restricted  
376 number of egg chambers were able to process transgenic RNAs into piRNAs, while this processing  
377 remained efficient on *T3* RNAs (**Fig. 3D, S6D**). When *P(lArB)* or *T3* were paternally inherited in  
378 the context of loss of *Rhi*, the quantity of RNA decreased when compared to wild type. Thus as  
379 for the maternal inheritance, *P(lArB)* and *T3* can be transcribed in a *Rhi* independent manner.  
380 Nevertheless, nearly no egg chambers were able to funnel the *P(lArB)* derived transcripts into the  
381 piRNA biogenesis machinery, resulting in low *lacZ* repression (**Fig. 3D**). On the contrary, all egg  
382 chambers were still able to silence *pRFP-T3* suggesting that *T3* derived transcripts were still  
383 substrate of the piRNA machinery to generate piRNAs (**Fig. 3D**).

384 As for the loss of *Rhi*, loss of *Moon* (**Fig. S6C**) induced the same *P(lArB)* RNA accumulation than  
385 in control when the transgenes were maternally inherited, while *lacZ* repression was impaired  
386 (**Fig. 3E**). However, when paternally inherited, *P(lArB)* RNA accumulation is strongly dependent  
387 on *Moon* (**Fig. 3E**). On the contrary, accumulation of *T3* RNA was increased in *Moon* KD in both  
388 lineages. However, these transcripts are properly tagged as piRNA precursors as the *RFP-T3*  
389 repression seems to be unaffected (**fig. 3E**).

390 Taken together, the quantity of piRNA precursor transcripts does not necessarily predict piRNA  
391 synthesis and silencing efficiency. This analysis also reveals that the loss of *Rhi* or *Moon* does not  
392 lead to the same consequences for the *P(lArB)* transgenes and the *T3* domain, although they are  
393 H3K9me3 enriched and included in the same *1A* locus and separated from few hundred of bp  
394 (**Fig. 3A**). Our data suggested that the primary properties of a sequence to become part of a piRNA  
395 cluster in absence of complementary maternally inherited piRNAs could depend on the genomic  
396 environment, on composition of chromatin but also on intrinsic features of sequences. We

397 therefore wondered if the size of the sequence to convert might be crucial in the cooption  
 398 efficiency.

399

400 **Frequency of cooption at 1A cluster might depend on the size of the sequence to convert**

401 The above results indicate that, at G1 of the paternal lineage, maternal piRNAs targeting the 1A  
 402 cluster (including *T1*, *T2*, *T4* and *INV4*) are more efficient to induce egg chamber conversion of  
 403 the 0.9 kb *T3* domain than the 18.3 kb *P(lArB)* (**Table 1**, **Fig. 1E** and **1F**).

404 **Table 1. Relation between sequence size and conversion efficiency between MI and PI lineages at G1.**

origin	sequence	size (kb)	Ratio of normalized piRNAs in MI over PI lineage at G1	Results in
<i>P-1152</i> strain in <i>Canton</i> background	<i>T3</i>	0.9	0.91	Fig. 1F
<i>P-1152</i> strain in <i>w1118</i> background	<i>T3</i>	0.9	0.93	Fig. S3C
<i>RS3</i> strain	<i>P(RS3)</i>	6	5.95	Fig. S4D
<i>P-1152</i> strain in <i>Canton</i> background	<i>P(lArB)</i>	18.3	6.35	Fig. 1E
<i>P-1152</i> strain in <i>w1118</i> background	<i>P(lArB)</i>	18.3	10.49	Fig. S3B

405 To further test if the sequence size might impact the conversion frequency, we took advantage of  
 406 1-the *BX2* strain that contains seven tandemly repeated *P(lacW)* transgenes that can be converted  
 407 into an active piRNA cluster by complementary maternal piRNAs in a Rhi and Moon dependent  
 408 manner (Casier et al., 2019; Hermant et al., 2015; de Vanssay et al., 2012) and 2-the *P-1152* or the  
 409 *RS3* strains that possess *P(lArB)* or *P(RS3)* transgenes inserted in the 1A or the 100F clusters,  
 410 respectively. In the following experiments, *P(lArB)* and *P(RS3)* were used as donors of maternal  
 411 piRNAs to convert *P(lacW)* (**Fig. 4A, 4F**). The *P(lArB)* and *P(lacW)* transgenes both encode the 3.5  
 412 kb *E. coli lacZ* gene and the 1.8 kb bacterial plasmid backbone (**Fig. 4A** and **4B**). The *P(RS3)* and  
 413 *P(lacW)* both encode the 4.1 kb *white* gene (**Fig. 4F** and **4G**). In addition, *P(lArB)*, *P(RS3)* and  
 414 *P(lacW)* have in common the 5' and 3' region of the *P* element. Heterozygous *P(lacW)* males were  
 415 crossed with hemizygous *P(lArB)* (**Fig. 4A, 4B** and **Fig S7A**) or heterozygous *P(RS3)* females (**Fig.**  
 416 **4F, 4G** and **Fig S7B**). The G1 progenies, that paternally inherited the *P(lacW)* transgenes and  
 417 maternally inherited the *P(lArB)* or *P(RS3)* piRNAs but not the DNA allele, were recovered and

418 crossed to each other for several generations (**Fig. S7**). Ovarian 23-29 nt RNAs were analyzed at  
 419 G1 and after several generations and mapped to the different parts of the *P(lacW)*. Maternally  
 420 inherited piRNAs complementary to *lacZ* and plasmid sequences produced from *P(lArB)* were  
 421 able to convert from the G1 their counterpart regions present in *P(lacW)* (**Fig. 4C** and **4D**).  
 422 Similarly, the *lacZ* gene, expressed as a transcriptional fusion with *P* first exons, and the *white* gene  
 423 of *P(lacW)* were converted from the first generation by complementary piRNAs synthesized from  
 424 the maternal *P(RS3)* allele (**Fig. 4H** and **4I**). However, piRNA synthesis was restricted in G1 for  
 425 the 4.1 kb *white* sequence absent from *P(lArB)* (2761 normalized piRNAs) and increased  
 426 significantly in G4 (6110 normalized piRNAs) (**Fig. 4E** and **Table 2**), whereas it occurs from the  
 427 first generation for the 1.8 kb plasmid sequence, not included in *P(RS3)* (**Fig. 4J** and **Table 2**).

428 **Table 2. Conversion efficiency between the first generation and after several generations, in absence of**  
 429 **inherited piRNAs related to sequence size**

origin	sequence	size (kb)	Ratio of normalized piRNAs numbers in Gn over G1	Result in
<i>P-1152</i> strain in <i>Canton</i> background	<i>T3</i>	0.9	0.84 <sup>a</sup>	Fig. 1F
♀ <i>RS3</i> x ♂ <i>BX2</i>	<i>plasmid</i>	1.8	1.25 <sup>b</sup>	Fig. 4J
♀ <i>P-1152</i> x ♂ <i>BX2</i>	<i>white</i>	4.1	2.25 <sup>a</sup>	Fig. 4E
<i>P-1152</i> strain in <i>Canton</i> background	<i>P(lArB)</i>	18.3	3.2 <sup>a</sup>	Fig. 1E

430 <sup>a</sup>: Gn corresponds to G4; <sup>b</sup>: Gn corresponds to G7

431  
 432 In order to challenge the hypothesis of size, we tried to take into account the composition of each  
 433 sequences. As all small RNA analyses were performed with 0 mismatch between small RNAs and  
 434 the tested sequence, we reanalyzed the libraries allowing 1, 2 and 3 mismatches to estimate the  
 435 number of 23-29 nt RNAs per kb in G1. No clear bias was identified in the number of piRNAs  
 436 towards *T3* and the plasmid sequences vs *P(lArB)* and the *white* gene (**Table S6**). This results  
 437 strongly suggests that the frequency of converted egg chambers is independent on sequence  
 438 composition but dependent on the size of the sequence to convert: in one generation, incomplete  
 439 conversion occurs for sequence longer than 4 kb, like *white* or *P(lArB)*), whereas complete  
 440 conversion occurs in one generation for a shorter sequence, like the plasmid sequence or *T3*

441 (summarized in **Tables 1** and **2**). Once fully established, *de novo* conversions are then stably  
442 transgenerationally inherited due to the newly maternally inherited piRNAs.

443

#### 444 **Conversion is restricted to sequences embedded into piRNA clusters**

445 The above results described the effect of maternal piRNAs targeting neighboring sequences  
446 within a piRNA cluster. We then looked if the spreading of conversion can be observed on  
447 flanking regions. The *P(lacW)* transgenes are inserted in a non-coding domain shared by the *AGO1*  
448 and the *mRPL53* genes (**Fig. S8A**). Analysis of 23-29 nt small RNAs shows that the libraries are  
449 weakly enriched in small RNAs complementary to the two genes (**Fig. S8B**). These piRNAs  
450 correspond for the majority to the transcript strands suggesting that they are produced primarily  
451 without amplification. The same analysis was performed on the *CG17636* gene, the first gene on  
452 the X chromosome (*i.e.* close to the *1A* cluster). Few and unchanged 23-29 nt matching to the  
453 *CG17636* gene were identified that were constant across generations (**Fig. S8C**). Similar results  
454 were observed on endogenous loci present in *P(lacW)* (**Fig. S9A**) or *P(lArB)* transgenes (**Fig. S9B**).  
455 Therefore, the spreading of conversion from transgenes sequences on *cis* observed here and in  
456 earlier study (Olovnikov et al., 2013) or on *trans* outside of piRNA clusters is limited suggesting  
457 the existence of a tight control that restricts heterochromatin and defines their borders.

458

#### 459 **Discussion**

460 It is now well established that TEs are harmful for their host genomes, although being source of  
461 evolution by providing a large range of genetic opportunities, like new tissue expression or new  
462 genes (Jangam et al., 2017). Few of these events can be then positively selected leading to the so  
463 called "molecular domestication" (Miller et al., 1999). TEs have the capacity to transpose into each  
464 other or close to each other (Gao et al., 2012) leading to highly enriched regions of TEs fragments  
465 (Bergman et al., 2006). These loci were shown to encode piRNAs involved in TEs silencing in



466 gonads of Metazoans (Brennecke et al., 2007; Gunawardane et al., 2007), the ultimate  
467 domestication event where TEs are controlling their own mobility. From now on, our  
468 understanding of the biochemistry and the genetics of the piRNA biogenesis is well advanced  
469 (Sato and Siomi, 2020). In *Drosophila*, it has been shown that maternally inherited piRNA can  
470 activate loci containing transgenes derived of *P* or *I* elements (Olovnikov et al., 2013; Shpiz et al.,  
471 2014; de Vanssay et al., 2012). Thanks to early studies on hybrids dysgenesis, it is now understood  
472 how newly horizontally transferred *P* elements can be maintained in the *Drosophila* population  
473 because of their capacity to insert into the *1A* cluster (Brennecke et al., 2008; Josse et al., 2007, 2008;  
474 Ronsseray et al., 1989). This piRNA cluster can be lost under laboratory environment without  
475 affecting fly survival despite the fact that *P* could be amplified until it gets stabilized most likely  
476 by new insertion into another piRNA cluster (Asif-Laidin et al., 2017). Analogous observations  
477 were found with a case of a *flamenco* locus deletion that derepressed the *Zam* element leading to  
478 neo insertion into germline piRNA cluster (Duc et al., 2019). It is well accepted that such new  
479 insertions will *in fine* produce new piRNAs. However, how and at what rate newly inserted TE  
480 copy will lose euchromatic identity to acquire piRNA cluster ones was unknown and was  
481 questioned in this study.

482

### 483 **Insights in early conversion of sequences inserted into piRNA clusters**

484 Variegation of ovarian *lacZ* repression strongly suggests that the progressive increasing number  
485 of transgenic piRNAs across generations might reflect an increasing number of fully converted  
486 egg chambers rather than a global increase of piRNAs synthesis by all egg chambers (**Fig. S1B**  
487 (Josse et al., 2007)). Our analysis revealed that, to be stably converted, a sequence not primarily  
488 targeted by maternal piRNAs needs to be embedded into piRNA cluster. This result is consistent  
489 with previous works showing mild effect on flanking sequence of *I* transgene insertion  
490 (Olovnikov et al., 2013) avoiding pervasive piRNA cluster marks propagation. Moreover, our  
491 results pinpoint that a sequence shorter than 1.8 kb (*T3* in *1A* cluster or plasmid in *P(lacW)*) is fully  
492 functionally converted in egg chambers in one generation, whereas a sequence longer than 4 kb

493 (*white* in *P(lacW)* or *P(lArB)* in *1A* cluster) needs about four generations (**Fig. 1, 3** and **Table 1**). It  
494 was previously shown that the amount of piRNAs is reduced between homozygous and  
495 hemizygous *P(lArB)* containing strains (de Vanssay et al., 2012). Here we report an unexpected  
496 kinetics where transgenerational inheritance includes accumulation of piRNAs at each  
497 generation, when *P(lArB)* is paternally inherited and a progressive decrease of piRNAs amount  
498 when it is maternally inherited starting from a homozygous locus (**Fig. 1E**). All these results  
499 emphasize that transgenerational inheritance is dependent on the number of piRNA cluster alleles  
500 but also on the size of the sequence to convert, thus determining the conversion rate persistence  
501 over four generations.

502

### 503 **Model of TE cooption by piRNA cluster: the secret of a successful horizontal transfer**

504 Horizontal transfer of TEs have been described in various organisms emphasizing that such  
505 events can occur more frequently than originally thought (Gilbert and Feschotte, 2018). However,  
506 the mechanisms and kinetics involved in the battle to silence these new invaders are not well  
507 understood. It can be modeled that upon new insertion into a piRNA cluster, a TE sequence is  
508 embedded in a locus that is targeted by maternal piRNAs (**Fig. 5**). In a sense, the paternally  
509 inherited *P(lArB)* transgene inserted in the *1A* cluster or the *white* gene in the context of *P(lacW)*  
510 activated by *P(lArB)* recapitulate a TE of few kb in length inserted into a piRNA cluster in a naïve  
511 genome. Therefore, the success for new TE invasion will depend, first on its capacity to be trapped  
512 into a piRNA cluster impeding any further transposition. Once integrated, our results suggest that  
513 at each generation, stochastic egg chamber conversion dependent on Piwi protein is occurring  
514 more frequently for shorter sequences, during which the TE copy will acquired stable piRNA  
515 cluster marks, like H3K9me3 enrichment and will be functional for silencing the active  
516 euchromatic copies (Casier et al., 2019; Le Thomas et al., 2013) **Fig. 3B**). Piwi associated with  
517 maternal piRNAs deposits H3K9me3 marks to complementary sequences during early  
518 embryogenesis that specify at each generation the piRNA clusters to maintain. Short sequences  
519 not included in the maternal piRNA repertoire but surrounded by Piwi-piRNA complexes will be

520 more efficiently converted than longer ones (Fig. 5). Once the conversion occurs, this epigenetic  
521 state is stably transmitted to the next generation. Therefore, the increasing number of converted  
522 egg chambers observed at each generation corresponds to the addition of newly converted plus  
523 the inherited ones.

#### 524 **Paramutation at piRNA clusters occurs at each generation to maintain genome stability**

525 Because piRNA clusters are composed for the most part by fragments of TEs, it is difficult to  
526 conduct an extensive analysis of their dynamics. We concentrated our effort on the *1A* cluster,  
527 because 1- strains devoid of this locus exist, 2- *T3* localization is restricted to the X subtelomere  
528 and 3- it is long enough to conduct simple molecular analysis (Asif-Laidin et al., 2017). Assuming  
529 that *T3* recapitulate the properties of the whole *1A* cluster, our analysis indicates that maternally  
530 inherited piRNAs targeting redundant subtelomeric sequences, such as *INV-4*, *T1*, *T2* and *T4*,  
531 produced by the autosomic subtelomeres, are able to epigenetically convert *T3* from the first  
532 generation. Such epigenetic conversion dependent on maternal piRNAs, known as paramutation  
533 (de Vanssay et al., 2012), might be the keystone to convert paternal alleles of germline piRNA  
534 clusters at each generation. This result reveals the selection forces that might exist to preserve  
535 hundreds of piRNA clusters with redundant sequences that secure efficient TE control at each  
536 generation. Moreover, we speculate that positive selection of small fragments of TEs in piRNA  
537 clusters, can favor transgenerational maintenance of active piRNA clusters.

#### 538 **Germline piRNA clusters are heterogeneous entities**

539 Dual-strand and uni-strand clusters are the two main types of germline piRNA clusters that have  
540 been described in *Drosophila* ovaries. The *42AB* cluster is the largest dual-strand cluster which is  
541 H3K9me3 enriched allowing recruitment of Rhi that binds to Deadlock interacting with Moon,  
542 one of the constituent of the germline RNA polymerase II required for the transcription in both  
543 strand. On the contrary, transcription of the *20A* cluster is independent of Rhi and Moon even  
544 though it is mildly enriched in H3K9me3 (Andersen et al., 2017; Mohn et al., 2014). So far, all  
545 descriptions of piRNA clusters considered these loci as homogeneous units. *P(IArB)* transgenes

546 and *T3* domain are included in the *1A* cluster and are separated from about 400 bp. Besides they  
547 are H3K9me3 enriched, our data indicates that they differ by the distribution of piRNAs on both  
548 strands, as *P(lArB)* results in a dual-strand cluster, while *T3* is an asymmetric dual-strand cluster  
549 (Asif-Laidin et al., 2017)), by their dependency on Rhi and Moon and by the frequency of their  
550 conversion induced by the surrounding cluster. These results suggest an unexpected feature that  
551 some of the piRNA clusters, like the *1A* cluster, are heterogeneous. Does this property is unique  
552 to the *1A* cluster? As a matter of fact, probably not. In the light of this results and by reanalyzing  
553 recent literature, the *80F* cluster, displays an heterogeneous piRNA profile distribution resembles  
554 to a dual-strand cluster for the most part of the locus and to a uni-strand cluster in the 3' region of  
555 the cluster (Andersen et al., 2017; ElMaghraby et al., 2019; Mohn et al., 2014). Furthermore, the  
556 germline dual-strand piRNA clusters *42AB* (Rhi and Moon dependent) and *38C1* (Rhi dependent  
557 and Moon independent) display, in loss of Moon, an increase number of piRNAs localized to the  
558 3' region of *42AB* cluster and to the whole *38C1* cluster. In the case of the *38C1* cluster, alternative  
559 promoters were identified (Andersen et al., 2017). Extrapolation from our results might indicate  
560 that silencing capacity in loss of Moon should be maintained for these specific regions if using  
561 appropriate target transgenes. Our RT-qPCR data revealed that Moon is required for the  
562 transcription and the production of *P(lArB)* piRNAs and prevents efficient *T3* transcription (**Fig.**  
563 **3E**) as for the *38C1* cluster (Andersen et al., 2017) responding to different properties depending on  
564 genomic environment (Komarov et al., 2020) and leading also to a heterogeneous cluster.

## 565 **Conclusion**

566 Depending on environmental conditions, it takes up to 20-40 generations in *D. simulans* to  
567 efficiently produce new piRNAs directed against *P* element when tested from the beginning of  
568 the invasion (Kofler et al., 2018). This result is echoed by early studies that identified a time period  
569 of about three decades to observe a global *P* activity decrease in natural populations (Kidwell,  
570 1983) and simulations based on *P* element literature in *D. melanogaster* (Kelleher et al., 2018).  
571 Computer simulations determined three distinct phases during TE invasion: a proliferative period  
572 followed by segregation of cluster insertions and finally fixation of cluster insertions (Kofler,

573 2019). The latter step requires reprogramming by imprinting piRNA cluster properties as shown  
574 for the ZAM element in *Drosophila* (Duc et al., 2019) that may take few generations. Here we are  
575 showing evidence that cooption of sequence of a size consistent with a TE by a piRNA cluster can  
576 happen as fast as four generations. From these results, it can be considered that the genome  
577 response to control a new TE horizontally transferred principally rely on its capacity to insert into  
578 an active piRNA cluster. Number of studies have also suggested that heterochromatic regions,  
579 like pericentromeric regions containing piRNA clusters, show large accumulation of TEs as well  
580 as strong suppression of recombination (Kent et al., 2017; Rizzon et al., 2002) limiting deletion  
581 within TEs. Thus, while currently speculative, we proposed that successful horizontal transfer  
582 might also drive positive selection toward TE fragments insertion, instead of full length copies, in  
583 piRNA cluster. Moreover, the progressive conversion for long sequences might be interpreted as  
584 a safeguard that might exist to define piRNA cluster borders, avoiding efficient propagation on  
585 flanking euchromatic regions. Future works will be required to identify factors involved to define  
586 the piRNA cluster borders. We took advantage of the *1A* cluster and *BX2* locus, one of the few  
587 clusters with functional assays, that we developed (this work and (de Vanssay et al., 2012)) to  
588 untangle the influence of size of fragments on conversion rate and heterogeneity of piRNA  
589 clusters. Finally, this work highlights that maternal piRNA legacy is a fundamental prerequisite  
590 for genome stability that has to be activated and once activated is inherited at each generation in  
591 order to maintain active maternal and paternal piRNA cluster alleles.

592

593

594

## 595 **Acknowledgements**

596 We thank the Bloomington Stock Center for providing stocks, FlyBase.org for providing valuable  
597 databases and the ARTbio/Bioinformatics analyses platform at the IBPS for the development of  
598 valuable tools (<http://artbio.fr/>). We thank Sandra Duharcourt for early comments on the  
599 manuscript, Ritha Zamy for technical assistance. This work was supported by fellowships from

600 the Ministère de l'Enseignement Supérieur et de la Recherche to A.A-L., K.C. and E.V., by  
601 fellowships from Ligue Nationale Contre le Cancer to K.C., and by grants from the Association  
602 de la Recherche contre le Cancer (Fondation ARC, SFI20121205921, SFI20131200470), the  
603 Fondation pour la Recherche Médicale, (FRM DEP20131128532), the Association Nationale de la  
604 Recherche (ANR, project "plastisipi") to S.R., and the University Pierre et Marie Curie (Emergence  
605 EME1223) to L.T. K.C. thanks the Réseau André Picard and the Société Française de Biologie du  
606 Développement for its support on travel fellowship grants. We thank the TRiP at Harvard Medical  
607 School (NIH/NIGMS R01-GM084947) for providing transgenic RNAi fly stocks used in this  
608 study."

609

610 **Figure legends**

611 **Figure 1. Conversion at the 1A cluster paternal allele.** **A.** Schematic representation of one of the  
612 X subtelomeric repeats found in the 1A cluster. The repeats are composed of domains also found  
613 in the 60F and 100F subtelomeres (solo LTRs of the *INV-4* retrotransposon (grey boxes), *T1* (blue  
614 box), *T2* (pink boxes) and *T4* (yellow box) domains). The *T3* domain is a 0.9 kb domain found only  
615 in the 1A cluster (green box). The 1A, 60F and 100F repeats are germline subtelomeric piRNA  
616 clusters. The asterisks show the two insertion sites of *P(lArB)* transgenes inserted in the same  
617 repeats of the *P-1152* strain. **B.** Structure of *P(lArB)* transgene inserted in the 1A cluster of the *P-*  
618 *1152* strain. The *P(lArB)* transgene contains the 5' and 3' *P* sequences, the *Bluescript* plasmid  
619 backbone, *rosy*, and *Adh* genes and the *lacZ* gene under the control of the weak *P* promoter. The  
620 transcription terminators of *rosy* and *lacZ* correspond to the terminators of *CG3281* (tCG) and  
621 *Hsp70Ab* (tHs) genes, respectively. The X subtelomeric repeats (1.8 kb) and the *P(lArB)* (18.3 kb)  
622 are not drawn to scale. Arrows represent the sense of transcription of genes. **C.** The 1A cluster  
623 Paternal Inheritance (*PI*) and Maternal Inheritance (*MI*) were obtained by two reciprocal crosses  
624 between the *P-1152* and the *Canton* strains. The names of each strain are indicated above the  
625 genotype in G0. The *T3* containing 1A cluster presents in the *P-1152* strain is referred as "*T3+*" and  
626 absent from the *Canton* strain and referred as "*T3-*". The genotypes show the maternal alleles  
627 above the fraction. Of note, the *P-1152* strain contains the 100F cluster and the *Canton* strain  
628 encodes the 60F and the 100F clusters. **D.** Scheme of experimental design is showing regions  
629 complementary to maternal piRNAs in G1 of each lineage. Below are shown size distribution and  
630 distribution of 23-29 nt piRNAs over the *P(lArB)* and the *T3* domain (right). The sense piRNAs are  
631 in red and the antisense piRNAs are in blue. **E, F.** Normalized 23-29 nt reads mapping to the  
632 *P(lArB)* transgene (**D**) and *T3* (**E**) in the paternal and maternal lineages are shown. The percentage  
633 of 23–29 nt *P(lArB)* small RNAs beginning with a 5' uridine (1U bias) are indicated. This  
634 enrichment is a signature of piRNAs. Error bars indicate standard deviation in G1, G2 and G3  
635 (n=2).

636 **Fig. 2. Functional conversion of paternal 1A cluster in one generation is dependent on piRNA**  
637 **pathway. A.** Schematic representation of the experiments: Flies from the *Oregon* strain containing  
638 the 1A cluster encoding *T3* and the *Gal 4* gene driven by the germinal *nanos* promoter (*nos* driver)  
639 were crossed with flies from the *w1118* strain devoid of the 1A cluster encoding the *pRFP-T3*  
640 reporter transgene. **B.** Strong ovarian germline *RFP* expression in progenies of a cross between  
641 females expressing the *nos* driver with males encoding the *pRFP-T3* reporter transgene in absence  
642 of 1A cluster. **C.** Maternally inherited *T3* is able to strongly represses ovarian germline expression  
643 of *pRFP-T3* transgene. **D.** Paternally inherited *T3* is able to strongly represses ovarian germline  
644 expression of *pRFP-T3* transgene. **E-F.** Ovarian *pRFP-T3* repression induced by maternally  
645 inherited *T3* (knockdown for *white* (control KD, **E**) is abolish in germline knockdown for Piwi  
646 (Piwi-KD, **F**). Repression was assayed by counting the percentage of RFP silenced egg chambers  
647 at stage 8-10. The total number of egg chambers are indicated in parenthesis. *RFP* expression (red)  
648 and *Dapi* staining (white) indicates nuclei. Parental crosses are indicated above pictures.

649  
650 **Fig. 3. Chromatin and transcription reveal an heterogeneous 1A cluster. A.** Schematic illustration  
651 of the *P(lArB)* transgene inserted in the 1A cluster and the position of the primers used in this  
652 study (Table S3). Of note, *P(lArB)* and subtelomeric domains are not drawn at scale. **B.** H3K9me3  
653 ovarian enrichment measured by ChIP-qPCR on three different regions of *P(lArB)* shown in A  
654 and on *T3* in maternal and paternal inheritance (MI and PI). Signal was normalized on the *42AB*  
655 region highly enriched in the H3K9me3 mark. The ChIP experiments were performed on  
656 independent biological samples followed by duplicate qPCRs on each sample. Unpaired t-test was  
657 used to calculate significance of the differences ( $p < 0.05$ ,  $n = 4$ ), ns: not significant. **C.** Schematic  
658 crosses used to test the effect of germline knockdown of Rhino (Rhi KD) and Moonshiner (Moon  
659 KD) on transcription of *P(lArB)* and *T3*. **D, E.** Measure of RNA accumulation of three different  
660 regions of *P(lArB)* and *T3* at G1 in Rhi-KD (**D**) and Moon KD (**E**) in ovaries determined by RT-  
661 qPCR and normalized to the expression of *rp49* gene. Significant p-values are given (bilateral  
662 Student's t-test,  $n=3$ ). Below the graphs are shown the percentage of egg chambers showing *lacZ*



663 or *RFP-T3* repression (**Fig. S6E**), *n* represents the total of egg chambers. This study reveals that  
664 *P(lArB)* is Rhi and Moon dependent and *T3* is Rhi and Moon independent, although they are  
665 located within the *1A* cluster. ns: not significant  $p>0.05$ , \*  $p<0.05$ , \*\*  $p<0.01$ , \*\*\*  $p<0.001$ , \*\*\*\*  
666  $p<0.0001$ .

667

668 **Fig. 4. Dynamics of conversion of the *P(lacW)* transgene by *P(lArB)* inserted in *1A* cluster or**  
669 ***P(RS3)* inserted in *100F* cluster.** Structures of *P(lArB)* transgene inserted in the *1A* cluster (**A**) and  
670 one of the seven *P(lacW)* transgene inserted in the *BX2* locus. (**B**). Normalized reads of 23-29 nt  
671 mapping to *lacZ* (**C**), plasmid sequence (**D**) and *white* (**E**) analyzed at the first (G1) and fourth  
672 generation (G4) are used to draw the graphs. While the numbers of 23-29 nt small RNAs are  
673 similar between G1 and G4 for *lacZ*, their quantity is more than 2 fold increased for *white* between  
674 G1 and G4. (**F**) Structures of the *P(RS3)* transgene inserted in the *100F* cluster and *P(lacW)*  
675 transgene inserted in the *BX2* locus (**G**). Normalized reads of 23-29 nt mapping to *lacZ* (**H**), *white*  
676 (**I**) and plasmid sequence (**J**) analyzed at the first (G1) and seventh generation (G7) are used to  
677 draw the graphs. The amount of normalized small RNAs reads under each panel as well as the  
678 fraction of 1U bias at 5' are indicated in each panel. The 5' 1U enrichment is a signature suggesting  
679 piRNA biogenesis origin. Note that *P(lArB)* (18.3 kb), *P(RS3)* (6 kb) and *P(lacW)* (10.7kb)  
680 transgenes are not drawn to scale. The numbers of 23-29 nt small RNAs are similar between the  
681 first and seventh generations for all the domains of *P(lacW)* including the 1 kb plasmid region that  
682 is not targeted by maternal piRNAs in G1. Small black lines represent the maternal piRNAs  
683 transmitted to the progenies produced by *P(lArB)* (**A**) or by *P(RS3)* (**F**) inserted in the *X* or *3R*  
684 subtelomeric piRNA clusters, respectively.

685

686 **Fig. 5. Cooption within piRNA cluster**

687 The 0.9 kb *T3* domain is considered to be the representative of the *1A* cluster as it is the unique  
688 sequence not redundantly found in the other piRNA clusters. In the *P-1152* strain, the *1A* cluster

689 contains a *P*-transgene. **A.** At the equilibrium, the *1A* cluster behaves as a Rhino and Moonshiner  
690 independent germline dual strand piRNA cluster, while the *P*-transgene is a Rhi-Moon dependent  
691 dual-strand piRNA cluster. **B, C.** In absence of maternal inheritance of complementary piRNAs,  
692 the *T3* sequence is converted into a piRNA cluster in all egg chambers (**B** and **C**) whereas the 18  
693 kb *P(lArB)* transgene is not converted (**B**) or converted in few egg chambers (**C**), according to a  
694 ON/OFF mechanism. NB: the drawing in not on scale.

695  
696 **Figure S1. ON/OFF lacZ egg chamber repression in MI and PI lineages.** **A.** Experimental  
697 schemes. Ovarian *lacZ* repression was assayed on the euchromatic *lacZ* target at G1, G2, G3 and  
698 G4 in both lineages by crossing females at each generation with euchromatic *lacZ* target containing  
699 males. **B.** Measure of ovarian *lacZ* repression in MI and *PI P(lArB)* lineages. Mean for the four sets  
700 of sublines and standard deviation (see **Table S4**).

701  
702 **Fig. S2. Normalized piRNA reads on the P(lArB) regions in MI and PI lineages.** Normalized 23-  
703 29 nt ovarian reads mapping to each the regions of *P(lArB)* in the *MI* and *PI* lineages were plotted  
704 on the same graph (**Fig. 1E, Table S5**). Graphs emphasize that all type of sequences, exogenous  
705 (*lacZ* (**A**) and *plasmid* (**B**)) and endogenous (*rosy* (**C**), *tCG/tHs* (**D**) and *Adh* (**E**)) of the drosophila  
706 genome, behave similarly by accumulating the capacity to produce more and more 23-29 nt small  
707 RNAs over generations. These results strongly suggest that endogenous and exogenous  
708 sequences to the *Drosophila* genome can be converted to produce 23-29 nt small RNAs after four  
709 generations. The G1, G2 and G3 correspond to the average of two independent lines. Error bars  
710 indicate standard deviation.

711 **Fig. S3. Kinetics of conversion of the 1A cluster in the w<sup>1118</sup> background.** **A.** Reciprocal crosses  
712 leading to *MI* and *PI* inherited *1A* cluster encoding the *P(lArB)* transgene and the *T3* domain. The  
713 names of each strain are indicated above the genotype. The *X* alleles are indicated as the genotype.  
714 The *w<sup>1118</sup>* strain is missing the *1A* cluster and contains the *60F* and *100F* subtelomeric clusters. **B**

715 and C. Normalized 23-29 nt small RNAs were mapped on the *P(lArB)* sequence (B) and on the *T3*  
716 sequence (C). Results on duplicates are shown for each analysis. It shows a one generation  
717 conversion for the *T3* domain for piRNA synthesis, that is not occurring for the *P(lArB)* transgene.  
718 Numbers in each panel represent the amount of normalized 23-29 nt reads identified in each  
719 condition.

720

721 **Fig. S4. Progressive conversion of the *white* gene included in the *P(RS3)* transgene inserted in**  
722 **the *100F* cluster. A.** One of the subtelomeric repeats is shown (0.98 kb). Each repeat is composed  
723 of several domains common with the *1A* and *60F* clusters (*INV-4*, *T1*, *T2* and *T4*, and specific for  
724 the autosomic subtelomeres, *T5*). Small black lines symbolize the maternal piRNAs produced by  
725 the *60F* and *100F* clusters. **B.** Structure of the *P(RS3)* transgene inserted in the *100F* (3R)  
726 subtelomere. It is encoding the *mini-white* gene (4 kb), interrupted by *FRT* sequence. The two  
727 regions are not drawn to scale. **C.** Reciprocal crosses performed to obtain *PI* and *MI P(RS3)*. “*3R-*  
728 *subt::P(RS3)*” and “*3R-subt*” stand for the 3R subtelomeric allele (*100F* cluster) with or without an  
729 insertion of *P(RS3)*, respectively. The names of each strain are indicated above the genotype. **D.**  
730 Ovarian 23-29 nt small RNAs mapping to *P(RS3)* transgene. The experiment shows that the *100F*  
731 piRNAs are unable to convert from the first generation *P(RS3)*. Numbers in each panel represent  
732 the amount of normalized 23-29 nt reads identified in each condition.

733

734 **Fig. S5. *P* derived sequences of the *P(lArB)* transgene is able to repress ovarian expression of**  
735 **a *pRFP* transgene devoid of *T3* sequence. A.** The *P(lArB)* transgenes inserted into the *1A* cluster  
736 of the *P-1152* represses the expression of the *RFP* transgene. **B.** Ovarian expression of the *RFP-T3*  
737 transgene is repressed by the *P(lArB)* and *T3* inserted in the *1A* cluster of the *P-1152* strain. **C.**  
738 Ovarian expression of *RFP* transgene is not affected with the *1A* cluster of the *Oregon* strain. **D.**  
739 Germline knockdown of Piwi silences ovarian *lacZ* expression. Parental crosses are indicated  
740 above pictures.

741  
742 **Fig S6. A.** Steady state transcription at G4 of the *P(lArB)* transgene and *T3* of the *P-1152 MI* and  
743 *PI*. RT-qPCR experiments revealed that the steady-state level of ovarian *P(lArB)* and *T3* RNA at  
744 G4 (n=4) are similar between *MI* and *PI*. Significant p-values are given (bilateral Student's t-test).  
745 ns: not significant. **B, C.** RT qPCR to control that *Rhi* (**B**, n=4) and *Moon* (**C**, n=6) are depleted in  
746 the experimental conditions. **D.** Crosses used to test the repression on the *lacZ* repression induced  
747 by *P(lArB)* inserted in the *1A* cluster of *P-1152* and *T3* repression induced the *1A* cluster of the  
748 Oregon strain.

749  
750 **Fig. S7. A.** Crosses between *P-1152* females carrying hemizygous *P(lArB)* transgenes inserted in  
751 the *1A* cluster over the *M5* balancer chromosome with *BX2* males carrying the heterozygous  
752 *P(lacW)* transgene inserted in the arm of the 2<sup>nd</sup> chromosome over the *Cy* balancer chromosome.  
753 **B.** Crosses between females carrying *P(RS3)* transgene inserted in the *100F* cluster over the *TM8*  
754 balancer chromosome with *BX2* males. In both crosses, G1 progenies devoid of either the *P(lArB)*  
755 or the *P(RS3)* were recovered and crossed to each other for several generations.

756  
757 **Fig. S8. No conversion in cis on flanking regions of the BX2 or the 1A clusters.** (A) Schematic  
758 representation of the *BX2* locus containing seven *P(lacW)* transgenes (only one is shown) inserted  
759 in tandem in the *Ago1/mRPL53* region. Broken arrows indicate the transcription direction of  
760 *mRPL53* and *Ago1*. (B) Normalized 23-29 nt reads matching the 5' and 3' regions of the *P(lacW)*  
761 transgenes insertion sites in a naïve context or G1 and G4. "Naïve" corresponds to the *BX2* locus  
762 not activated by maternal *P(lacW)* piRNAs. (C) Normalized 23-29 nt small RNAs were mapped  
763 on *CG17636* sequence, the most proximal gene on the X chromosome at G1 and G4 in the two  
764 *P(lArB)* lineages. For *P(lacW)* and *P(lArB)*, a mild impact can be observed as few piRNAs are  
765 identified, but no conversion into piRNA cluster of both regions were identified. Numbers in each  
766 panel indicate the number of normalized 23-29 nt reads.

767

768 **Fig. S9. No conversion in *trans* is induced on endogenous sequences.** (A) Normalized 23-29 nt  
769 reads identified when the *P(lacW)* transgene is converted by *P(lArB)* at G1 and G4 mapping to the  
770 endogenous region of the *white* gene highlighted by the red frames. Grey boxes highlight the *white*  
771 sequence included in the *P(lacW)* transgene, showing a large deletion of the first intron. (B, C, and  
772 D) The raw values of 23 to 29 nt small RNAs have been plotted against *D. melanogaster* genome  
773 (R5.9) under IGV (Integrative Genomes Viewer) in 6 small RNA libraries *Canton*, *P(lArB)-PI* at G1  
774 and G4 and *P(lArB)-MI* at G1. The red frames highlight the exact regions present on the transgene:  
775 terminal regions of *CG3281* and *Hsp70* genes (B), *rosy* (C) and *Adh* (D). Red and blue reads are for  
776 the sense and anti-sense reads respectively. Grey reads correspond to reads mapping on multiple  
777 sites, here on the duplicated *Hsp70Ab* and *Hsp70Aa* genes. Black stars represent the overlapping  
778 small RNA between exon and intron. In all four cases, the maternally inherited transgenes  
779 piRNAs have no or a minor effect on the piRNAs produced from the endogenous loci.

780

## 781 References

- 782 Akkouche, A., Mugat, B., Barckmann, B., Varela-Chavez, C., Li, B., Raffel, R., Péliçon, A., and  
783 Chambeyron, S. (2017). Piwi Is Required during *Drosophila* Embryogenesis to License Dual-Strand  
784 piRNA Clusters for Transposon Repression in Adult Ovaries. *Mol. Cell* 66, 411-419.e4.
- 785 Andersen, P.R., Tirian, L., Vunjak, M., and Brennecke, J. (2017). A heterochromatin-dependent  
786 transcription machinery drives piRNA expression. *Nature* 549, 54–59.
- 787 Anxolabéhère, D., Kidwell, M.G., and Periquet, G. (1988). Molecular characteristics of diverse  
788 populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile  
789 P elements. *Mol. Biol. Evol.* 5, 252–269.
- 790 Asif-Laidin, A., Delmarre, V., Laurentie, J., Miller, W.J., Ronsseray, S., and Teyssset, L. (2017). Short and  
791 long-term evolutionary dynamics of subtelomeric piRNA clusters in *Drosophila*. *DNA Res.* 24, 459–472.
- 792 Bergman, C.M., Quesneville, H., Anxolabéhère, D., and Ashburner, M. (2006). Recurrent insertion and  
793 duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome.  
794 *Genome Biol.* 7, R112.

795 Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007).  
796 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* *128*,  
797 1089–1103.

798 Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An  
799 epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* *322*, 1387–1392.

800 Casier, K., Delmarre, V., Gueguen, N., Hermant, C., Viode, E., Vaury, C., Ronsseray, S., Brasset, E.,  
801 Teyssset, L., and Boivin, A. (2019). Environmentally-induced epigenetic conversion of a piRNA cluster.  
802 *Elife* *8*.

803 Chen, Y.-C.A., Stuwe, E., Luo, Y., Ninova, M., Le Thomas, A., Rozhavskaia, E., Li, S., Vempati, S.,  
804 Laver, J.D., Patel, D.J., et al. (2016). Cutoff Suppresses RNA Polymerase II Termination to Ensure  
805 Expression of piRNA Precursors. *Mol. Cell* *63*, 97–109.

806 Czech, B., and Hannon, G.J. (2016). One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-  
807 Guided Silencing. *Trends Biochem. Sci.* *41*, 324–337.

808 Dorer, D.R., and Henikoff, S. (1994). Expansions of transgene repeats cause heterochromatin formation  
809 and gene silencing in *Drosophila*. *Cell* *77*, 993–1002.

810 Duc, C., Yoth, M., Jensen, S., Mounié, N., Bergman, C.M., Vaury, C., and Brasset, E. (2019). Trapping a  
811 somatic endogenous retrovirus into a germline piRNA cluster immunizes the germline against further  
812 invasion. *Genome Biol.* *20*, 127.

813 Dufourt, J., Dennis, C., Boivin, A., Gueguen, N., Théron, E., Goriaux, C., Pouchin, P., Ronsseray, S.,  
814 Brasset, E., and Vaury, C. (2014). Spatio-temporal requirements for transposable element piRNA-  
815 mediated silencing during *Drosophila* oogenesis. *Nucleic Acids Res.* *42*, 2512–2524.

816 ElMaghraby, M.F., Andersen, P.R., Pühringer, F., Hohmann, U., Meixner, K., Lendl, T., Tirian, L., and  
817 Brennecke, J. (2019). A Heterochromatin-Specific RNA Export Pathway Facilitates piRNA Production.  
818 *Cell* *178*, 964-979.e20.

819 Gao, C., Xiao, M., Ren, X., Hayward, A., Yin, J., Wu, L., Fu, D., and Li, J. (2012). Characterization and  
820 functional annotation of nested transposable elements in eukaryotic genomes. *Genomics* *100*, 222–230.

821 Gilbert, C., and Feschotte, C. (2018). Horizontal acquisition of transposable elements and viral sequences:  
822 patterns and consequences. *Curr. Opin. Genet. Dev.* *49*, 15–24.

823 Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and  
824 Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in  
825 *Drosophila*. *Science* *315*, 1587–1590.

826 Hermant, C., Boivin, A., Teyssset, L., Delmarre, V., Asif-Laidin, A., van den Beek, M., Antoniewski, C.,  
827 and Ronsseray, S. (2015). Paramutation in *Drosophila* Requires Both Nuclear and Cytoplasmic Actors of  
828 the piRNA Pathway and Induces Cis-spreading of piRNA Production. *Genetics* *201*, 1381–1396.

829 Hill, T., and Betancourt, A.J. (2018). Extensive exchange of transposable elements in the *Drosophila*  
830 *pseudoobscura* group. *Mob DNA* *9*, 20.

- 831 Jangam, D., Feschotte, C., and Betrán, E. (2017). Transposable Element Domestication As an Adaptation  
832 to Evolutionary Conflicts. *Trends Genet.* *33*, 817–831.
- 833 Josse, T., Teyssset, L., Todeschini, A.-L., Sidor, C.M., Anxolabéhère, D., and Ronsseray, S. (2007).  
834 Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin  
835 formation. *PLoS Genet.* *3*, 1633–1643.
- 836 Josse, T., Maurel-Zaffran, C., de Vanssay, A., Teyssset, L., Todeschini, A.-L., Delmarre, V., Chaminade,  
837 N., Anxolabéhère, D., and Ronsseray, S. (2008). Telomeric Trans-Silencing in *Drosophila melanogaster*:  
838 Tissue Specificity, Development and Functional Interactions between Non-Homologous Telomeres. *PLoS*  
839 *ONE* *3*, e3249.
- 840 Kelleher, E.S., Azevedo, R.B.R., and Zheng, Y. (2018). The Evolution of Small-RNA-Mediated Silencing  
841 of an Invading Transposable Element. *Genome Biol Evol* *10*, 3038–3057.
- 842 Kent, T.V., Uzunović, J., and Wright, S.I. (2017). Coevolution between transposable elements and  
843 recombination. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* *372*.
- 844 Kidwell, M.G. (1983). Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc.*  
845 *Natl. Acad. Sci. U.S.A.* *80*, 1655–1659.
- 846 Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S.,  
847 Nowosielska, A., et al. (2009). The *Drosophila* HP1 homolog Rhino is required for transposon silencing  
848 and piRNA production by dual-strand clusters. *Cell* *138*, 1137–1149.
- 849 Kofler, R. (2019). Dynamics of Transposable Element Invasions with piRNA Clusters. *Mol. Biol. Evol.*  
850 *36*, 1457–1472.
- 851 Kofler, R., Hill, T., Nolte, V., Betancourt, A.J., and Schlötterer, C. (2015). The recent invasion of natural  
852 *Drosophila simulans* populations by the P-element. *Proc. Natl. Acad. Sci. U.S.A.* *112*, 6659–6663.
- 853 Kofler, R., Senti, K.-A., Nolte, V., Tobler, R., and Schlötterer, C. (2018). Molecular dissection of a natural  
854 transposable element invasion. *Genome Res.* *28*, 824–835.
- 855 Komarov, P.A., Sokolova, O., Akulenko, N., Brassat, E., Jensen, S., and Kalmykova, A. (2020).  
856 Epigenetic Requirements for Triggering Heterochromatinization and Piwi-Interacting RNA Production  
857 from Transgenes in the *Drosophila* Germline. *Cells* *9*.
- 858 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment  
859 of short DNA sequences to the human genome. *Genome Biol.* *10*, R25.
- 860 Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin,  
861 A.A., and Tóth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a  
862 repressive chromatin state. *Genes Dev.* *27*, 390–399.
- 863 Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., Chen, Y.-C.A., Luo, Y.,  
864 Sachidanandam, R., Toth, K.F., et al. (2014). Transgenerationally inherited piRNAs trigger piRNA  
865 biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev.*  
866 *28*, 1667–1680.

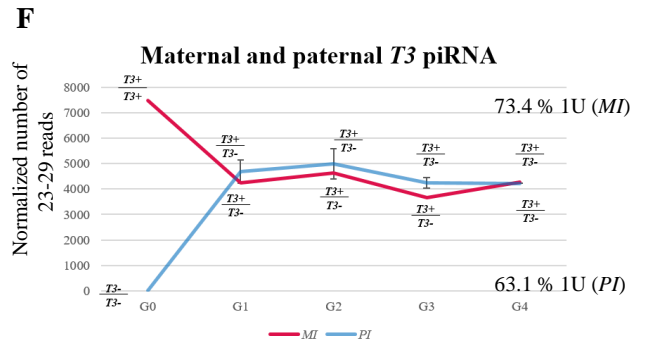
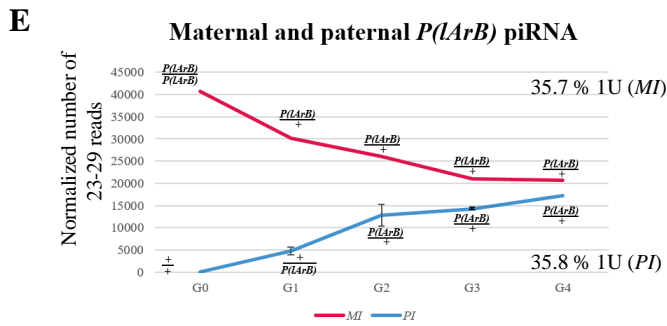
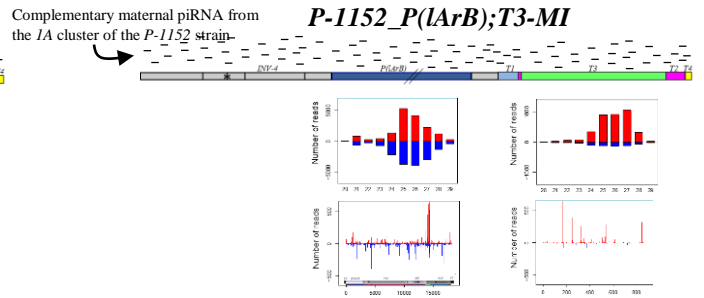
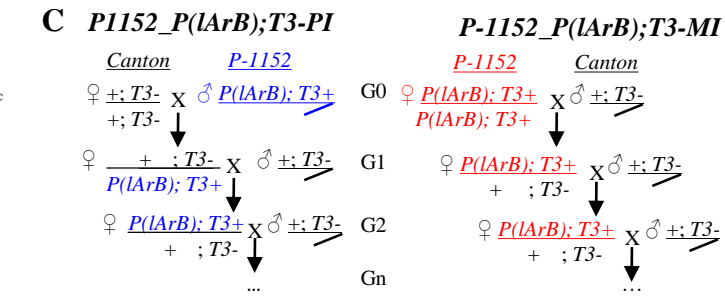
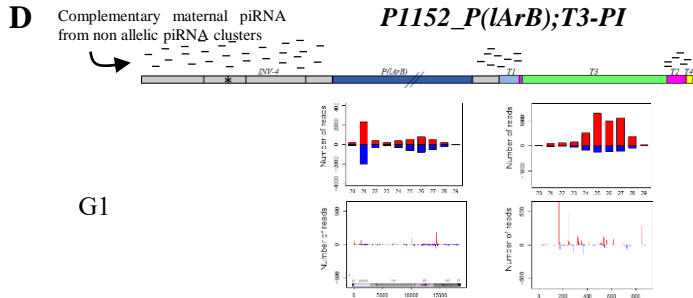
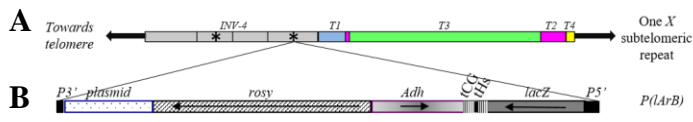
- 867 Loreto, E.L.S., Carareto, C.M.A., and Capy, P. (2008). Revisiting horizontal transfer of transposable  
868 elements in *Drosophila*. *Heredity (Edinb)* *100*, 545–554.
- 869 Marie, P.P., Ronsseray, S., and Boivin, A. (2017). From Embryo to Adult: piRNA-Mediated Silencing  
870 throughout Germline Development in *Drosophila*. *G3 (Bethesda)* *7*, 505–516.
- 871 Miller, W.J., McDonald, J.F., Nouaud, D., and Anxolabéhère, D. (1999). Molecular domestication--more  
872 than a sporadic episode in evolution. *Genetica* *107*, 197–207.
- 873 Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The rhino-deadlock-cutoff complex licenses  
874 noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* *157*, 1364–1379.
- 875 Muerdter, F., Olovnikov, I., Molaro, A., Rozhkov, N.V., Czech, B., Gordon, A., Hannon, G.J., and  
876 Aravin, A.A. (2012). Production of artificial piRNAs in flies and mice. *RNA* *18*, 42–52.
- 877 O’Kane, C.J., and Gehring, W.J. (1987). Detection in situ of genomic regulatory elements in *Drosophila*.  
878 *Proc. Natl. Acad. Sci. U.S.A.* *84*, 9123–9127.
- 879 Olovnikov, I., Ryazansky, S., Shpiz, S., Lavrov, S., Abramov, Y., Vaury, C., Jensen, S., and Kalmykova,  
880 A. (2013). De novo piRNA cluster formation in the *Drosophila* germ line triggered by transgenes  
881 containing a transcribed transposon fragment. *Nucleic Acids Res.* *41*, 5757–5768.
- 882 Poyhonen, M., de Vanssay, A., Delmarre, V., Hermant, C., Todeschini, A.L., Teyssset, L., and Ronsseray,  
883 S. (2012). Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3*  
884 (Bethesda) *2*, 331–338.
- 885 Rangan, P., Malone, C.D., Navarro, C., Newbold, S.P., Hayes, P.S., Sachidanandam, R., Hannon, G.J.,  
886 and Lehmann, R. (2011). piRNA production requires heterochromatin formation in *Drosophila*. *Curr.*  
887 *Biol.* *21*, 1373–1379.
- 888 Reiss, D., Josse, T., Anxolabehere, D., and Ronsseray, S. (2004). aubergine mutations in *Drosophila*  
889 *melanogaster* impair P cytotype determination by telomeric P elements inserted in heterochromatin. *Mol*  
890 *Genet Genomics* *272*, 336–343.
- 891 Rizzon, C., Marais, G., Gouy, M., and Biémont, C. (2002). Recombination Rate and the Distribution of  
892 Transposable Elements in the *Drosophila melanogaster* Genome. *Genome Res.* *12*, 400–407.
- 893 Roche, S.E., and Rio, D.C. (1998). Trans-silencing by P elements inserted in subtelomeric  
894 heterochromatin involves the *Drosophila* Polycomb group gene, Enhancer of zeste. *Genetics* *149*, 1839–  
895 1855.
- 896 Ronsseray, S., Lehmann, M., and Anxolabéhère, D. (1989). Copy number and distribution of P and I  
897 mobile elements in *Drosophila melanogaster* populations. *Chromosoma* *98*, 207–214.
- 898 Ronsseray, S., Marin, L., Lehmann, M., and Anxolabéhère, D. (1998). Repression of hybrid dysgenesis in  
899 *Drosophila melanogaster* by combinations of telomeric P-element reporters and naturally occurring P  
900 elements. *Genetics* *149*, 1857–1866.
- 901 Rørth, P. (1998). Gal4 in the *Drosophila* female germline. *Mech. Dev.* *78*, 113–118.



- 902 Rozhkov, N.V., Hammell, M., and Hannon, G.J. (2013). Multiple roles for Piwi in silencing *Drosophila*  
903 transposons. *Genes Dev.* 27, 400–412.
- 904 Sato, K., and Siomi, M.C. (2020). The piRNA pathway in *Drosophila* ovarian germ and somatic cells.  
905 *Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci.* 96, 32–42.
- 906 Shpiz, S., Ryazansky, S., Olovnikov, I., Abramov, Y., and Kalmykova, A. (2014). Euchromatic  
907 transposon insertions trigger production of novel Pi- and endo-siRNAs at the target sites in the *drosophila*  
908 germline. *PLoS Genet.* 10, e1004138.
- 909 Srivastav, S.P., Rahman, R., Ma, Q., Pierre, J., Bandyopadhyay, S., and Lau, N.C. (2019). Har-P, a short  
910 P-element variant, weaponizes P-transposase to severely impair *Drosophila* development. *Elife* 8.
- 911 Todeschini, A.-L., Teysset, L., Delmarre, V., and Ronsseray, S. (2010). The Epigenetic Trans-Silencing  
912 Effect in *Drosophila* Involves Maternally-Transmitted Small RNAs Whose Production Depends on the  
913 piRNA Pathway and HP1. *PLoS ONE* 5, e11032.
- 914 de Vanssay, A., Bougé, A.-L., Boivin, A., Hermant, C., Teysset, L., Delmarre, V., Antoniewski, C., and  
915 Ronsseray, S. (2012). Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus.  
916 *Nature* 490, 112–115.
- 917 Zanni, V., Eymery, A., Coiffet, M., Zytnicki, M., Luyten, I., Quesneville, H., Vaury, C., and Jensen, S.  
918 (2013). Distribution, evolution, and diversity of retrotransposons at the flamenco locus reflect the  
919 regulatory properties of piRNA clusters. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19842–19847.
- 920 Zhang, Z., Wang, J., Schultz, N., Zhang, F., Parhad, S.S., Tu, S., Vreven, T., Zamore, P.D., Weng, Z., and  
921 Theurkauf, W.E. (2014). The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA  
922 precursor splicing. *Cell* 157, 1353–1363.

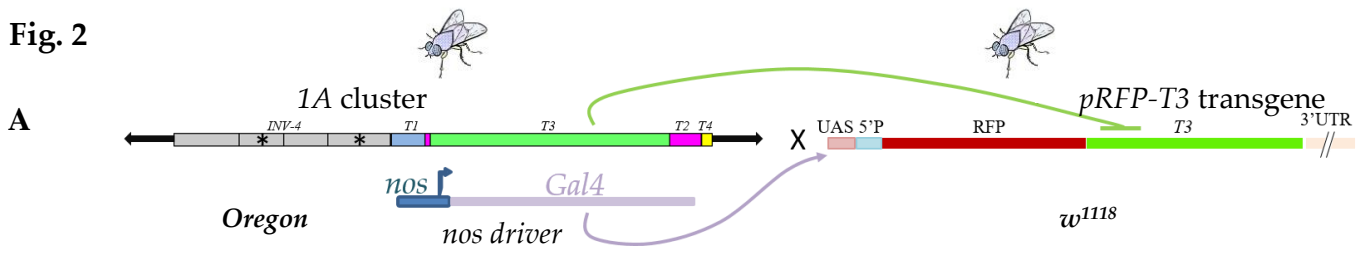
923

**Fig. 1**



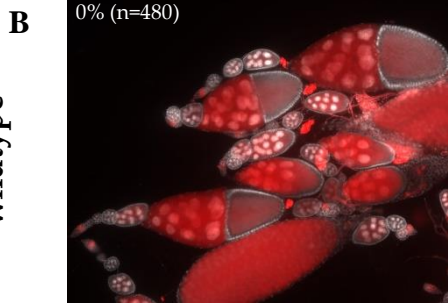
Asif-Laidin et al.,

**Fig. 2**



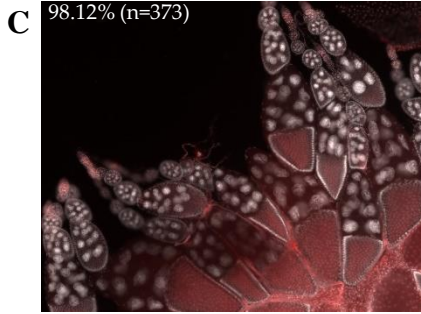
**No 1A cluster**

♀ *nos* x ♂ *pRFP-T3*



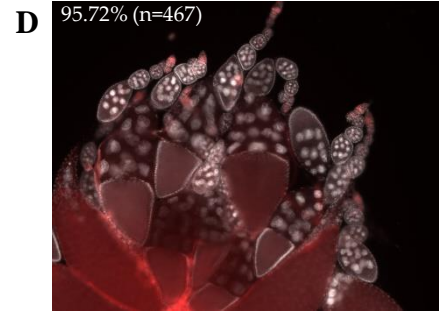
**1A cluster-MI**

♀ 1A cluster; *nos* x ♂ *pRFP-T3*

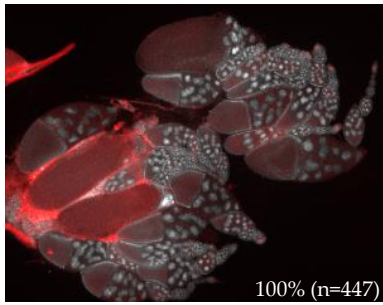


**1A cluster-PI**

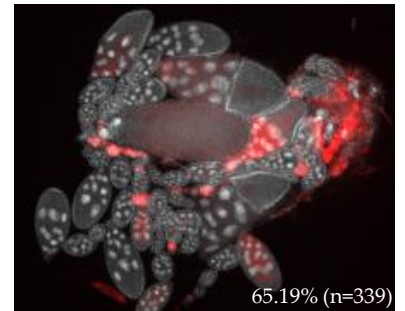
♀ *pRFP-T3* x ♂ 1A cluster; *nos*



**E** ♀ 1A cluster; *nos*; *pRFP-T3* x ♂ white KD

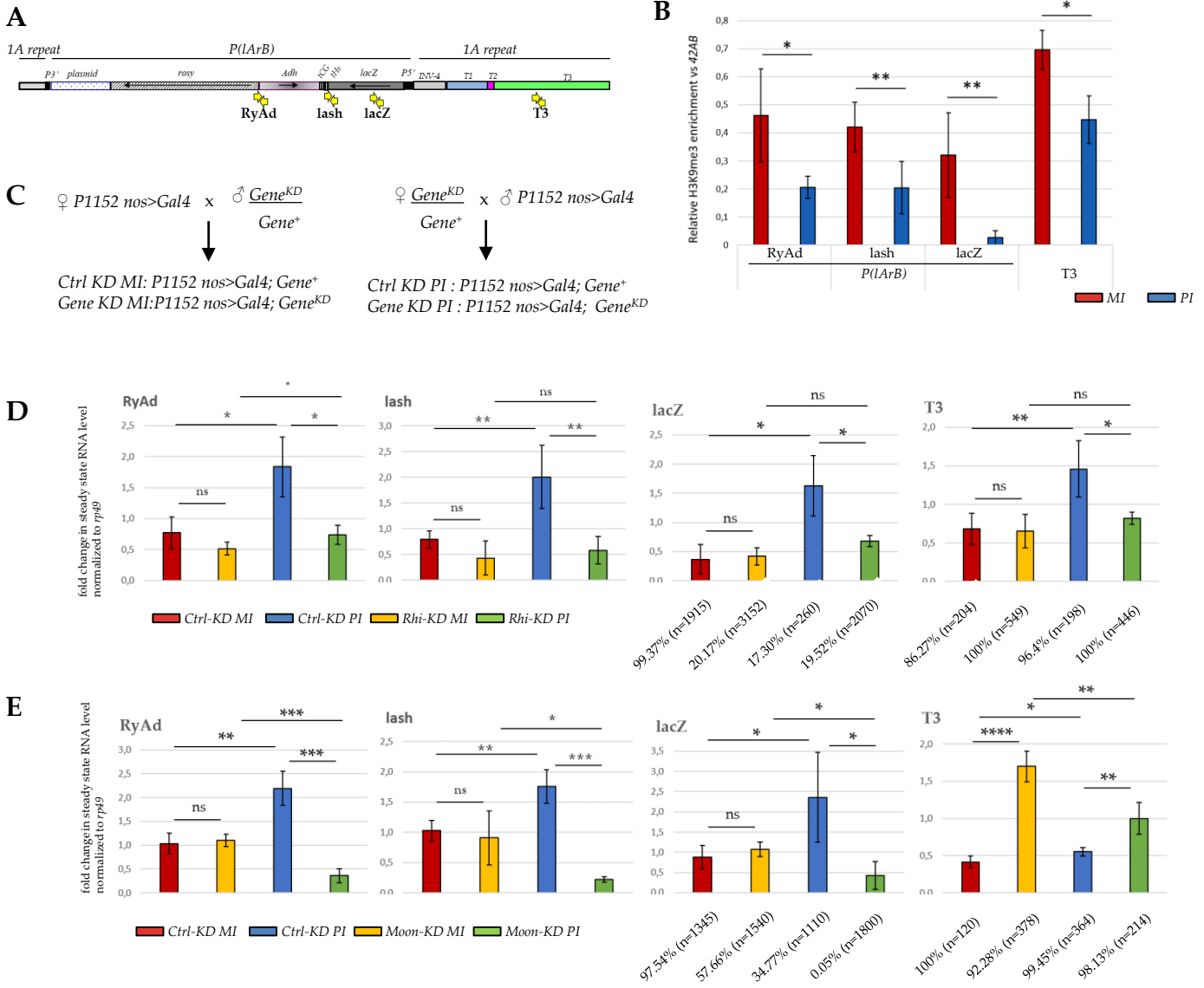


**F** ♀ 1A cluster; *nos*; *pRFP-T3* x ♂ piwi KD



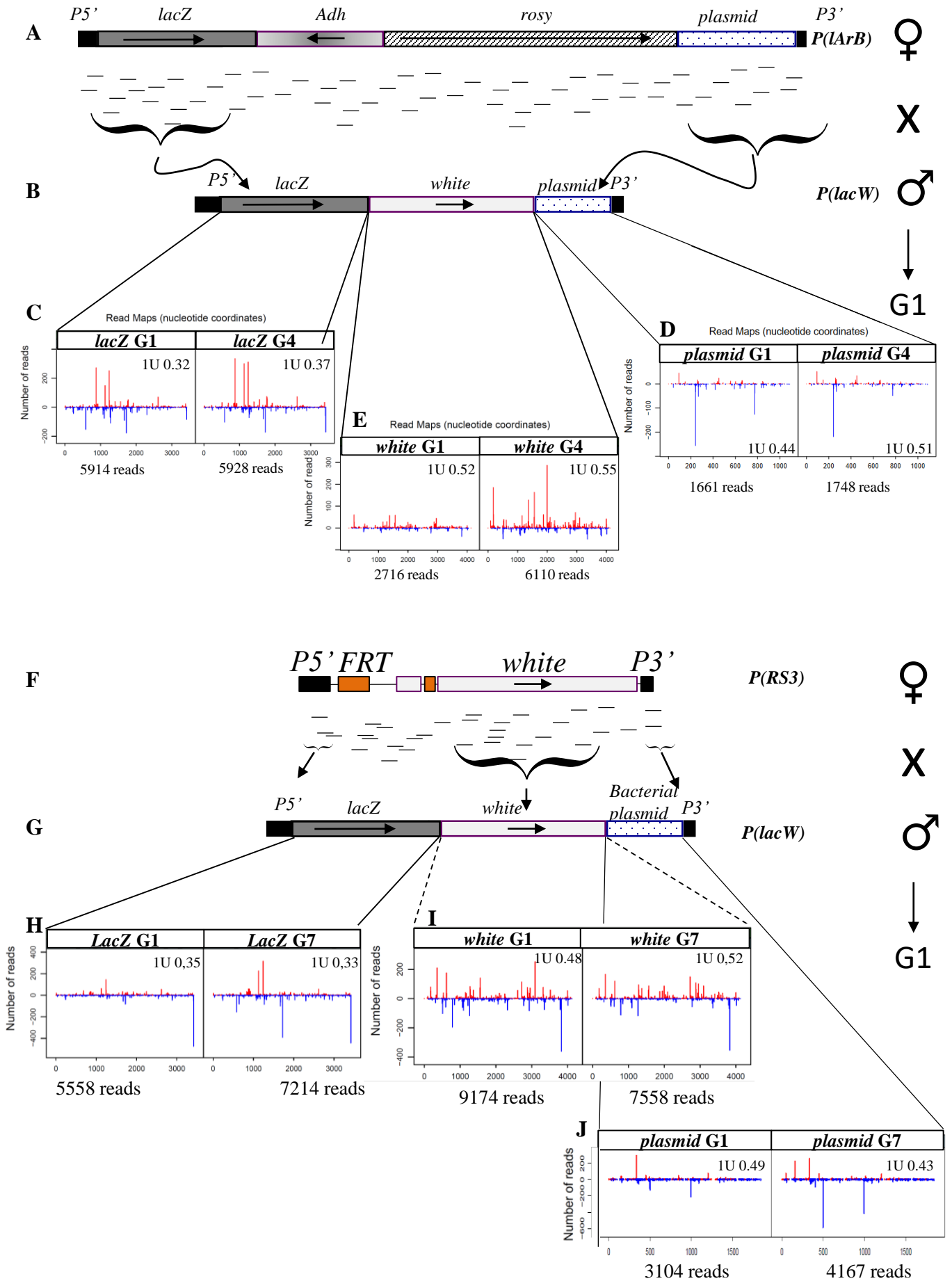
Asif-Laidin et al.,

**Fig. 3**



Asif-Laidin et al.,

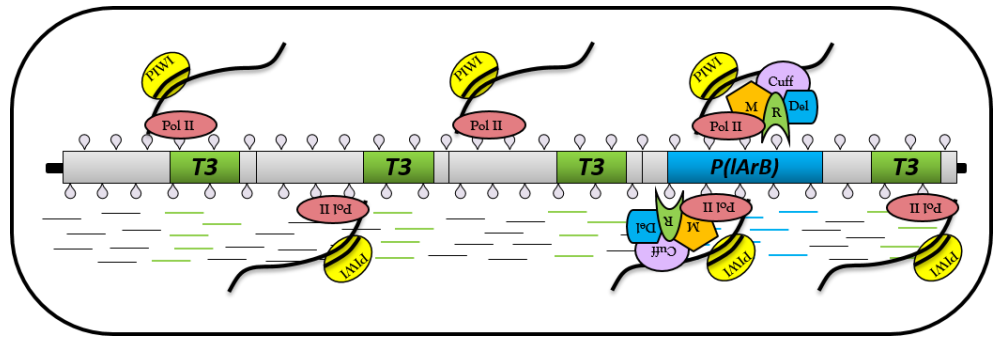
**Fig. 4**



**Fig. 5**

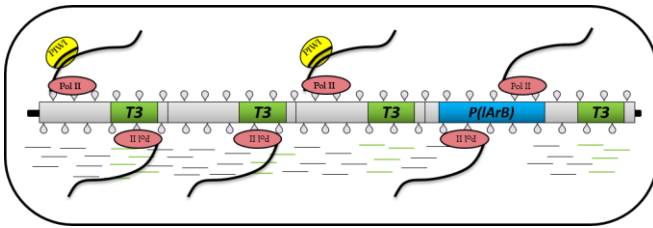
A. Maintenance of activated *1A* cluster containing a *P(IArB)* by maternally inherited piRNAs

All egg chambers



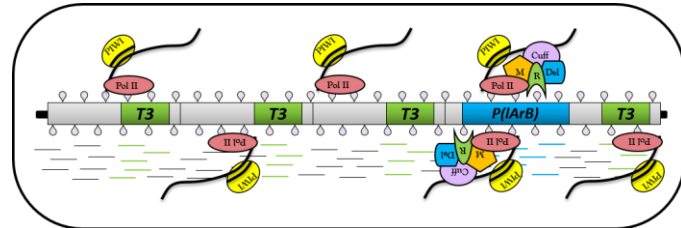
In absence of maternally inherited piRNAs

B. *T3* is converted, but not *P(IArB)*

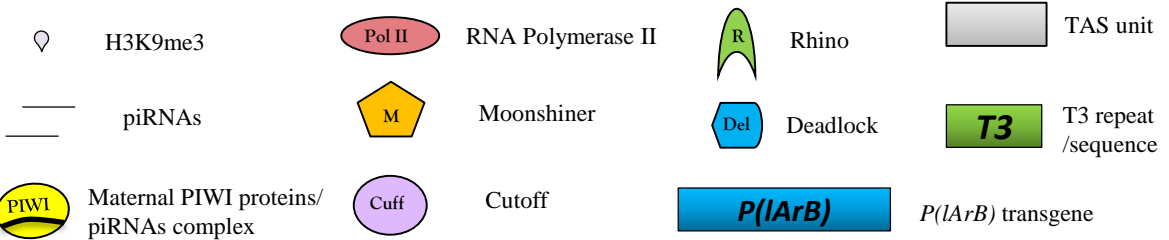


Majority of egg chambers

C. *T3* and *P*-transgene are converted



Few egg chambers



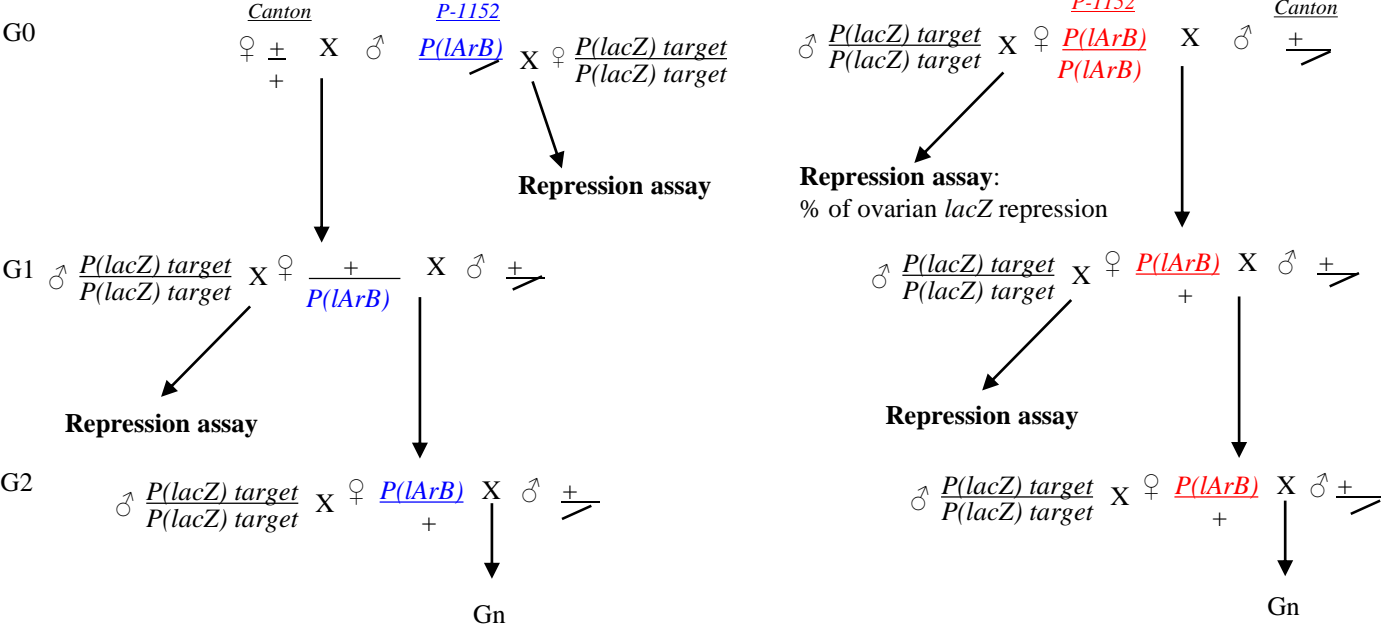
Asif-Laidin et al.

**Fig. S1**

**A**

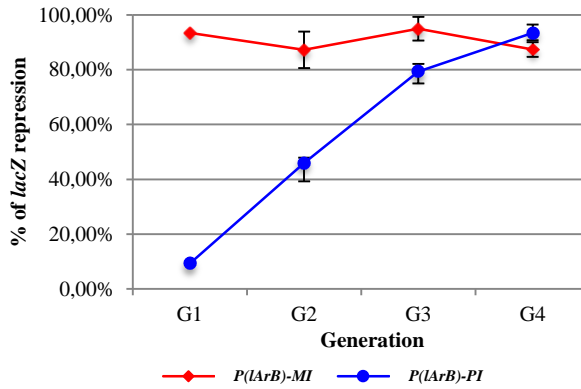
*P-1152\_P(lArB)-PI*

*P-1152\_P(lArB)-MI*



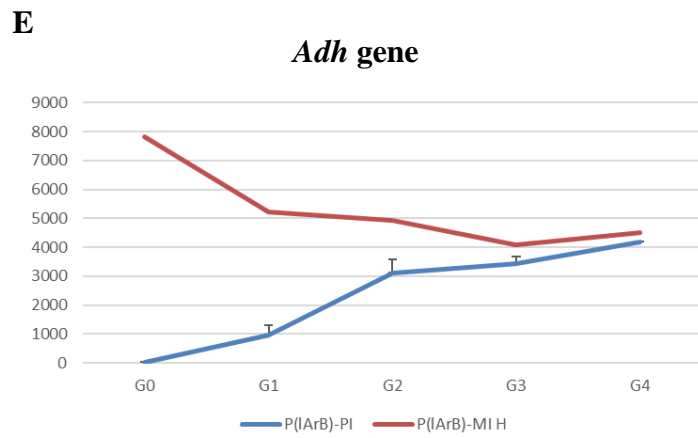
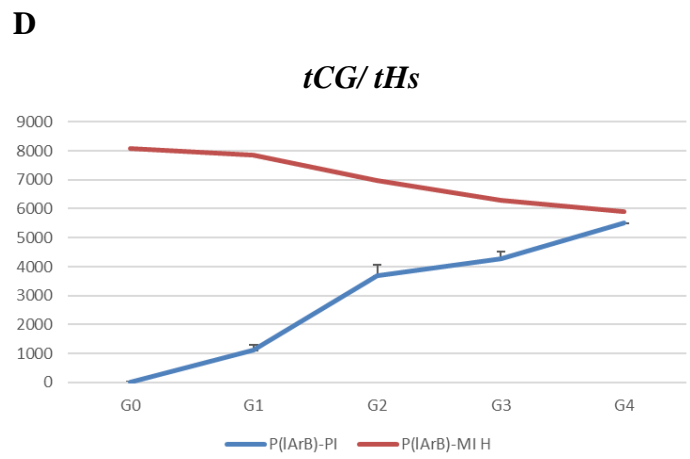
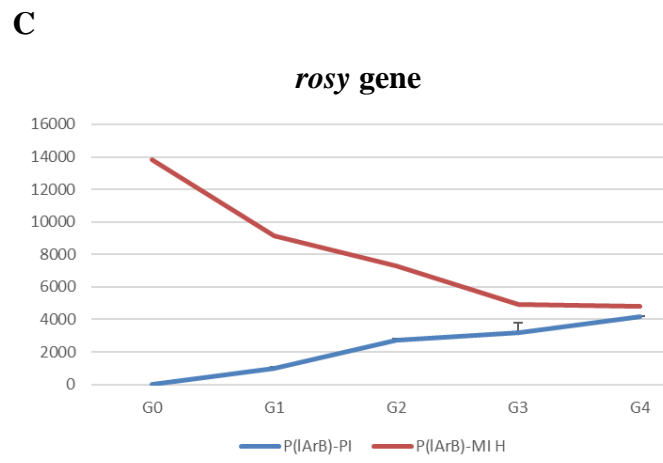
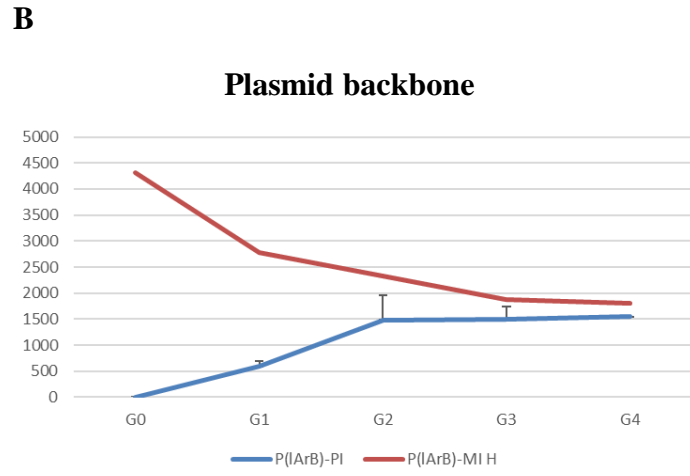
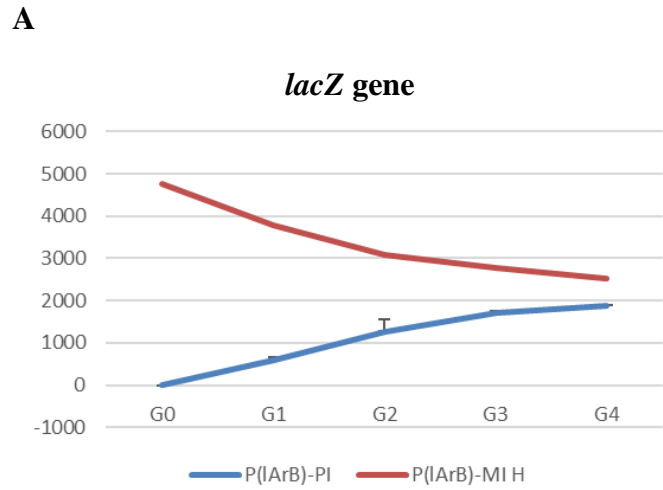
**B**

**Repression assay measured in *P(lArB)* maternal and paternal lineages**



Asif-Laidin et al.,

**Fig. S2**



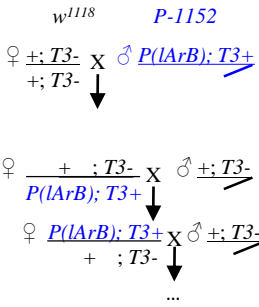
Asif-Laidin et al.,



**Fig. S3**

**A**

*P1152\_P(lArB);T3-PI*



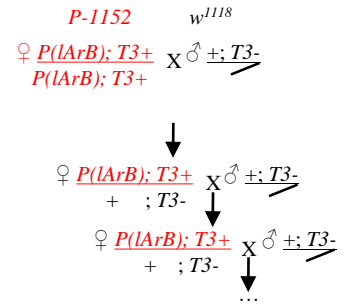
G0

G1

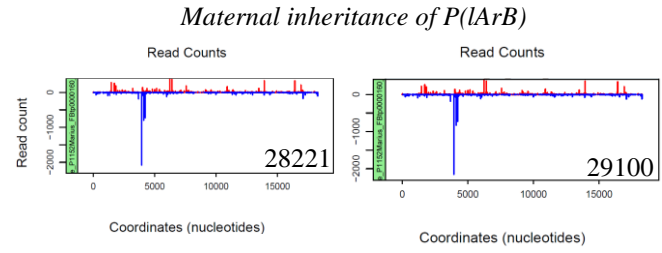
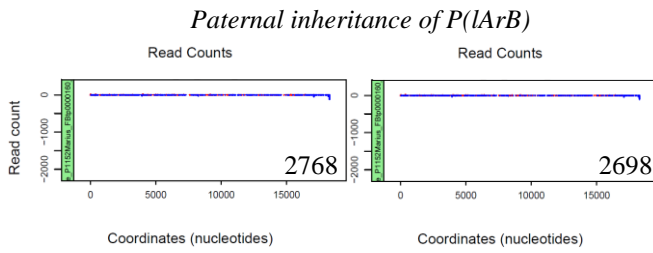
G2

Gn

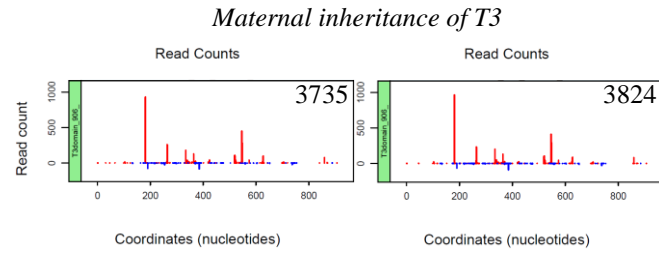
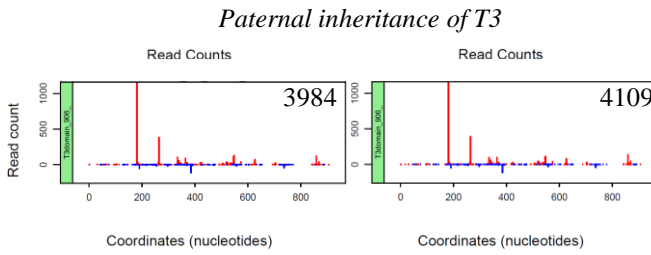
*P-1152\_P(lArB);T3-MI*



**B**

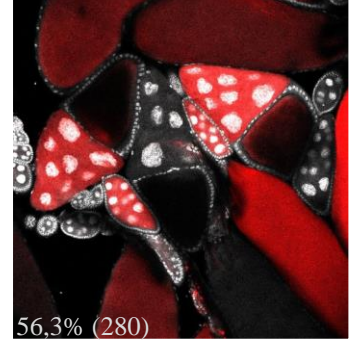
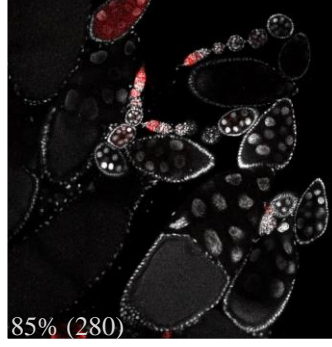
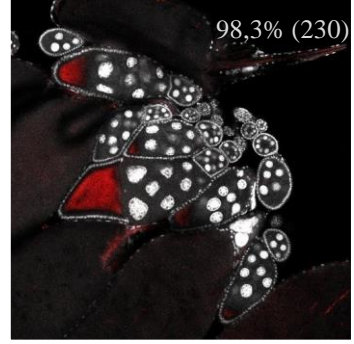
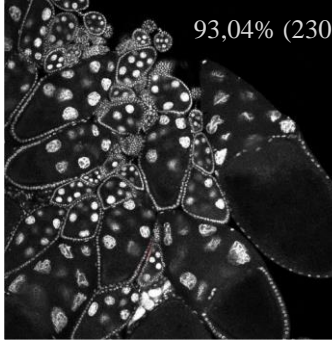
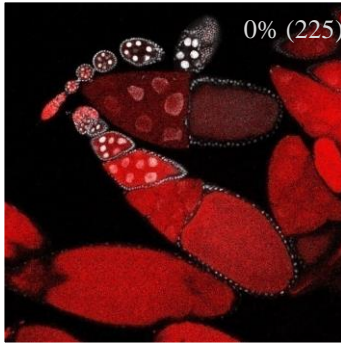
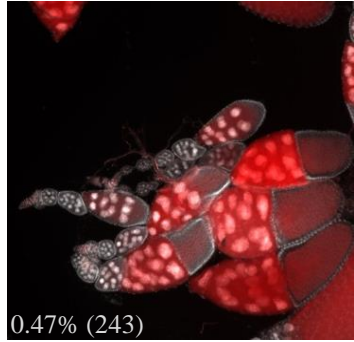
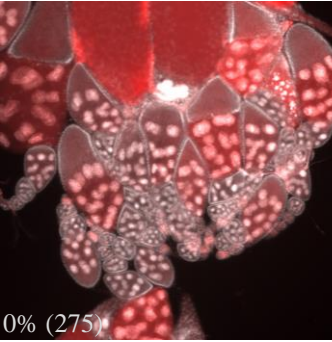
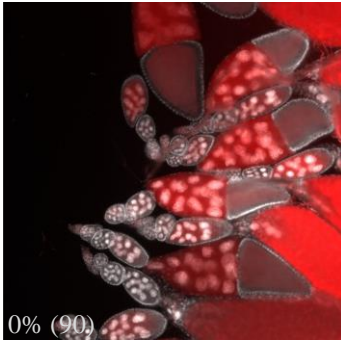
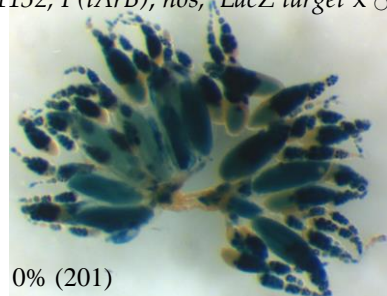
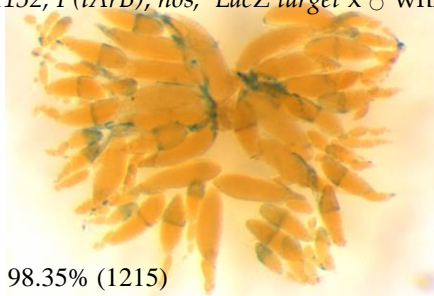


**C**



Asif-Laidin et al.,

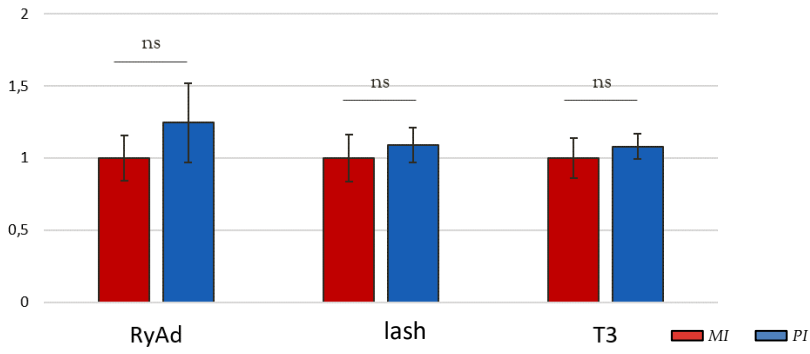


**Fig. S5****A****No 1A cluster****1A cluster-MI****1A cluster-PI**♀ *nos* × ♂ *pRFP*♀ *P-1152; nos* × ♂ *pRFP*♀ *pRFP* × ♂ *P-1152; nos**P-1152\_P(lArB);T3<***B**♀ *nos* × ♂ *pRFP-T3*♀ *P-1152; nos* × ♂ *pRFP-T3*♀ *pRFP-T3* × ♂ *P-1152; nos**P-1152\_P(lArB);T3***C**♀ *nos* × ♂ *pRFP*♀ *Oregon; nos* × ♂ *pRFP*♀ *pRFP* × ♂ *Oregon; nos**Oregon;T3***D**♀ *P-1152; P(lArB); nos; LacZ target* × ♂ *white KD*♀ *P-1152; P(lArB); nos; LacZ target* × ♂ *piwi KD**P-1152\_P(lArB);T3*

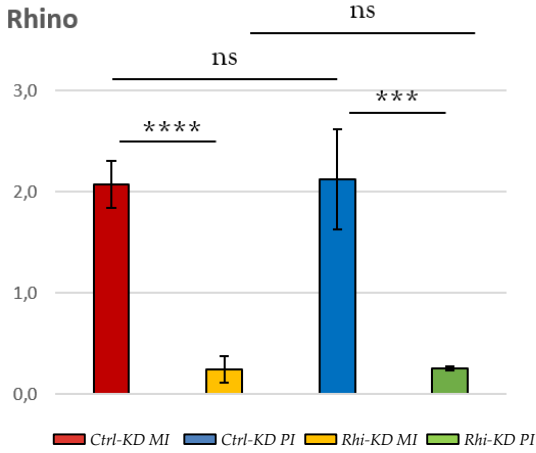
**Fig. S6**

**G4**

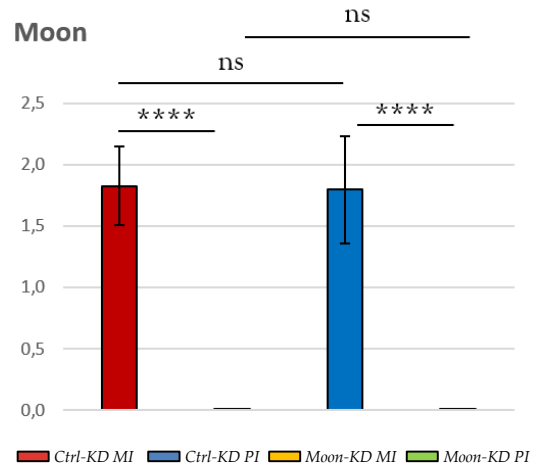
**A**



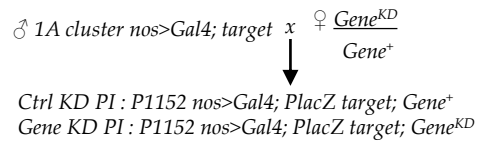
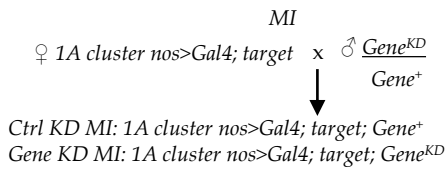
**B**



**C**

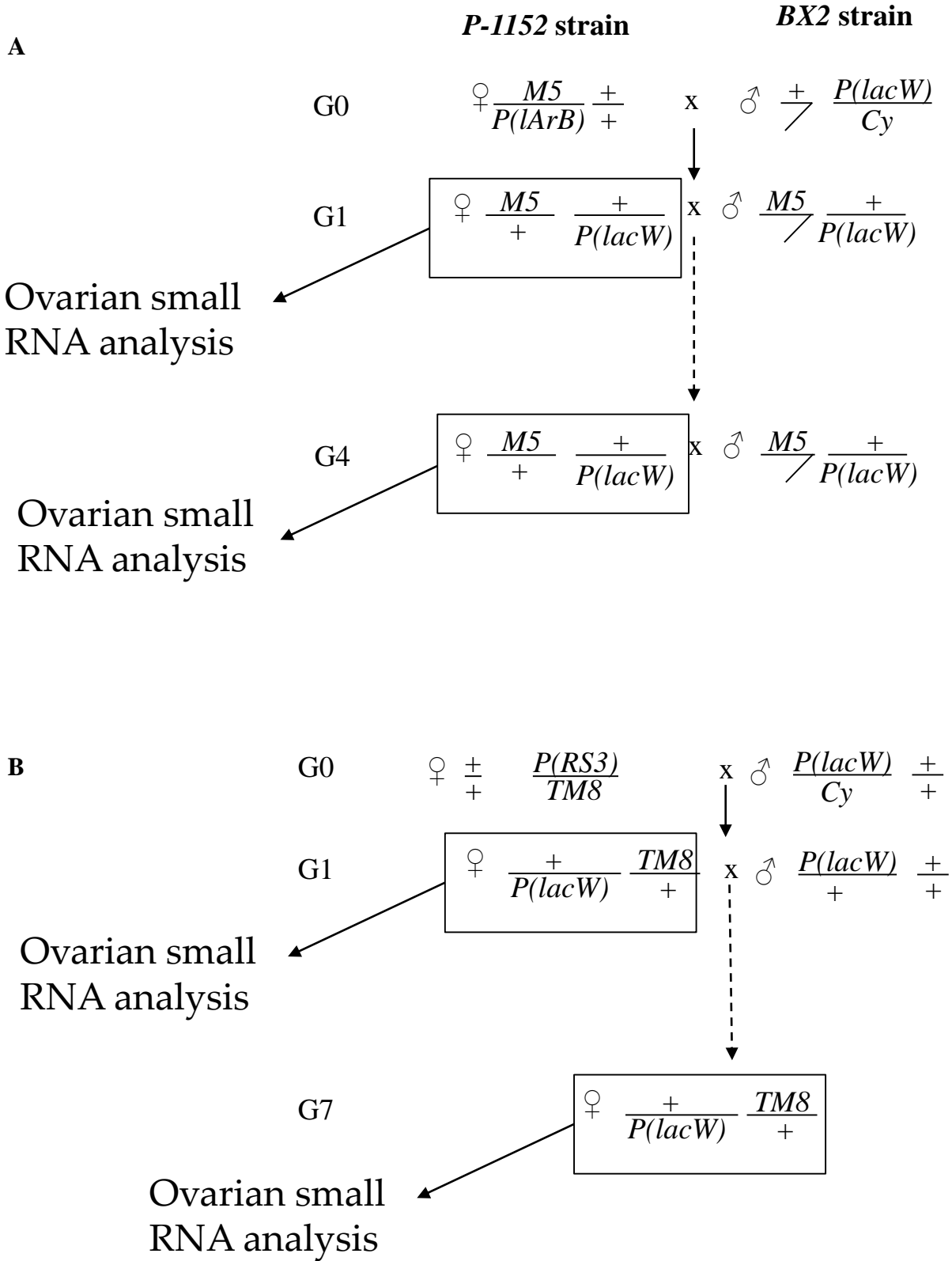


**D**



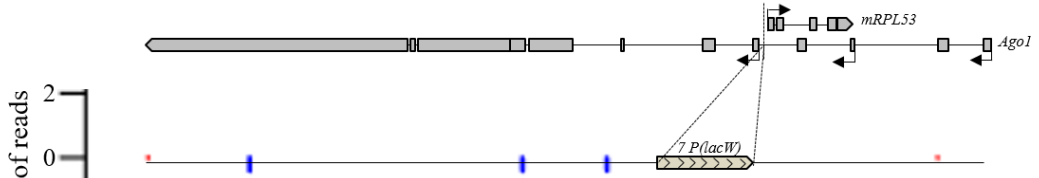
Asif-Laidin et al.,

Fig. S7

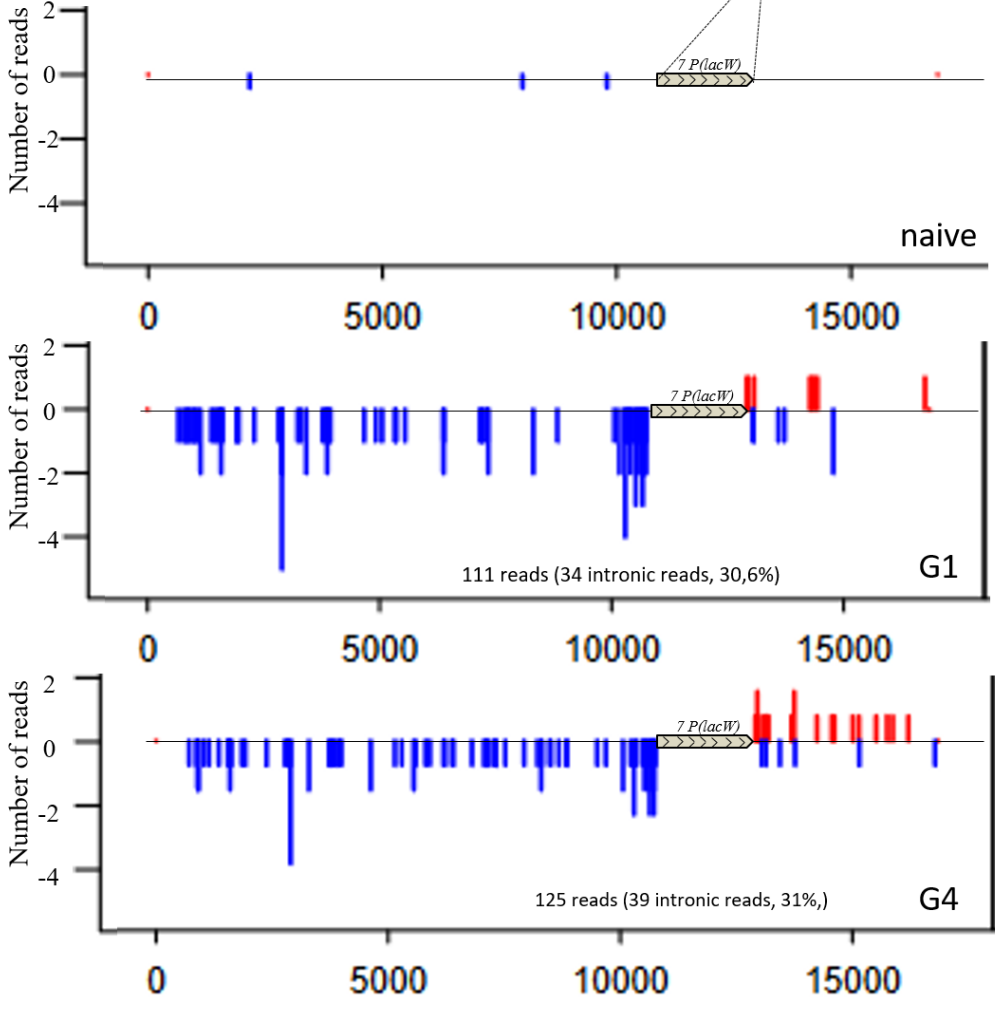


**Fig. S8**

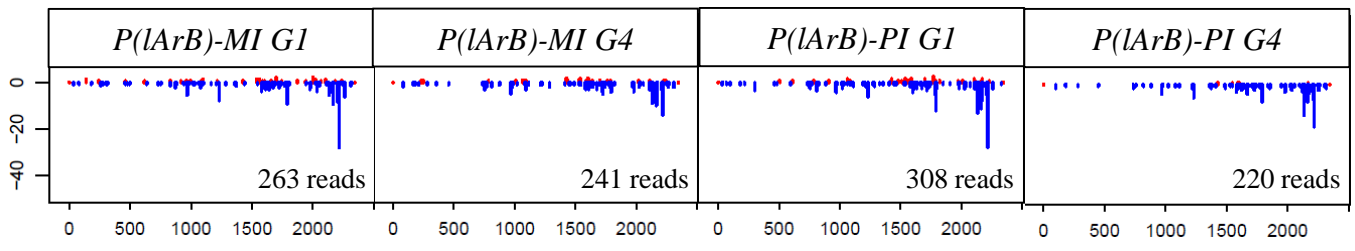
**A**



**B**



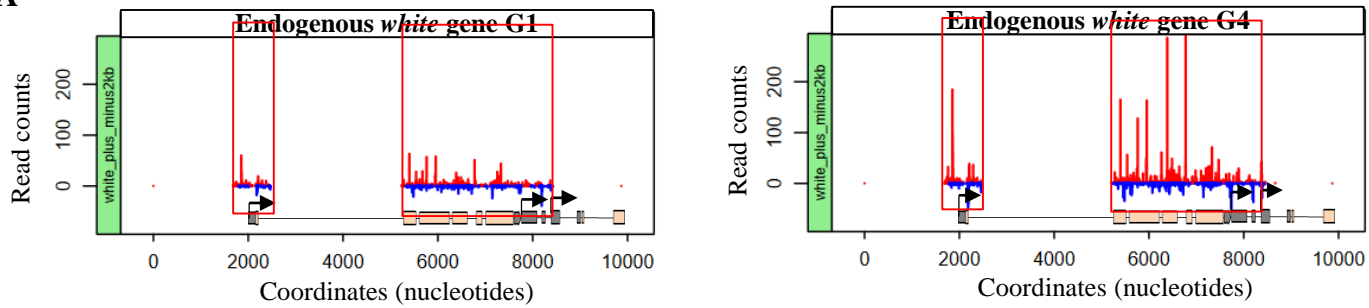
**C**



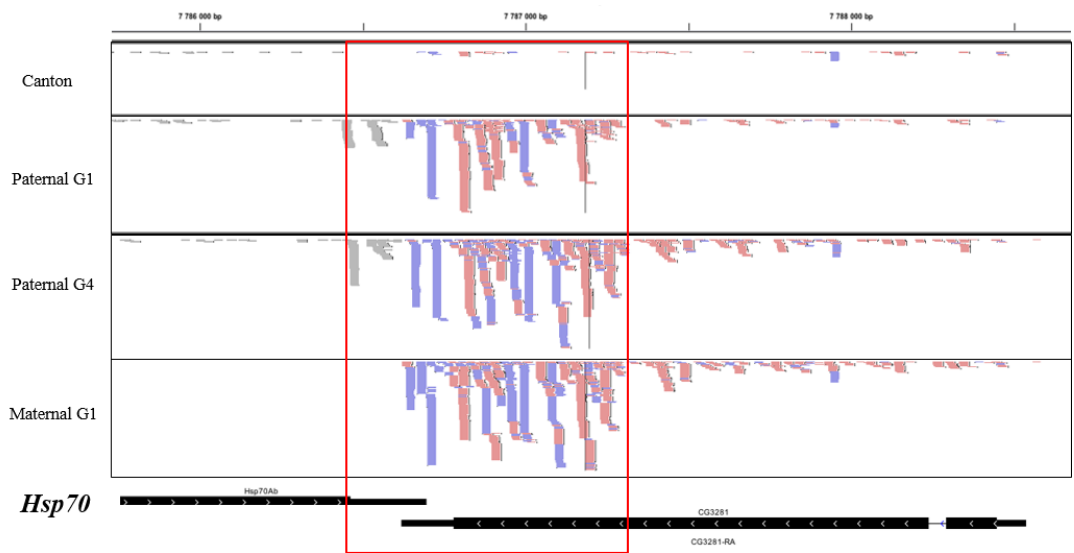
Asif-Laidin et al.

**Fig. S9**

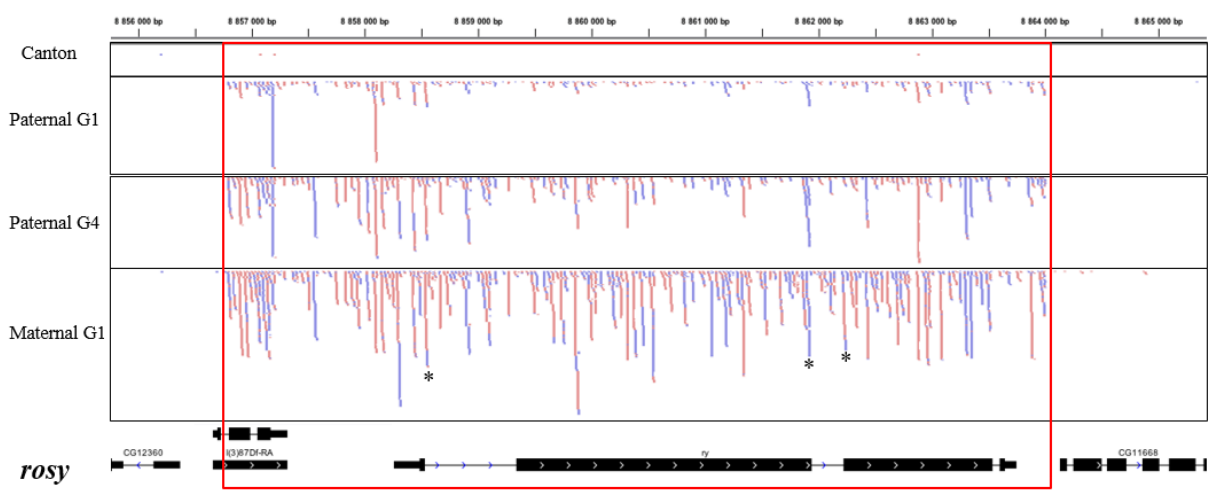
**A**



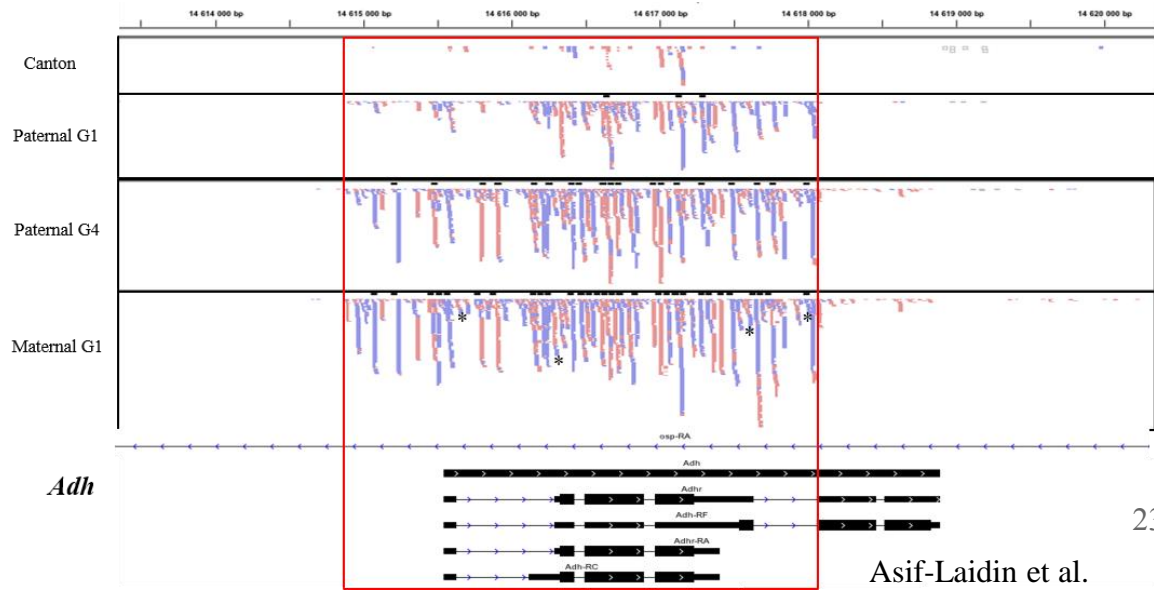
**B**



**C**



**D**



**Table S1. List of *Drosophila* strains used in this study**

Strains	Contains the 1A piRNA cluster (including T3 sequence)	Contains the 60F and/or 100F piRNA cluster	Transgene	Localization of transgene/references
<i>P-1152</i>	Yes	Yes	2 <i>P(lArB)</i>	1A piRNA cluster, (Josse et al., 2007)
<i>Canton</i>	No	Yes	-	-
<i>w<sup>1118</sup></i>	No	Yes	-	-
<i>Oregon<sup>RC</sup></i>	Yes	Yes	-	-
<i>BX2</i>	No	Yes	7 <i>P(lacW)</i>	Euchromatic (50C) (de Vanssay et al., 2012)
PRS3	No	Yes	<i>P(RS3)</i>	100F piRNA cluster (Dufour et al., 2014)
nosGal4	No	ND <sup>a</sup>	PBac{w[+mW.hs]=GreenEye.nosGAL4}	Third chromosome (Marie et al., 2017)
pRFP	No	ND <sup>a</sup>	pUASp-RFP	Third chromosome (this study)
pRFP-T3	No	ND <sup>a</sup>	<i>pUASp-RFP</i> transcriptionally fused to T3 sequence	Third chromosome (this study)
<i>BQ16</i>	No	ND <sup>a</sup>	<i>P(92)</i> encoding an euchromatic <i>lacZ</i> target	Third chromosome (64C)
<i>Piwi KD</i>	No	ND <sup>a</sup>	<i>Stock 22235</i>	Third chromosome
<i>Rhino KD</i>	No	ND <sup>a</sup>	<i>Stock 35171</i>	Third chromosome
<i>moonKD</i>	No	ND <sup>a</sup>	<i>PA61</i>	Third chromosome (Andersen et al., 2017)

<sup>a</sup>: Not Determined. These transgenic lines were obtained in a *w<sup>1118</sup>* genetic background, it is therefore assumed that they possess autosomic subtelomeres piRNA clusters.



**Table S2: Annotation of small RNA libraries**

Small RNA were prepared from ovaries of females of the indicated strains. Values for the different categories of sequences matching to the *D. melanogaster* genome 5.49 in each library are indicated. For comparisons, libraries were normalized (normalization factor) relatively to the number of sequence reads aligning to the *D. melanogaster* genome but not to miscRNAs (including rRNA and snoRNA) or tRNAs (effective depth) (Asif-Laidin et al., 2017).

<b>A. Libraries</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	
<b>alignment reference</b>	GRH117	GRH118	GRH119	GRH120	GRH121	GRH122	GRH123	GRH124	GRH126	GRH127	GRH128	GRH129	GRH133	
<b>Subline</b>	<b>Maternal H</b>				<b>Paternal</b>									
<b>Generations</b>	G1	G2	G3	G4	B G1	D G1	B G2	D G2	B G3	D G3	D G4	<i>Canton<sup>v</sup></i>	<i>P-1152</i>	
<b>D. melanogaster (Flybase r5.49)</b>	8749771	8010061	7614649	7092541	8792420	8365118	9448234	9092202	9421670	7925059	7751589	10153341	10011295	
<b>Dmel_all-miscRNA</b>	401121	964415	502779	542585	5455	839017	550691	764979	908722	55603	454261	924278	1416324	
<b>Dmel_all-tRNA</b>	122722	215323	108824	118454	63501	138242	192744	20465	212698	117016	99062	107897	557997	
<b>Dmel_all-transposon</b>	4 206 089	3 362 562	3 447 748	3 219 305	4 058 452	3 707 867	4 502 347	4 084 503	4 100 466	3 537 130	3 417 567	4423806	4067114	
<b>Dmel_miR_r20</b>	1 959 665	1 714 231	1 803 742	1 546 512	1 998 328	1 732 825	1 879 526	1 965 293	2 049 340	1 885 826	2 119 232	1638666	1235457	
<b>Remaining Unmatched</b>	2 060 174	1 753 530	1 751 556	1 665 685	2 126 639	1 947 167	2 322 926	2 072 777	2 150 444	1 829 057	1 661 467	477717	477418	
<b>effective depth</b>	8225928	6830323	7003046	6431502	8723464	7387859	8704799	8306758	8300250	7752440	7198266	9121166	8036974	
<b>Normalization factor</b>	0,78	0,94	0,92	1,00	0,74	0,87	0,74	0,77	0,77	0,83	0,89	0,71	0,80	

<b>B. Libraries</b>	1	2	3	4	1	2
<b>alignment reference</b>	GRH97	GRH98	GRH99	GRH100	GRH15	GRH16
<b>Subline</b>	P1152 x W1118	P1152 x W1118	W1118 x P1152	W1118 x P1152	RS3 x BQ16	BQ16 x RS3
<b>D. melanogaster (Flybase r5.49)</b>	8972279	10144872	12760054	12631941	19991599	22763040
<b>Dmel_all-miscRNA r5.49</b>	1159528	1422030	990966	1152945	3380686	2941295
<b>Dmel_all-tRNA r5.49</b>	81301	101940	101888	93653	1102954	1502668
<b>Dmel_all-transposon-r5.49</b>	4344546	4972906	6732299	6620332	6457261	7217350
<b>Dmel_miR_r20</b>	2069196	2130922	2986509	2855154	4382403	5669436
<b>effective depth</b>	7731450	8620902	11667200	11385343	15507959	18319077
<b>Normalization factor</b>	1,00	0,90	0,66	0,68	1,00	0,85
<b>C. Libraries</b>	1	2	3	1	2	
<b>alignment reference</b>	GRH107	GRH108	GRH12	GRH150	GRH148	
<b>Subline</b>	BX2_P1152	BX2_P1152	BX2/ +	RS3_BX2	RS3_BX2	
<b>Generations</b>	G1	G4	naïve	G1	G7	
<b>D. melanogaster (Flybase r5.49)</b>	7862750	10357598	20 140 166	10462432	10975285	
<b>Dmel_all-miscRNA r5.49</b>	526863	806014	1 393 137	1134092	1298089	
<b>Dmel_all-tRNA r5.49</b>	160045	250354	1 101 672	556522	538921	
<b>Dmel_all-transposon-r5.49</b>	3938988	5213229	7 687 715	4627974	4849380	
<b>Dmel_miR_r20</b>	1931216	2370154	6 848 463	2228288	2353898	
<b>effective depth</b>	7175842	9301230	17 645 357	8771818	9138275	
<b>Normalization factor</b>	1,00	0,77	0,41	1,00	0,96	

**Table S3. Primers used in this study**

<b>Primers</b>	<b>Sequences 5' to 3'</b>
RyAd	For GGGATCTGCGCCACATCG Rev CGGGCTGCAGGAATTCGATA
AdCG	For CTTGGCTGCAGGTCGACG Rev GGAGTACGGCCTTTGACCC
lash	For CTTTGGCCTTAGTCGACGGA Rev TCCTGGAGCCCGTCAGTATC
LacZ1	For GAGAATCCGACGGGTTGTTA Rev AAATTCAGACGGCAAACGAC
lacZ2	For ACTATCCCGACCGCCTTACT Rev GTGGGCCATAATTCAATTCG
42AB	For AAGACCCAATTTTTGCGTCGC Rev CAAGGATAGGGATTGGTCC
rp49/RpL32	For CCGCTTCAAGGGACAGTATCTG Rev ATCTCGCCGCAGTAAACGC
Moon	For TCGATCATCACATCCTCCGG Rev AGAACCAGTCATTTCAACCACA
Rhino	For TCTTGGTCTGGTCGATGCAC Rev AAAACCGCTCCACTTCACCA
TAS_X3_1	LEFT CGTATGCGAGAGGAGTGTA RIGHT TCCAACGCCCTAAAGAGAGA

**Table S4. Measure of ovarian *lacZ* repression for the maternally and paternally inherited *P(lArB)* transgene**

Sublines	Paternally transmitted <i>P(lArB)</i> silencer transgene: <i>P(lArB)-PI</i>				Maternally transmitted <i>P(lArB)</i> silencer transgene: <i>P(lArB)-MI</i>			
	A	B	C	D	E	F	G	H
G1	9.25% (n=1000)				93.5% (n=1000)			
G2	55.7% (n=1100)	44.35% (n=1100)	42.5% (n=1350)	41.2% (n=750)	89.7% (n=1100)	85.1% (n=1350)	86.75% (n=1450)	87.6% (n=1550)
G3	82.25% (n=1300)	83.75% (n=1500)	75.1% (n=1400)	76.25% (n=1450)	96.75% (n=1350)	93.55% (n=1400)	91.8% (n=1500)	98.0% (n=1400)
G4	96.3% (n=1400)	90.5% (n=1350)	91.95% (n=1300)	95.1% (n=1600)	91.35% (n=1800)	88.1% (n=1650)	84.15% (n=1700)	86.15% (n=1350)
G5	95.93% (n=1350)	NT	89.6% (n=600)	88.35% (n=1300)	92.4% (n=1550)	85.55% (n=1250)	NT	90.9% (n=1250)
G6	94.5% (n=850)	92.1% (n=600)	92.7% (n=1050)	98.0% (n=1000)	91.8% (n=850)	91.18% (n=800)	96.9% (n=1000)	91.85% (n=650)
G7	77.4% (n=1200)	98.75% (n=1200)	NA	NA	89.5% (n=1100)	80.0% (n=2000)	NA	71.0% (n=1550)
G8	13.5% (n=1450)	73.5% (n=1000)	NA	NA	78.1% (n=1350)	69.75% (n=800)	NA	73.1% (n=1450)
G9	46.4% (n=700)	87.4% (n=700)	NA	NA	96.5% (n=800)	79.5% (n=400)	NA	

To measure the level of *lacZ* repression, females of each crosses shown in **Fig. S1** were crossed with males encoding an euchromatic *P-lacZ* transgene target expressed in female germline. The *lacZ* germline repression was scored by calculating the % of repressed egg chambers. Twelve females were tested for each cross. n is the number of total egg chambers counted.

NT: Not Tested

NA: Not Applicable, when the subline was unfertile

**Table S5. Normalized number of 23 to 29 reads found in the 3 sublines sequenced**

<i>P(lArB)</i> -PI D	<i>P(lArB)</i>	T3 domain	<i>lacZ</i>	<i>Hsp70</i>	<i>Adh</i>	<i>rosy</i>	<i>Plasmid</i>
G1	5394	5012	649,02	1233,66	1208,43	1057,92	667,29
G2	11035,64	4562,25	1051,05	3391,85	2769,69	2769,69	1148,84
G3	14109,17	4098,54	1733,87	4447,14	3605,52	3605,52	1319,7
G4	17277,57	4223,94	1888,58	5498,42	4183,89	4183,89	1543,26

<i>P(lArB)</i> -PI B	<i>P(lArB)</i>	T3 domain	<i>lacZ</i>	<i>Hsp70</i>	<i>Adh</i>	<i>rosy</i>	<i>Plasmid</i>
G1	4117,36	4360,08	526,14	990,86	738,52	882,08	510,60
G2	14475,14	5405,70	1481,48	3955,30	3445,44	2666,22	1812,26
G3	14433,65	4376,68	1691,69	4082,54	3284,05	2795,10	1663,20

<i>P(lArB)</i> -MI H	<i>P(lArB)</i>	T3 domain	<i>lacZ</i>	<i>Hsp70</i>	<i>Adh</i>	<i>rosy</i>	<i>Plasmid</i>
G1	30211,74	4252,56	3774,42	7861,62	5204,16	9133,80	2786,94
G2	26079,36	4624,80	3091,66	6969,16	4926,54	7284,06	2334,96
G3	20991,64	3658,84	2771,04	6275,32	4085,72	4928,44	1874,04
G4	20680,00	4275,00	2512,00	5900,00	4516,00	4782,00	1796,00

<b>Control Strains</b>	<i>P(lArB)</i>	T3 domain	<i>lacZ</i>	<i>Hsp70</i>	<i>Adh</i>	<i>rosy</i>	<i>Plasmid</i>
<i>Canton (G0)</i>	60	-	-	24	31	4	1
<i>P-1152 (G0)</i>	40769	7477	4773	8090	7825	13807	4316

**Table S6. Number of 23-29 nt reads with 0, 1, 2 and 3 mismatches matching to *P(lArB)*, *T3* and *P(lacW)*.**

Libraries	<i>Canton</i> <sup>a</sup>		<i>P-1152-MI G1</i>	<i>RS3_BQ16</i>
	<i>P(lArB)</i> 18 kb	<i>T3</i> (0.9 kb)	<i>white</i> ( <i>P(lacW)</i> ) 4.1 kb	<i>pBR322</i> ( <i>P(lacW)</i> ) 1.8 kb
0 mm	60 (3.3)	0	4 (0.97)	13 (7.2)
1 mm	92 (5.1)	11 (12.2)	4 (0.97)	16 (8.8)
2 mms	98 (5.4)	197 (218.8)	4 (0.97)	16 (8.8)
3 mms	147 (8.1)	1407 (1563.3)	61 (14.8)	16 (8.8)

Number of piRNAs per kb are indicated in bracket

<sup>a</sup>: *Canton* does not encode the *1A* cluster



## **DISCUSSION ET PERSPECTIVES**



Les piRNA clusters sont les régulateurs majeurs de l'activité des ET dans la lignée germinale des animaux ; ils contrôlent la mobilisation des ET présents mais également l'invasion du génome par de nouveaux ET. A ce titre, les piRNA clusters sont les gardiens de l'intégrité génomique. De nombreuses questions relatives à la biologie et à la dynamique des piRNA clusters restent encore en suspens.

Au cours de ma thèse, j'ai étudié le rôle de facteurs environnementaux et génétiques dans l'activation d'un piRNA cluster, *BX2*, pour la production de piRNA. En effet, nous avons découvert que l'augmentation de la température au cours du développement (29°C au lieu de 25°C) et la perte de fonction de deux gènes, *hfp* et *JHDM2*, sont capables d'activer la production de piRNA par le cluster *BX2<sup>OFF</sup>*. Dans les deux cas, l'activation de la production de piRNA ne nécessite pas l'héritage de piRNA maternels. Néanmoins, comme pour la paramutation issue de l'héritage de piRNA maternels, la conversion épigénétique de *BX2* par la température ou par la perte de fonction de *hfp* est stable à travers les générations et ce même en l'absence de l'inducteur. Parallèlement, j'ai pu participer à l'étude de la dynamique d'un piRNA cluster subtélomérique, récapitulant les premières étapes de la mise en place de la régulation d'un nouvel ET issu d'un transfert horizontal.

## I. Comment définir un piRNA cluster

Qu'est-ce qu'un piRNA cluster ? Une réponse simple, et simplifiée, serait qu'un piRNA cluster est une séquence qui produit des piRNA ; mais est-ce que la production de piRNA seule est suffisante pour identifier une séquence comme étant un piRNA cluster ? Peut-on définir un piRNA cluster autrement que par sa seule production de piRNA ? Possède-t-il des caractéristiques moléculaires (transcription, chromatine ou localisation) spécifiques ?

### 1. Définition biologique

Nous savons qu'il existe différents types de séquences capables de produire des piRNA dans le génome chez la drosophile : les piRNA clusters uni-strand et dual-strand mais également certains gènes qui sont capables de produire des piRNA au niveau de leur 3'UTR (Brennecke *et al*, 2007; Robine *et al*, 2009; Saito *et al*, 2009). Même si ces séquences sont toutes capables

de produire des piRNA, leur biologie, leur composition et les facteurs chromatinien qui y sont associés ne sont pas exactement les mêmes : les piRNA clusters (uni-strand ou dual-strand) sont composés de nombreux fragments d'ET et enrichis par la marque hétérochromatique H3K9me3 ce qui n'est pas le cas des gènes dont les 3'UTR produisent des piRNA ; seuls les piRNA clusters dual-strand sont majoritairement enrichis en Rhino et sont transcrits dans les deux sens (sens et antisens) ; les facteurs impliqués de la biogénèse des piRNA produit par les piRNA cluster uni-strand et par les 3'UTR de gènes sont les mêmes mais ils ne sont pas tous identiques à ceux impliqués dans la biogénèse des piRNA cluster dual-strand. En outre, d'un point de vue fonctionnel, les piRNA issus des piRNA clusters participent essentiellement à la régulation des ET tandis que les piRNA géniques, issus de 3'UTR de gènes, semblent plutôt impliqués dans la régulation d'autres gènes (Brennecke *et al*, 2007; Czech *et al*, 2013; Goriaux *et al*, 2014; Handler *et al*, 2013; Klattenhoff *et al*, 2009; Mohn *et al*, 2014; Muerdter *et al*, 2013; Robine *et al*, 2009; Saito *et al*, 2009; Zhang *et al*, 2014). Des travaux rapportent également que des insertions d'ET à différentes positions génomiques sont capables d'induire la production de piRNA par l'ET lui-même et au niveau des séquences adjacentes. Les auteurs estiment que ces insertions correspondent à la formation de nouveaux piRNA clusters, même si le profil chromatinien (marques H3K9me3, présence ou non de Rhino) n'est pas analysé (Shpiz *et al*, 2014). Ainsi, proposer une définition unique n'apparaît pas si évident puisque la manière de définir un piRNA cluster peut varier selon les propriétés moléculaires qu'on y associe. Nous admettons communément qu'un piRNA cluster est une séquence répétée composée de fragments d'ET, capable de produire des piRNA et enrichie en marques H3K9me3. A ces propriétés communes s'ajoutent, dans le cas particulier des piRNA clusters dual-strand, un enrichissement en Rhino et une transcription dans les deux sens dépendante d'une machinerie de transcription spécifique impliquant Moon et l'ARN polymérase II (Andersen *et al*, 2017; Brennecke *et al*, 2007; Klattenhoff *et al*, 2009; Mohn *et al*, 2014; Zhang *et al*, 2014).

## 2. Définition bioinformatique

Si définir les piRNA clusters biologiquement n'est pas si aisé, les définir d'un point de vue bioinformatique semble plus complexe encore. Définir bioinformatiquement un piRNA cluster se rapporte à la manière dont on peut, à partir de données de séquençages (DNA-seq, RNA-seq, small RNA-seq et CHIP-seq), établir la localisation génomique précise d'un piRNA

cluster. En ce sens, définir les piRNA clusters en se basant sur la définition biologique apparaît plutôt simple puisqu'il s'agit d'identifier des régions riches en ET et produisant des piRNA. Néanmoins, selon la méthode d'analyse et le génome de référence utilisé, les résultats peuvent varier.

En 2007, Julius Brennecke et ses collaborateurs ont été les premiers à identifier et définir les piRNA clusters chez la drosophile. Pour ce faire, les petits ARN ovariens associés aux protéines PIWI ont été immunoprécipités puis séquencés. Les piRNA issus de ces banques de données ont été alignés sur le génome de référence de la drosophile (Flybase, Release 5), les reads normalisés en fonction du nombre d'occurrence et filtrés, ainsi seules les reads uniques ont été analysées permettant d'identifier les régions produisant des piRNA uniques. Les régions chevauchantes sont regroupées en une seule région et ne sont considérées comme étant des piRNA clusters que les régions pour lesquelles il y a un minimum de 1 piRNA unique par kilobase pour une fenêtre de 5 kilobases. L'analyse du contenu en ET révèle par ailleurs que ces régions produisant des piRNA sont associées à une densité d'ET importante. Cette méthode a permis l'identification de 142 piRNA clusters répartis sur les différents chromosomes mesurant entre quelques centaines jusqu'à plusieurs centaines de milliers de paires de bases. Parmi ces piRNA clusters, sept se situent dans des régions euchromatiques tandis que les autres sont localisés dans des régions hétérochromatiques, c'est-à-dire au niveau des péri-centromères et des sub-télomères, ou dans des régions du génome non annotées (près d'un tiers) lorsque ces travaux sont parus (Brennecke *et al*, 2007).

Dans une étude plus récente visant à comparer la répartition des piRNA clusters entre le moustique *Anopheles gambiae* et *Drosophila melanogaster*, d'autres chercheurs ont voulu redéfinir les piRNA clusters chez la drosophile (George *et al*, 2015). Pour ce faire, ils ont utilisé des banques de données de petits ARN issues de la lignée de référence *w<sup>1118</sup>* provenant d'une publication antérieure (Brennecke *et al*, 2008). Les petits ARN de 24 à 29 nucléotides ont été alignés sur le génome de référence de la drosophile (Flybase, Release 6.04) et les reads normalisés selon le nombre d'occurrence et filtrés pour ne conserver que les piRNA uniques. Dans cette étude aussi, les régions chevauchantes sont rassemblées en une seule région et celles dont plus d'une centaine de piRNA uniques ont été identifiés sont considérées comme des piRNA clusters. Avec cette méthode, 155 piRNA clusters ont été identifiés mais notons que les frontières de ces piRNA clusters ne sont pas définies très exactement. En effet, les auteurs rapportent que le but de l'étude est de comparer la répartition des piRNA clusters entre deux espèces distinctes et non d'établir précisément les limites de ces séquences. Parmi

ces piRNA clusters, 76 sont localisés dans l'hétérochromatine péricentromérique, 26 au niveau des frontières entre l'hétérochromatine et l'euchromatine, 51 dans l'euchromatine et les 2 restants dans l'hétérochromatine du chromosome Y (George *et al*, 2015).

On peut constater que la répartition des piRNA clusters est différente entre ces deux études car même si la majorité des piRNA clusters est hétérochromatique dans les deux cas, la proportion des piRNA clusters euchromatiques est plus importante dans la seconde étude que dans la première, respectivement 32% contre environ 5% (Brennecke *et al*, 2007; George *et al*, 2015). Cette différence pourrait être en partie due au génome de référence utilisé ; ainsi les piRNA clusters présents dans les régions non annotées pour la première étude donc probablement localisés dans la seconde étude. Lorsqu'on examine les coordonnées génomiques des piRNA clusters pour ces deux études on peut constater que certains piRNA clusters sont proches. Parce que les auteurs de la seconde étude attestent eux-mêmes que les frontières des piRNA clusters qu'ils ont définis ne sont pas précises (George *et al*, 2015), nous prendrons quelques exemples issus de la première étude (Brennecke *et al*, 2007). Sur le chromosome X, le cluster 2 et le cluster *flamenco* sont séparés de 74 kb. Notons d'ailleurs que le *flamenco* est divisé en deux piRNA clusters (clusters 8 et 9) séparés de 75 kb. Sur le chromosome 3, dans l'hétérochromatine du 3R, les clusters 30 et 93 ne sont séparés que de 100 pb tandis que sur le 3L, les clusters 65 et 91 sont séparés de 31 kb. Sur le chromosome 2, au niveau du 2R les clusters 39 et 73 sont séparés de 66 kb ; au niveau du 2L, les clusters 23 et 79 sont séparés de 100 pb et on peut localiser également le cluster 31 à 37 kb du cluster 79 (Brennecke *et al*, 2007). Ainsi, en changeant certains paramètres dans l'analyse de la répartition des piRNA (augmentation ou diminution du seuil de un piRNA/kb, plus ou moins de stringence à l'étape d'alignement, choix de regroupement des régions chevauchantes seulement ou des régions proches), le nombre et/ou la taille des piRNA clusters peut varier, rendant très difficile la définition exacte de leurs frontières.

Le séquençage et l'annotation du génome de la drosophile a permis de mettre en évidence que les régions hétérochromatiques sont essentiellement constituées de régions répétées (Adams *et al*, 2000; Hoskins *et al*, 2007, 2002). En conséquence, l'identification de régions denses en ET associées à un enrichissement en marques H3K9me3 n'est pas nécessairement une signature moléculaire propre aux piRNA clusters. Quant à l'analyse de la répartition de Rhino dans les cellules germinales par ChIP-seq, elle ne permet que l'identification des piRNA clusters dual-strand. Ainsi, l'identification des piRNA clusters nécessite des analyses moléculaires complémentaires puisque l'étude de la production de piRNA ou de la répartition des marques

chromatiniennes au niveau du génome seule n'est pas suffisante. Par ailleurs, les propriétés moléculaires des piRNA clusters uni-strand et dual-strand étant différentes, leur détermination implique de prendre en compte des critères distincts lors de l'analyse.

Puisque qu'il est difficile de s'accorder sur une définition biologique ou bioinformatique des piRNA clusters, nous proposons une définition fonctionnelle : un piRNA cluster est une séquence qui produit des piRNA capables de réprimer une séquence homologue.

## II. Activation d'un piRNA cluster

*BX2* est un outil transgénique mimant un piRNA cluster dual-strand et dont nous pouvons moduler l'état épigénétique facilement. Le laboratoire a montré que le cluster *BX2<sup>OFF</sup>* pouvait être activé pour la production de piRNA grâce à l'héritage de piRNA maternels ; à l'inverse, l'absence d'héritage de piRNA maternels engendre l'inactivation d'un cluster *BX2<sup>ON</sup>* (Casier et al., 2019 / Article I; de Vanssay et al., 2012). Cet exemple de conversion épigénétique d'un piRNA cluster illustre que l'état épigénétique du cluster dépend de la présence ou non de piRNA homologues à la séquence. Autrement dit, il faut des piRNA initiaux pour permettre à un piRNA cluster de produire de nouveaux piRNA. Ce constat pose la question de l'émergence des premiers piRNA. Nous avons mis en évidence deux types de facteurs capables d'induire la production de piRNA par *BX2<sup>OFF</sup> de novo* : l'augmentation de la température (29°C au lieu de 25°C) et la perte de fonction de gènes (*hfp* et *JHDM2*).

Lorsque la température est de 29°C ou qu'on inactive les gènes *hfp* et *JHDM2*, le cluster *BX2<sup>OFF</sup>* est converti en une seule génération pour la production de piRNA. L'activation de *BX2* peut également être maintenue au cours des générations même en l'absence de l'inducteur qui a permis l'activation, à savoir le retour à 25°C ou le rétablissement de l'expression du gène *hfp*. En effet, une fois les premiers piRNA produits, l'héritage de piRNA maternels permet d'activer le cluster *BX2* à chaque génération grâce à la paramutation. Ce changement d'état épigénétique s'accompagne de modifications chromatiniennes : le cluster activé par la chaleur ou par la perte de fonction des deux gènes est enrichi en marques H3K9me3 et Rhino. Par conséquent, deux facteurs différents sont capables de convertir *BX2* de façon stable au cours des générations. Cependant, les mécanismes moléculaires à l'origine

de l'activation de *BX2<sup>OFF</sup>* par l'augmentation de la température et par la perte de fonction de gènes ne sont pas les mêmes.

Dans le cas de l'augmentation de la température, le taux de conversion en une génération est relativement faible (de l'ordre de 2% des chambres ovariennes). Par ailleurs, la valeur du taux de conversion en une génération pourrait expliquer pourquoi la conversion complète d'une lignée *BX2<sup>OFF</sup>* nécessite plusieurs dizaines de générations à 29°C. A 29°C on observe plus de transcrits précurseurs antisens de *BX2* comparé à 25°C. Or, les données de small RNA-seq indiquent que cette augmentation de la quantité de transcrits antisens n'est pas associée à une augmentation de la quantité de piRNA lorsque la séquence cible est absente du génome. En effet, nos résultats montrent qu'en l'absence d'une séquence cible et, le cas échéant, si la séquence n'est pas transcrite, la conversion de *BX2* à 29°C ne se fait pas. Notre hypothèse est qu'à 29°C les transcrits précurseurs antisens de *BX2* en excès interagissent avec les transcrits – sens – de la séquence cible, générant des ARN dual-strands qui pourraient être reconnus et clivés par des endonucléases spécifiques. Une partie des petits ARN générés pourrait alors être prise en charge par les protéines PIWI (Piwi ou Aub). Compte tenu de la valeur du taux de conversion on peut supposer que l'apparition de ce type de phénomène hypothétique est rare puisqu'il ne survient que dans 2% des cas.

Dans le cas de la perte de fonction des gènes *hfp* et *JHDM2*, le taux de conversion en une génération est relativement élevé (entre 85 et 90% des chambres ovariennes). Nous avons établi un lien entre ces deux gènes puisque *hfp* est impliqué dans l'épissage des deux isoformes transcriptionnelles de *JHDM2*. Les expériences d'immunoprécipitation d'ARN indiquent que la protéine Hfp interagit directement avec les transcrits de *JHDM2* mais ne se lie pas aux précurseurs de *BX2*. De plus, le sauvetage fonctionnel montre que l'activation de *BX2* en contexte *hfp GLKD* dépend également de la perte de fonction de *JHDM2*. *JHDM2* est une déméthylase d'histone spécifique des résidus H3K9me1 et H3K9me2. En conséquence, notre hypothèse est que la perte de fonction de *JHDM2* - induite par une mutation du gène lui-même ou par la perte de fonction de *hfp* – empêche la déméthylation active entraînant alors une accumulation des marques H3K9me2 et/ou H3K9me3 sur le cluster *BX2<sup>OFF</sup>*. Or, il a été montré que le chromodomaine de Rhino possède une affinité pour la marque H3K9me3 mais également une affinité, moins forte, pour la marque H3K9me2 (Le Thomas *et al*, 2014). En conséquence, l'accumulation des marques H3K9me2 ou H3K9me3 au niveau du cluster entraînerait alors le recrutement de Rhino suivi des facteurs impliqués dans la transcription des piRNA clusters dual-strand et l'adressage vers la voie des piRNA.

Ainsi, l'activation de *BX2* par ces deux types de facteurs, environnementaux et génétiques, se fait par des mécanismes distincts. La chaleur pourrait entraîner une activation du cluster *via* un mécanisme dépendant de l'ARN tandis que la perte de fonction des gènes *hfp* et *JHDM2* engendre l'activation de *BX2* *via* un mécanisme qui semble dépendant de la chromatine. Il est à noter que contrairement à l'activation par la chaleur qui nécessite une séquence cible transcrite, l'activation par perte de fonction de gènes ne requiert pas nécessairement de séquence homologue. La nature des mécanismes moléculaires impliqués dans l'activation de *BX2* pourrait être à l'origine de la différence du taux de conversion en une génération.

### **III. Effets pan-génomiques**

Si la chaleur et la perte de fonction de gènes sont capables d'activer la production de piRNA de *BX2*, un cluster composé de transgènes répétés, d'autres régions génomiques pourraient être converties pour la production de piRNA.

#### **1. Augmentation de température et production de piRNA**

L'analyse des données de small RNA-seq de lignées *BX2* converties à 29°C a révélée qu'il n'existe pas d'autres régions, hormis *BX2*, capables de produire *de novo* et de manière stable des piRNA (Casier *et al*, 2019/Article I). En effet, les données analysées sont issues de lignées *BX2* dont le cluster a été activé à 29°C pour la production de piRNA puis stabilisées à 25°C pendant plusieurs générations. En conséquence, cette analyse ne permet pas de mettre en évidence l'existence de régions produisant des nouveaux piRNA à 29°C mais plutôt d'identifier s'il existe des régions converties de manière stable pour la production de piRNA après une exposition à 29°C pendant plusieurs générations.

A l'échelle du génome, on ne voit pas de différence de quantité de piRNA produits entre 25°C et 29°C. En revanche, on détecte un peu moins de piRNA sur le piRNA cluster *42AB* à 29°C comparé à 25°C (Casier *et al*, 2019/Article I). Il existe peu de travaux abordant l'effet de la température sur la dynamique ou la quantité de piRNA globale. Dans une étude parue en 2017, des analyses de small RNA-seq d'ovaires de drosophiles suggèrent qu'il y a moins de piRNA à 29°C par rapport à 18°C. Parallèlement, les auteurs montrent que certains gènes de

la voie des piRNA sont dérégulés : *Zuc*, *Armi*, *Aub* et *Ago3* sont plus exprimés tandis que *Piwi*, *Yb*, *Hsp83* et *Shu* sont moins exprimés à 29°C ; et l'analyse de la signature Ping-Pong indique qu'il y a plus de piRNA secondaires à 29°C comparé à 18°C. Les auteurs émettent l'hypothèse que l'amplification Ping-Pong est plus active à 29°C, d'une part parce que les gènes *Aub* et *Ago3* sont plus exprimés, et d'autre part car les structures secondaires d'ARN sont plus accessibles. Cependant, les analyses de small RNA-seq qui indiquent qu'il y a moins de piRNA à 29°C et de la signature Ping-Pong qui suggère une amplification du mécanisme semblent contradictoires et ne permettant pas de conclure clairement (Fast *et al*, 2017). Une autre étude réalisée chez *Cordylochernes scorpioides* montre que les piRNA sont globalement moins exprimés à une température élevée (3.5°C au-dessus de la température contrôle, mais la température contrôle n'est pas précisée) dans les testicules (Su-Keene *et al*, 2018). Enfin, chez *Caenorhabditis elegans*, des résultats indiquent qu'à 25°C, l'abondance des précurseurs de piRNA diminue comparé à 20°C. Cette diminution de la quantité de précurseurs est associée à la dérégulation d'un transgène rapporteur GFP à 25°C, indiquant que la quantité de piRNA diminue à cette température (Belicard *et al*, 2018).

Ces études semblent indiquer qu'à l'échelle de génome entier, une température plus élevée entraîne la diminution de la quantité de piRNA produits, probablement par les piRNA clusters. Pourtant, nos travaux montrent que l'augmentation de la température affecte l'état épigénétique de *BX2*, permettant au cluster de produire des piRNA. Ainsi, la température n'affecte peut-être pas de la même manière un piRNA cluster actif et un piRNA cluster inactif, en capacité d'être converti pour la production de piRNA. La conversion du cluster *BX2* en réponse à ce stress environnemental apparaît être un cas très spécifique et nous amène à nous questionner sur ce qui fait la particularité de *BX2*, autrement dit : pourquoi *BX2* peut être activé pour la production de piRNA et pas d'autres régions génomiques ? Plusieurs pistes non exclusives peuvent être considérées. La première piste est relative à la structure de *BX2* qui est un cluster constitué de sept transgènes *P[lacW]* répétés en tandem, représentant une séquence de près de 75 kilobases. On peut facilement imaginer que l'émergence de grandes structures répétées, proches de *BX2* et prêtes à être converties pour la production de piRNA doit être rare. Une fois de telles structures formées elles pourraient être converties très rapidement (quelques générations), notamment grâce à des stress environnementaux, et ne sont par conséquent pas détectables à l'échelle de notre étude. Une seconde possibilité concerne la localisation génomique de *BX2*, qui est inséré dans un intron du gène *ago1*. Nous avons pu mettre en évidence qu'à 29°C, il y a plus de transcrits précurseurs antisens de *BX2*



mais également plus de transcrits d'*ago1*. Or, *ago1* est transcrit en antisens comparé à l'orientation de *BX2*, ce qui pourrait suggérer que les transcrits antisens de *BX2* sont peut-être en partie des ARN chimères *ago1-BX2* issus de la transcription d'*ago1*. Dans une étude récente, des chercheurs ont généré des lignées qui diffèrent pour la localisation d'un transgène dérivé de l'élément *P* contenant des fragments d'élément *I* (transgène *I-TG*) (Komarov *et al*, 2020). Notons que dans le génome de la drosophile, l'élément *I* est régulé par les piRNA issus d'anciennes copies de l'élément insérées dans des piRNA clusters, par conséquent les copies du transgène *I-TG* peuvent être converties pour produire elles-mêmes des piRNA par un mécanisme de paramutation. On peut constater que pour ces différentes lignées, certaines insertions produisent peu de piRNA, localisés sur le transgène tandis que d'autres produisent beaucoup de piRNA à la fois sur le transgène mais également au niveau des séquences adjacentes, indiquant qu'une même séquence ne répond pas de la même manière en fonction du site d'insertion (Komarov *et al*, 2020). L'ensemble de ces résultats suggère que la localisation et l'environnement génomique sont importants pour l'émergence d'un piRNA cluster.

## 2. Environnement, stress et piRNA clusters

Nous avons pu montrer que la température affecte la biologie de *BX2*. En effet, une température élevée (29°C) peut entraîner l'activation du cluster tandis que des températures plus basses peuvent (18°C ou 10°C), à l'inverse, induire l'inactivation de *BX2*. Ainsi, la température semble moduler l'état épigénétique de *BX2*. Dans les deux cas, nous pouvons constater que le pourcentage de conversion est relativement faible. En effet, à 29°C l'activation de *BX2<sup>OFF</sup>* n'apparaît que dans 2% des chambres ovariennes ; à des températures plus basses, l'inactivation est plus importante quand la température est la plus basse (10°C) mais le pourcentage de répression n'excède pas les 1%. En l'état actuel, nous n'envisageons pas d'entreprendre de nouvelles expériences sur le rôle de la température sur la biologie des piRNA cluster (de *BX2* en particulier). En effet, avec ces pourcentages de conversion faibles, les analyses moléculaires (Small RNA-seq, ChIP-seq...), en particulier pour l'étude des températures basses, s'avèrent extrêmement difficiles, limitant la compréhension des mécanismes moléculaires impliqués.

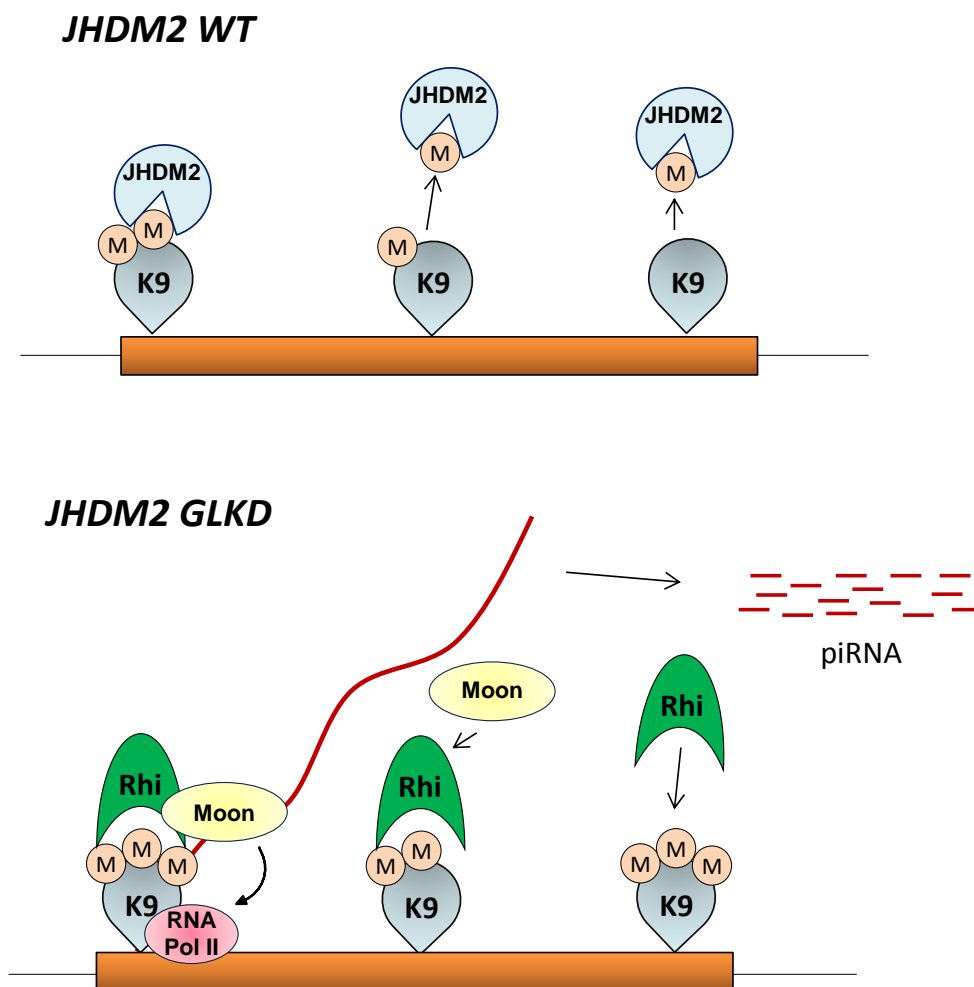
Nous n'avons pas pu identifier d'autres stress capables de modifier l'état épigénétique de *BX2* parmi ceux que nous avons testé (sel, cuivre et bisphénol A). Ces résultats indiquent que l'effet de la température est spécifique et que *BX2* ne peut pas être activé par n'importe quel stress environnemental. Cependant, il serait intéressant de tester d'autres stress sur l'état épigénétique de *BX2*, notamment le peroxyde d'hydrogène (H<sub>2</sub>O<sub>2</sub>) et l'oxide nitrique (NO). Il a été montré dans différents types de lignées cellulaires que le stress oxydatif (H<sub>2</sub>O<sub>2</sub>, NO) est capable de moduler l'activité des déméthylases à domaine Jumonji-C dont *JHDM2* fait partie (Hancock *et al*, 2017; Hickok *et al*, 2013; Niu *et al*, 2015; Shmakova *et al*, 2014). De cette manière, un stress environnemental pourrait entraîner l'activation de *BX2* en modulant l'activité de *JHDM2*.

### 3. Inactivation de gènes et production de piRNA

A la différence de la chaleur, l'analyse des données de small RNA-seq indique que l'inactivation de *JHDM2* dans les ovaires modifie la production de piRNA à l'échelle du génome. En effet, nous avons pu identifier des régions génomiques qui produisent des piRNA en contexte *JHDM2 GLKD* alors qu'elles n'en produisent pas chez le contrôle. Ces régions sont de tailles variables et réparties sur les différents chromosomes. Des expériences de RT-qPCR et de RNA-seq montrent également que la production de piRNA est associée à une transcription aberrante de ces régions qui, de manière étonnante, contiennent majoritairement des facteurs de transcription impliqués dans des processus développementaux qui ne s'expriment pas - ou s'expriment très peu - dans les ovaires en contexte sauvage. Alors que *JHDM2* est régulé par *hfp*, on n'identifie pas ces régions génomiques produisant de nouveaux piRNA en contexte *hfp GLKD*. De plus, le fait que les descendants des *hfp GLKD* soient fertiles semble également indiquer que le génome est moins affecté. Il s'avère que le contexte *hfp GLKD* n'affecte pas autant l'expression de *JHDM2* que le contexte *JHDM2 GLKD* lui-même, ce qui pourrait expliquer, d'une part pourquoi on ne retrouve pas dans le contexte *hfp GLKD* les régions identifiées pour le contexte *JHDM2 GLKD* ; et d'autre part pourquoi les individus *hfp GLKD* sont fertiles alors que les individus *JHDM2 GLKD* sont stériles.

Des résultats de RTqPCR nous montrent que dans des mutants *JHDM2 GLKD* ; *moon*<sup>-/-</sup>, la transcription de ces régions est fortement diminuée, indiquant que Moon est impliqué dans la transcription non-canonique de ces régions. Moon est un acteur important de la voie des

piRNA, il interagit avec les facteurs du complexe RDC et recrute l'ARN polymérase II permettant la transcription des piRNA clusters dual-strand. Ces données suggèrent qu'en contexte *JHDM2 GLKD*, des régions génomiques se comportent comme des piRNA clusters. Ainsi, de la même manière que pour la conversion du cluster *BX2*, l'hypothèse est que la perte de fonction de *JHDM2* empêche la déméthylation active de ces régions entraînant alors une accumulation des marques H3K9me2 et/ou H3K9me3. Le chromodomaine de Rhino possède une affinité pour les marques H3K9me2 et H3K9me3 (Le Thomas *et al*, 2014). En conséquence, l'accumulation des marques H3K9me2 et/ou H3K9me3 entrainerait alors le recrutement de Rhino suivi des facteurs impliqués dans la transcription des piRNA clusters dual-strand et l'adressage vers la voie des piRNA au niveau des différentes régions génomiques (Figure 41).



**Figure 41 : Modèle sur le rôle de *JHDM2*.** En contexte sauvage (en haut), *JHDM2* reconnaît et déméthyle spécifiquement les marques H3K9me1 et H3K9me2 (M : groupe méthyl, K9 : lysine 9), empêchant l'accumulation d'hétérochromatine au niveau de certaines régions dans les ovaires. En contexte *JHDM2 GLKD* (en bas), *JHDM2* n'est plus présent pour déméthyliser les marques H3K9me1 et H3K9me2, ce qui entraîne une accumulation des marques

hétérochromatiniennes H3K9me2 et/ou H3K9me3. Ces marques sont alors reconnues par Rhino (Rhi) qui recrute Moon et l'ARN polymérase II (RNA Pol II) permettant la transcription non-canonique de la région. Des piRNA sont alors produits à partir des ARN précurseurs issus de cette transcription.

Pour autant, nous ne savons pas pourquoi certaines régions particulières produisent des nouveaux piRNA en contexte *JHDM2 GLKD* et pas d'autres. En effet, nous n'avons pas encore d'évidence selon laquelle JHDM2 serait capable de reconnaître des motifs particuliers afin de démétyler spécifiquement certaines régions en contexte sauvage. Par ailleurs, si nous confirmons que ces régions sont effectivement enrichies en marques H3K9me2 et/ou H3K9me3, la raison pour laquelle ces marques seraient reconnues par Rhino plutôt que par HP1 reste à élucider. Une hypothèse serait que ces régions se trouvent à proximité de piRNA clusters existants et qu'il y a une extension des domaines hétérochromatiniens sur les régions adjacentes en contexte *JHDM2 GLKD*. Cependant, définir correctement les piRNA clusters s'avère parfois difficile. Pour l'heure, nous manquons de données de ChIP sur les marques H3K9me2, H3K9me3, Rhino et JHDM2 pour valider l'ensemble de ces hypothèses.

L'analyse des small RNA-seq des *JHDM2 GLKD* a aussi révélée qu'environ 155 régions génomiques produisent un peu moins de piRNA par rapport au contrôle. De manière étonnante, on retrouve une partie des piRNA clusters parmi ces régions. Ainsi, tandis que certaines régions produisent des nouveaux piRNA, d'autres produisent moins de piRNA. Cette observation pourrait suggérer que, en contexte *JHDM2 GLKD*, les régions produisant des nouveaux piRNA sont en compétition pour les facteurs de la voie des piRNA avec les piRNA clusters qui, en conséquence, produisent moins de piRNA.

L'ensemble de nos résultats suggère que *JHDM2* joue un rôle dans la spécification de l'identité des piRNA clusters en maintenant dans un état inactif, non producteur de piRNA, certaines régions génomiques particulières grâce à une déméthylation active des marques H3K9me1 et H3K9me2.

#### **4. Fonctions de *JHDM2***

Le rôle de *JHDM2* dans la détermination des piRNA clusters et les mécanismes moléculaires associés restent encore à établir clairement. Des expériences de ChIP-seq sur les marques H3K9me2, H3K9me3 et Rhino nous permettraient de mettre en évidence les changements

globaux au niveau de la chromatine en contexte *JHDM2 GLKD* et qui pourraient expliquer comment certaines régions produisent des piRNA de manière aberrante. Pour autant, ces données de séquençages seules ne permettent pas d'expliquer pourquoi ces régions en particulier et pas d'autres sont converties pour la production de piRNA. En contexte sauvage, *JHDM2* pourrait interagir avec ces régions pour les empêcher de produire des piRNA en les déméthylant activement. Des expériences de ChIP-seq nous permettraient de déterminer si *JHDM2* interagit avec ces régions et s'il existe ou non un motif reconnu par la protéine à l'origine de sa spécificité. En parallèle, des expériences de spectrométrie de masse seront nécessaires pour identifier de potentiels partenaires moléculaires permettant à *JHDM2* d'assurer sa fonction. Cependant, nous ne possédons pas d'anticorps capable de reconnaître *JHDM2* pour réaliser ces expériences de ChIP-seq et de spectrométrie de masse. Pour pallier l'absence de cet anticorps, Julie Autaa, étudiante en M2, a commencé la construction d'une lignée *JHDM2*-GFP en utilisant une technique dérivée de CRISPR-Cas9 permettant l'insertion de la séquence de la *GFP* en phase avec *JHDM2*. En effet, il existe des anticorps anti-GFP efficaces, vendus dans le commerce, pour les expériences nécessitant des étapes d'immunoprécipitation.

Un autre aspect important dans l'analyse des individus *JHDM2 GLKD* est de comprendre pourquoi la descendance est stérile. En effet, les individus pondent des œufs qui n'éclosent pas. Tous les descendants des *JHDM2 GLKD*, incluant aussi bien les individus qui portent la mutation et ceux qui ne la portent pas, sont incapables de se développer au-delà du stade L1. Cette observation suggère qu'il y a un effet maternel, or nous savons que les individus héritent des piRNA maternels. Les individus qui se développent jusqu'au stade L1 ont des phénotypes qui ressemblent aux mutants perte de fonction des gènes *labial* et *Oc*, deux gènes insérés dans les régions produisant de nouveaux piRNA chez les mères *JHDM2 GLKD*. Notre hypothèse est donc que les piRNA produits de manière aberrante dans les ovaires des femelles *JHDM2 GLKD* sont hérités dans la descendance et entraîne la répression de gènes essentiels au développement correct de l'individu par hétérochromatinisation locale (TGS).

Pour vérifier cette hypothèse, nous devons vérifier fonctionnellement si des mutants *JHDM2 GLKD*-voies des piRNA permettent aux descendants de se développer, mais obtenir des mutants homozygotes viables pour la voie des piRNA s'avère difficile. Nous devons également mesurer l'expression (par RT-qPCR ou/et par RNA-seq) et déterminer l'état chromatinien par des expériences de ChIP-seq (marques H3K9me3 et Rhino) sur différents gènes dans des embryons descendants des *JHDM2 GLKD*.

## IV. Être ou ne pas être un piRNA cluster : des facteurs pour spécifier l'identité des piRNA clusters

Définir précisément ce qu'est un piRNA cluster et où sont ses limites peut s'avérer difficile. En conséquence, l'identification de facteurs impliqués dans la définition des piRNA clusters constituerait des marqueurs moléculaires des frontières des piRNA clusters. Nos résultats suggèrent que *JHDM2* pourrait être l'un de ces facteurs en participant à la déméthylation des marques H3K9me1/2 afin d'empêcher certaines régions génomiques de produire des piRNA, mais l'ensemble des mécanismes moléculaires mis en jeu restent encore à préciser. Le maintien de l'intégrité des piRNA clusters est essentiel pour contrôler l'activité des ET et ainsi assurer la protection du génome au cours du développement. Cependant, les frontières des piRNA clusters doivent être finement établies afin d'empêcher ou de limiter l'extension des marques chromatiniennes sur les régions adjacentes qui peuvent entraîner une modification de l'expression des gènes affectés.

Actuellement, très peu d'études rapportent le rôle de facteurs impliqués dans la détermination des piRNA clusters. Les insulateurs, ces barrières moléculaires impliquées dans la délimitation des domaines chromatiniens, apparaissent comme de bons candidats dans la délimitation de ces frontières. Certains rétrotransposons contiennent des insulateurs, c'est notamment le cas pour *gypsy* qui possèdent des sites de liaisons pour Suppressor of Hairy Wing (*Su(Hw)*) dont l'activité est capable de modifier l'expression de certains gènes (Gdula *et al*, 1996; Geyer & Corces, 1992; Soshnev *et al*, 2013). Dans une étude récente, des chercheurs ont étudié l'effet des insulateurs de *gypsy* sur la production de piRNA par les piRNA clusters dans les ovaires de la drosophile (Radion *et al*, 2019). Pour ce faire, ils ont inséré un transgène dérivé de l'élément *P* contenant deux gènes rapporteurs ainsi que des séquences insultrices de *gypsy* dans deux piRNA clusters endogènes du chromosome 2R, l'un localisé au niveau du péricentromère (cluster *42AB*) et l'autre au niveau des séquences subtélomériques (*TAS-2R*). Les insertions dans ces piRNA clusters ne permettent pas au transgène de produire des piRNA dans la condition contrôle, en revanche la mutation de *Su(Hw)* entraîne une production de piRNA primaires par le transgène. Cette production de piRNA en contexte mutant pour *Su(Hw)* est associée à une quantité de transcrits précurseurs plus importante et à un enrichissement en marques H3K9me3, Rhino et HP1 au niveau du transgène, suggérant que *Su(Hw)* bloque la propagation des marques hétérochromatiques des piRNA clusters sur le transgène chez le contrôle. Par ailleurs, chez le contrôle l'insertion du

transgène dans *42AB* entraîne une diminution de la quantité de transcrits précurseurs et de piRNA produit par le cluster en amont du transgène tandis que dans le mutant *Su(Hw)* la production de transcrits précurseurs et de piRNA par *42AB* en amont de l'insertion du transgène sont restaurées (Radion *et al*, 2019).

L'étude au niveau du génome entier a révélé que plus de 100 gènes sont surexprimés dans le mutant *Su(Hw)* par rapport au contrôle, la grande majorité de ces gènes n'étant pas exprimés dans les ovaires en contexte sauvage (Soshnev *et al*, 2013). On a détecté également 66 gènes produisant des piRNA chez le mutant *Su(Hw)* alors qu'ils n'en produisent pas chez le contrôle (Radion *et al*, 2019). Ces phénotypes moléculaires d'expression ectopique et de production aberrante de piRNA rappellent en partie les effets pan-génomiques en contexte *JHDM2 GLKD*. Néanmoins il est important de noter qu'en contexte mutant pour *Su(Hw)*, les gènes surexprimés ne sont pas ceux qui produisent des piRNA (Radion *et al*, 2019; Soshnev *et al*, 2013). Dans le cas des *JHDM2 GLKD*, les gènes surexprimés sont également ceux qui produisent des piRNA (Partie II des résultats). De plus, les gènes affectés par le mutant *Su(Hw)*, aussi bien pour la transcription que pour la production de piRNA, ne sont pas les mêmes que ceux affectés en contexte *JHDM2 GLKD* (Radion *et al*, 2019; Soshnev *et al*, 2013).

Nous manquons encore de données pour établir la fonction des insulateurs et des facteurs chromatinien dans la détermination des piRNA clusters. Tous les ET ne possèdent pas d'insulateurs et il a également été montré que différents transgènes dérivés de l'élément *P* insérés dans des piRNA clusters endogènes sont convertis pour la production de piRNA et n'empêche pas la production de piRNA par le reste du cluster (Muerdter *et al*, 2012; Olovnikov *et al*, 2013; Radion *et al*, 2018).

## **V. Dynamique des piRNA clusters et protection du génome**

En 1996, Stéphane Ronsseray a identifié des insertions de copies d'élément *P* dans les séquences subtélomériques permettant de réguler d'autres copies d'éléments *P* dans le génome (Ronsseray *et al*, 1996). On sait désormais que ces séquences subtélomériques sont essentiellement composées de piRNA clusters et que la régulation des ET s'effectue grâce à l'activité des piRNA (Brennecke *et al*, 2007). Selon le modèle actuel, la mise en place de la

répression d'un ET se fait en plusieurs étapes : l'arrivée de l'ET dans un génome, l'invasion par transposition puis l'initiation et le maintien de la régulation de l'ET. Par la suite, des évènements de délétion dans la séquence des piRNA clusters peuvent conduire à la dérégulation puis la remobilisation de certains ET (Kofler, 2019; Zanni *et al*, 2013).

## 1. Arrivée d'un ET et invasion du génome

Les transferts horizontaux peuvent être responsables de l'apparition d'un nouvel ET dans un génome. Le premier cas référencé de transfert horizontal d'un ET chez le genre *Drosophila* est celui de l'élément *P*. Il apparaît que l'élément *P* provenant de *Drosophila willistoni* a envahi le génome de *Drosophila melanogaster* au cours du 20<sup>ème</sup> siècle (Daniels *et al*, 1990; Kidwell *et al*, 1977). Plus récemment, il a été montré que l'élément *P* a également envahi le génome de *Drosophila simulans* après un évènement de transfert horizontal (Kofler *et al*, 2015a). Plus d'une centaine de cas de transferts horizontaux ont été identifiés pour le genre *Drosophila*, indiquant que ce phénomène n'est pas rare (Loreto *et al*, 2008). Des mécanismes d'échange de matériel génétique entre un vecteur portant l'ET (virus, bactéries endosymbiotiques ou parasites) et l'espèce receveuse ont été proposés pour expliquer le phénomène de transfert horizontal mais les évidences expérimentales manquent encore pour confirmer ces hypothèses (Gilbert & Cordaux, 2017; Loreto *et al*, 2008).

Lorsqu'un nouvel ET apparaît dans un génome son activité n'est pas régulée immédiatement, notamment parce qu'il n'existe pas de piRNA complémentaires capables de reconnaître cette nouvelle séquence. Cette absence de régulation permet au nouvel ET de transposer et génère des mutations qui entraînent l'atrophie des ovaires dans la majorité des cas. Cet état persiste jusqu'à la mise en place de la régulation de l'ET qui intervient dès lors que celui-ci s'insère dans un piRNA cluster.

Il est communément admis que le taux de transposition est assez faible, mais calculer précisément le taux de transposition s'avère difficile. En effet, les modèles suggèrent que le taux de transposition est très hétérogène car il varie selon de nombreux paramètres, notamment en fonction de l'espèce, du tissu, du fond génétique, de l'âge, du type et du nombre de copies d'ET, de la présence ou non d'un système de régulation des ET, ou encore des conditions environnementales (Charlesworth & Charlesworth, 1983; Díaz-González *et al*, 2011; Kelleher *et al*, 2018; Kofler *et al*, 2018, 2015b; Sousa *et al*, 2013). Des modèles



théoriques et des simulations tentent de déterminer le délai existant entre le moment où un nouvel ET apparaît dans un génome et celui où l'ET est régulé efficacement. D'après les modèles proposés par deux équipes différentes, il apparaît que la valeur du taux de transposition influence beaucoup le temps requis pour la mise en place de la régulation d'un nouvel ET (Kelleher *et al*, 2018; Kofler, 2019). Dans ces deux études, la mobilité des ET est estimée indirectement grâce au nombre de copies de l'ET (Kelleher *et al*, 2018) ou par le nombre d'insertion dans le génome (Kofler, 2019) ; dans les deux cas il est admis que la mobilité est bloquée lorsqu'un plateau est atteint. Quelle que soit la méthode utilisée, on constate que lorsque le taux de transposition est élevée, la mobilité des ET est bloquée plus rapidement ; à l'inverse la régulation se met en place plus tardivement si le taux de transposition est faible. Ainsi, le nombre de générations nécessaires pour la mise en place de la répression d'un nouvel ET peut varier considérablement, entre 200 générations pour un taux de transposition élevé jusqu'à 2000 générations pour un taux de transposition faible (Kelleher *et al*, 2018; Kofler, 2019). Certains travaux ont montré que de nombreux stress environnementaux induisent une augmentation importante de la transposition, événement qu'on appelle 'burst' de transposition (Esnault *et al*, 2019; García Guerreiro, 2012; Kofler *et al*, 2018; Miousse *et al*, 2015). Ces bursts de transposition sont plus susceptibles d'entraîner des dommages à l'ADN mais augmentent également les chances pour qu'une copie d'un nouvel ET s'insère dans un piRNA cluster existant, permettant la mise en place de sa régulation plus rapidement.

## **2. Mise en place et maintien de la régulation**

Lorsqu'une copie d'un nouvel ET s'insère dans un piRNA cluster, il est converti pour la production de piRNA, devenant partie intégrante du piRNA cluster dans lequel il est localisé. Cette insertion permet ainsi la régulation de la mobilité des copies actives présentes dans le génome. En effet, les études montrent que des transgènes dérivés de l'élément *P* acquièrent la signature moléculaire du piRNA cluster dans lequel ils sont insérés, à savoir des capacités répressives de copies actives associées à la production de piRNA ainsi qu'à un enrichissement en marques H3K9me3 et Rhino (Khurana *et al*, 2011; Muerdter *et al*, 2012; Olovnikov *et al*, 2013; Radion *et al*, 2019, 2018). Cependant, le temps nécessaire à l'acquisition de ces caractéristiques moléculaires par un ET nouvellement inséré dans un piRNA cluster n'est pas clairement établi.

L'utilisation de différents transgènes dérivés de l'élément *P* (*BX2*, *P(lArB)*...) nous permet d'étudier cette dynamique. *P(lArB)* est un transgène contenant le gène *lacZ* inséré dans le subtélomère du chromosome *X* (cluster *IA*) capable de réprimer en *trans* l'expression d'un transgène *lacZ* dans le génome (Roche & Rio, 1998). Le laboratoire a montré que lorsque le transgène *P(lArB)* est hérité maternellement, on observe un phénotype de variévation pour lequel le transgène cible est réprimé dans la majorité des chambres ovariennes dès la première génération, ce pourcentage de répression se maintient ensuite aux générations suivantes. En revanche, lorsque *P(lArB)* est hérité paternellement, on constate que peu de chambres ovariennes présentent de la répression à la première génération ; le nombre de chambres ovariennes présentant de la répression augmente progressivement au cours des générations suivantes jusqu'à atteindre approximativement le même que celui d'un transgène hérité maternellement. Il est à noter que dans ces expériences, l'héritage maternel ou paternel du transgène *P(lArB)* n'intervient qu'à la première génération, par la suite le transgène est toujours hérité maternellement aux générations suivantes (Josse *et al*, 2007).

Nous avons reproduit ces résultats plusieurs fois. En moyenne il faut quatre générations pour que le pourcentage de répression issu d'un héritage paternel du transgène à la première génération atteigne le pourcentage de répression issu d'un héritage maternel du transgène. L'analyse de small RNA-seq confirme que *P(lArB)* produit des piRNA dans la lignée *P-1152* qui porte 2 copies du transgène à l'état homozygote, tandis que la lignée de référence *Canton*, qui ne porte pas le transgène ni de subtélomère du chromosome *X* ne produit, par conséquent, par de piRNA dérivés de *P(lArB)*. Ces analyses révèlent également que lorsque le transgène *P(lArB)* est hérité maternellement, il produit encore des piRNA aux générations suivantes. On peut noter que la quantité de piRNA produit par le transgène diminue progressivement de la 1<sup>ère</sup> à la 4<sup>ème</sup> génération, passant de 40000 à 20000 reads environ. Cette diminution peut s'expliquer par le fait qu'on passe de l'état homozygote à la génération 0 à l'état hétérozygote pour *P(lArB)* à la génération 1. A l'inverse, lorsque le transgène *P(lArB)* est hérité paternellement, la quantité de piRNA produit augmente jusqu'à la 4<sup>ème</sup> génération pour atteindre une quantité de piRNA comparable à celle de la 4<sup>ème</sup> génération pour un héritage maternel (environ 20000 reads) (Article II). Ces résultats indiquent que lorsque *P(lArB)* est hérité paternellement à la première génération, il faut plusieurs générations pour que le transgène acquière complètement les caractéristiques moléculaires d'un piRNA cluster au sein d'une population. Dans les subtélomères du chromosome *X*, on trouve une séquence de 0.9 kb unique dans le génome appelée *T3* (Asif-Laidin *et al*, 2017). On constate que *T3*, qu'il soit

hérité maternellement ou paternellement, est capable de réprimer l'expression d'un transgène rapporteur *pRFP-T3* et produit des piRNA en quantités comparables dès la première génération. De manière étonnante, il y a plus de transcrits de *P(lArB)* et de *T3* lorsque ceux-ci sont hérités paternellement comparé à un héritage maternel à la première génération. Ces résultats indiquent que la quantité de transcrits ne reflète pas nécessairement la quantité de piRNA produits. On peut également noter la grande différence de taille pour ces séquences, 0.9 kb pour *T3* et 18.3 kb pour *P(lArB)*. Ces résultats suggèrent que la taille de la séquence influence l'efficacité de la conversion en une génération ; ainsi plus la séquence est petite, plus la conversion est efficace. Selon ce principe une séquence de petite taille comme *T3* est susceptible d'être convertie dans la majorité des cas (plus de 95%) en une seule génération alors qu'une séquence de grande taille comme *P(lArB)*, difficile à convertir en une génération, n'est alors convertie que dans une minorité de cas (environ 10%). Le transgène *P(lArB)* et *T3* ne sont séparés que de 400 pb, pourtant il apparaît de manière inattendue que leurs propriétés sont très différentes. *P(lArB)* produit des piRNA en sens et en antisens dont la formation semble dépendante de Rhino et de Moon, puisqu'en l'absence de ces facteurs le pourcentage de répression d'un transgène cible à la première génération est faible. En revanche, *T3* produit majoritairement des piRNA sens dont la formation ne dépend ni de Rhino ni de Moon, l'absence de ces facteurs n'affectant pas le pourcentage de répression d'un transgène cible.

Le cluster *BX2* possède des séquences communes avec le transgène *P(lArB)*, à savoir *lacZ* et le plasmide *pBR322*, mais également avec un autre transgène appelé *P(RS3)*, dérivé de l'élément *P* et inséré dans les séquences subtélomériques du chromosome *3R*, à savoir *white*. Lorsque les piRNA provenant des transgènes *P(lArB)* ou *P(RS3)* sont hérités maternellement, ils permettent la conversion des séquences communes au niveau de *BX2* dès la première génération : ainsi, l'héritage des piRNA de *P(lArB)* entraîne la production de piRNA pour *lacZ* et *pBR322* mais pour *white* on détecte peu de piRNA ; à l'inverse l'héritage des piRNA de *P(RS3)* entraîne la production de piRNA par *white* mais peu de production pour *lacZ* et *pBR322*. A la septième génération, l'ensemble du cluster *BX2* produit des piRNA de manière homogène, indiquant qu'il faut plusieurs générations pour convertir totalement le cluster *BX2* dans une population. Nous savons que *BX2<sup>OFF</sup>* ne peut pas être converti en l'absence de piRNA maternels, or avec ces résultats nous pouvons constater que dans certains cas des séquences qui ne sont pas directement ciblées par des piRNA homologues peuvent être converties pour la production de piRNA, même si l'efficacité de la conversion est faible.

L'ensemble de ces résultats nous permet d'émettre une hypothèse quant à la conversion de *P(lArB)* hérité paternellement. Les piRNA ciblant les régions flanquantes de *P(lArB)* seraient impliqués dans la conversion des séquences non ciblées par des piRNA homologues, puisque non hérités maternellement. La valeur du pourcentage de répression à la première génération issue de l'héritage paternel de *P(lArB)* nous permet de dire que la conversion par les séquences adjacentes n'intervient que dans une minorité de cas. Aux générations suivantes, *P(lArB)* peut être converti de deux manières : soit grâce à un héritage de piRNA maternels produit par le transgène converti à la génération précédente ; soit grâce à la conversion dépendante des piRNA ciblant les séquences adjacentes. Ainsi, lors de l'insertion d'un nouvel ET dans un piRNA cluster, les piRNA des séquences flanquantes pourraient convertir l'élément en une seule génération ; l'efficacité de la conversion du nouvel ET semblant dépendre de sa taille. Une fois convertie, les piRNA produits par l'ET peuvent être transmis maternellement aux générations suivantes et convertir la séquence à chaque génération selon le processus de paramutation, nécessaire pour le maintien de la régulation de toutes les copies de l'ET dans le génome (Brennecke *et al*, 2007; de Vanssay *et al*, 2012).

Une étude révèle que l'insertion d'un nouvel ET dans un piRNA cluster est un événement relativement fréquent ; c'est le cas notamment pour le rétrotransposon *ZAM* pour lequel plusieurs événements d'insertions indépendants dans des piRNA clusters germinaux ont été suggérés grâce à l'étude de sa régulation (Duc *et al*, 2019). Des modèles suggèrent également qu'il existe des contraintes évolutives sur la taille des piRNA clusters. La taille cumulée de l'ensemble des piRNA clusters dans un génome ne doit pas être inférieure à 0.2% du génome pour permettre une régulation efficace des ET, sous ce seuil les risques d'extinction de la population sont élevés (Kofler, 2020).

### **3. Structures des *TAS* et régulation du génome**

Le subtélomère du chromosome *X* (ou *TAS-X*) est un piRNA cluster (cluster *IA*) constitué d'ET mais aussi d'autres domaines (*T1*, *T2* et *T4*) qu'on retrouve également répétés au niveau des subtélomères des autres chromosomes et d'un domaine unique dans le génome appelé *T3* (Asif-Laidin *et al*, 2017). Laure Teyssset a entrepris des expériences de séquençage avec la technologie MinION afin de mettre en évidence la structure exacte du *TAS-X* et les premières analyses révèlent des structures, notamment pour le domaine *T3*, qui contiennent plusieurs

répétitions et de nombreux fragments d'ET typiques des piRNA clusters (Laure Teyssset, communication personnelle).

En comparant des lignées de laboratoire et des lignées issues de la nature, Amna a pu constater qu'une partie des lignées de laboratoire a perdu le *TAS-X* alors que toutes les lignées de la nature possèdent le *TAS-X* (Asif-Laidin *et al*, 2017). Ces données suggèrent que le *TAS-X* n'est pas indispensable pour les lignées qui vivent en laboratoire. Avant son départ du laboratoire, Amna a réalisé des expériences de RNA-seq sur des lignées possédant des *TAS-X* et d'autres n'en possédant pas. Les analyses montrent que de nombreux gènes sont dérégulés entre ces deux conditions (avec ou sans *TAS*). Ces données posent la question du rôle des séquences *TAS*, en particulier le *TAS-X* et de l'implication des piRNA dans ces fonctions.

## VI. Comprendre la biologie des piRNA clusters

La biologie des piRNA est complexe et il reste de nombreux aspects à aborder. Les différents outils génétiques que nous possédons nous permettent d'appréhender divers aspects de la biologie des piRNA clusters relatifs à leur formation, leur régulation ou encore leur dynamique.

Définir précisément les piRNA clusters est difficile mais pourtant essentiel si l'on veut aller plus avant dans leur compréhension. Leur dynamique apparaît dépendante de très nombreux facteurs, aussi bien environnementaux (avec l'exemple de la chaleur) que génétiques (*hfp* et *JHDM2* entre autres). Par ailleurs, l'ensemble des travaux tend à dévoiler des différences de régulation subtiles entre les piRNA clusters. Nous savons depuis plusieurs années que la biologie des piRNA clusters somatiques et germinaux est différente (Brennecke *et al*, 2007; Goriaux *et al*, 2014). Au sein même des piRNA clusters germinaux, des travaux semblent indiquer que la régulation des piRNA clusters péricentromériques (*42AB* par exemple) et subtélomériques (*TAS*) diffère également, certains mutants n'affectant pas de la même façon la production de piRNA et les marques chromatiniennes de ces deux types de clusters (Komarov *et al*, 2020; Radion *et al*, 2019). De manière inattendue, nos travaux semblent pointer des différences de régulation au sein d'un même piRNA cluster. En effet, au niveau du subtélomère du chromosome X, le transgène *P(lArB)* et *T3* ont des propriétés différentes : ces deux séquences n'ont ni les mêmes propriétés, ni la même dynamique (Article II). Ces

différences de propriétés devront être prises en compte car elles pourraient révéler des caractéristiques fonctionnelles.

Les ET, de par leurs propriétés, structurent les génomes et participent activement à l'évolution, mais leur activité est également associée au développement de nombreuses maladies et à des phénomènes de dysgénésies gonadiques. Ainsi, le rôle des piRNA et des piRNA clusters est de contrôler finement cette mobilité afin de maintenir l'intégrité du génome au cours des générations d'une part et permettre la survie des espèces d'autre part. Pour autant, le contrôle des ET ne doit pas être trop strict pour permettre aux ET de créer de la diversité génétique. En ce sens, les piRNA clusters sont des séquences qui participent indirectement à l'évolution des génomes et pourraient être impliqués dans la spéciation (Casier *et al*, 2019a; Kotov *et al*, 2019; Parhad *et al*, 2017). Des études comparatives et des simulations sur la dynamique des ET et des piRNA clusters semblent indiquer que les ET et les piRNA clusters co-évoluent (Luo *et al*, 2020). Ces données ainsi que la structure des piRNA clusters, constitués de fragments d'ET, amènent à se questionner sur l'origine des piRNA clusters qui pourrait être commune à celle des ET.

Maintenir l'intégrité des génomes et empêcher le développement de maladies ou d'anomalies génétiques impliquent de maintenir l'intégrité des piRNA clusters, grâce à des facteurs auxquels *JHDM2* pourrait appartenir. En effet, une production aberrante de piRNA par des séquences qui ne sont pas définies comme des piRNA clusters en contexte sauvage pourrait induire des phénomènes de létalité embryonnaire, comme le suggèrent les phénotypes observés chez les descendants des *JHDM2 GLKD*. Des études révèlent également une corrélation entre la dérégulation des nombreux piRNA et le développement de plusieurs types de cancers (Law *et al*, 2013; Martinez *et al*, 2015; Mei *et al*, 2013; Ng *et al*, 2016; Yin *et al*, 2017). De ce fait, les piRNA clusters doivent être régulés strictement pour ne pas entraîner, *via* l'activité des piRNA, la dérégulation de gènes essentiels aussi bien pour la génération en cours que pour la (ou les) génération(s) suivante(s).



# **MATERIEL ET METHODES**



## I. Ressources et outils d'analyses

Au laboratoire et parmi la communauté des drosophilistes, nous disposons de nombreux outils et ressources nous permettant d'avancer dans nos recherches.

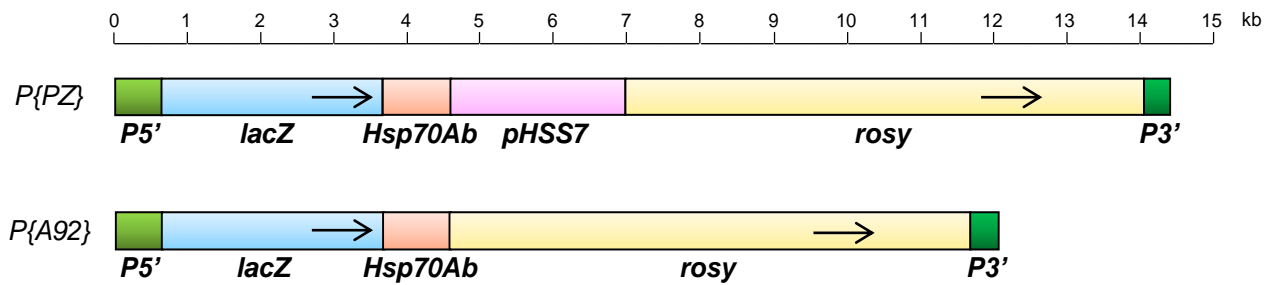
- La base de données **FlyBase** (<https://flybase.org/>) répertorie de nombreuses informations concernant le modèle drosophile, incluant le séquençage complet et annoté des génomes de plusieurs espèces de drosophiles. FlyBase donne ainsi accès à de nombreuses données pour chaque gène : séquences, fonctions, profils d'expression, allèles/lignées disponibles, phénotypes des mutants, interactions moléculaires, références bibliographiques...
- Le '**Bloomington Drosophila Stock Center**' (BDSC) de l'université d'Indiana aux Etats-Unis met à disposition de nombreuses lignées de drosophiles (lignées de références, RNAi, mutants perte ou gain de fonction, balanceurs, transgènes...).
- **Fasteris** (<https://www.fasteris.com>) est une entreprise suisse avec laquelle nous travaillons qui réalise les séquençages à haut-débit dont nous avons besoin pour différents types d'expériences (RNA-seq, small RNA-seq, ChIP-seq...).
- La plateforme **ARTBio** de l'Institut de Biologie Paris-Seine, dont l'URM7622 fait partie, développe et met à disposition des outils d'analyses bioinformatiques *via* le serveur Galaxy. Ces outils nous permettent d'analyser de nombreuses données issues des techniques de séquençage à haut-débit (RNA-seq, small RNA-seq, ChIP-seq...).

## II. Génétique

### 1. Balanceurs et transgènes

Un chromosome balancier est un chromosome portant des inversions multiples ainsi qu'un marqueur phénotypique généralement dominant et qui est létal à l'état homozygote. Cet outil génétique nous permet de suivre et de garder des insertions ou des mutations à l'état hétérozygote. Les chromosomes balanceurs suivants ont été utilisés :

- **Cy** est un balancier du chromosome 2 qui donne un phénotype d'ailes recourbées,
  - **CyRoi** est un balancier du chromosome 2 avec un phénotype d'ailes recourbées et d'œil rugueux,
  - **Sb** est un balancier du chromosome 3 avec un phénotype de soies courtes et rugueuses,
  - **Xa** est un balancier des chromosomes 2 et 3, pour lequel les chromosomes 2 et 3 sont attachés, et donnant un phénotype d'encoche au bout des ailes,
  - **hu** est un marqueur de chromosome balancier du chromosome 3 pour lequel le nombre de soies sur l'épaule est plus important,
  - **Ubx** est un marqueur de chromosome balancier du chromosome 3 pour lequel on observe des altères plus grosse portant des soies.
- 
- Le transgène **P[lacW]** est dérivé de l'élément *P* ; il contient les pieds de *P* aux extrémités 5' et 3' ainsi que le gène *lacZ*, une partie du gène *lacY*, le gène *miniwhite* (construction ne contenant que les exons du gène *white*), le terminateur du gène *Hsp70Bb* et du plasmide *pBR322*. Le transgène mesure 10691 paires de bases (Voir figure 26A). FlyBase ID : FBtp0000204.
  - Le cluster **BX2** est constitué de sept copies du transgène *P[lacW]* répétées en tandem, est inséré dans un intron du gène *Ago1* et en amont des gènes *mRpL53* et *CG3155*, sur le chromosome 2 (site cytologique 50C) (Voir figure 26B). On estime la taille de *BX2* à environ 75 kb (10691 x 7 copies de *P[lacW]*). Ce cluster a été obtenu grâce à des événements de transposition successifs du transgène *P[lacW]* chez des drosophiles possédant une mutation du gène *white* endogène (Dorer & Henikoff, 1994).
  - Le transgène **P(lArB)** est dérivé de l'élément *P*. Il contient les pieds de *P* en 5' et en 3', le plasmide *Bluescript*, ainsi que les gènes *rosy*, *Adh* et *lacZ*. Les terminateurs de transcription de *Adh* et *lacZ* correspondent respectivement aux terminateurs des gènes *CG8231* et *Hsp70Ab* (Voir figure 25A).
  - Les transgènes dits « cibles » **P(PZ)** et **P(A92)** sont deux transgènes rapporteurs dérivés de l'élément *P*, contenant les gènes *lacZ* et *rosy* (Figure 42). Ces transgènes sont aussi appelés *P(TARGET)*. Flybase ID *P(PZ)* : FBtp0000210 ; *P(A92)* : FBtp0000154.



**Figure 42: Structures des transgènes « cibles » (aussi appelés  $P(TARGET)$ ).** Les transgènes  $P(PZ)$  et  $P(A92)$  sont dérivés de l'élément  $P$  ; ils contiennent les pieds de  $P$  aux extrémités 5' et 3', les gènes  $lacZ$  et  $rosy$  et le terminateur du gène  $Hsp70Ab$ . Le transgène  $P(PZ)$  possède en plus une séquence du plasmide  $pHSS7$ . Les flèches représentent le sens de transcription des gènes  $lacZ$  et  $rosy$ . L'échelle représentée est en kb.

## 2. Lignées

- $w^{1118}$  est une lignée de référence pour laquelle le gène *white*, impliqué dans la pigmentation de l'œil, porte une mutation entraînant la perte de fonction du gène. FlyBase ID : FBal0018186.
- La lignée « *nosGAL4* » contient un transgène inséré sur le chromosome 3 dérivé de l'ET *PiggyBac* où la séquence du gène *gal4* est sous le contrôle du promoteur *nanos*, permettant une expression de GAL4 spécifiquement dans la lignée germinale. L'insertion est viable à l'état homozygote. FlyBase ID : FBti0131635, BDSC #32180.
- Les lignées **BX2** ( $BX2^{OFF}$  et  $BX2^{ON}$ ) portent le cluster *BX2* inséré dans le premier intron du gène *Ago1* et en amont des gènes *mRpL53* et *CG3155*, sur le chromosome 2 (site cytologique 50C). Ces lignées sont hétérozygotes pour le cluster *BX2*, qui est balancé par un chromosome *Cy* car les individus homozygotes sont très peu viables et stériles. FlyBase ID : FBti0016766.
- La lignée **P-1039** (aussi appelée  $P(TARGET)^{GS}$  dans l'article I) porte un transgène rapporteur  $P(PZ)$  sur le chromosome 2, inséré dans le gène *eEF5* (site cytologique 60B). Cette lignée est hétérozygote pour le transgène  $P(PZ)$  qui est balancé par un chromosome *Cy* car les individus homozygotes pour cette insertion ne sont pas

viables. Dans cette lignée, le transgène *P(PZ)* s'exprime à la fois dans les cellules germinales et somatiques de l'ovaire. Flybase ID : FBst0011039, BDSC #11039.

- La lignée **BQ16** (aussi appelée *P(TARGET)<sup>G</sup>* dans l'article I) porte un transgène rapporteur *P(A92)* sur le chromosome 3, inséré dans le gène *shep* (site cytologique 64C). La lignée est viable à l'état homozygote pour le transgène *P(A92)* mais la lignée peut également être balancée par un chromosome *Sb*. Dans la lignée, le transgène *P(A92)* s'exprime uniquement dans les cellules germinales de l'ovaire. FlyBase ID : FBti0003435.
- La lignée **BA37** (aussi appelée *P(TARGET)<sup>S</sup>* dans l'article I) porte un transgène rapporteur *P(A92)*. Les publications dans lesquelles la lignée est référencée indiquent que le transgène est localisé sur le chromosome 2 (Audibert & Simonelig, 1999; Higuet *et al*, 1992; Lemaitre *et al*, 1993; Marin *et al*, 2000). Parce que le site d'insertion du transgène n'était pas clairement établi, Laure Teyssset l'a localisé par PCR inverse. Elle a découvert que le transgène n'était pas localisé sur le chromosome 2 mais sur le chromosome 3, inséré dans le gène *kmr* (site cytologique 87F). Les individus homozygotes pour le transgène *P(A92)* ne sont pas viables. La lignée est balancée par un chromosome balanceur qui n'est plus utilisé actuellement ; il s'agit d'une fusion des chromosomes 2 et 3 qui porte le marqueur phénotypique du chromosome *Cy*, ce qui peut expliquer la mauvaise interprétation de sa localisation dans les publications précédentes. Dans la lignée, le transgène *P(A92)* s'exprime uniquement dans les cellules somatiques de l'ovaire. FlyBase ID : FBti0003418.
- Les lignées **BX2**, **P-1039** ; **BX2**, **BQ16** ; **BX2**, **BA37** et **BX2**, **nosGAL4** ont été construites à partir des lignées d'origine (**BX2**, **P-1039**, **BQ16**, **BA37** et **nosGAL4**). La séquence commune entre **BX2** et les transgènes cibles est située entre la position 87 du pied 5' de *P* et la fin du gène *lacZ*, soit une région de 3547 paires de bases.
- Les lignées **hfp<sup>9</sup>** et **hfp<sup>13</sup>** portent un allèle hypomorphe pour le gène *hfp*, localisé sur le chromosome 3. Ces allèles ont été obtenus par excision de l'élément *P* dans la région 5'UTR de *hfp* (Van Buskirk & Schüpbach, 2002). Les homozygotes *hfp<sup>13/13</sup>* *hfp<sup>13</sup>* sont viables et fertiles, mais l'allèle est maintenu à l'état hétérozygote par un balanceur marqué par *Ubx* ou *hu*. Les homozygotes *hfp<sup>9/9</sup>* *hfp<sup>9</sup>* sont stériles, la lignée est donc maintenue grâce à un balanceur marqué par *Ubx*.

- La lignée **BX2**, *hfp-GFP* a été construite à partir d'une lignée *BX2<sup>OFF</sup>* (balancée par *Cy*) et d'une lignée *hfp-GFP* obtenue par le laboratoire de Trudi Schüpbach (Howard Hughes Medical Institute, Princeton) (Van Buskirk & Schüpbach, 2002) pour laquelle un transgène GFP est en phase avec la séquence codante de *hfp*. La protéine de fusion générée fait environ 95 kDa. La lignée est viable et fertile.
- La lignée **JHDM2<sup>KO</sup>** a été obtenue par recombinaison homologue, en remplaçant l'ORF de *JHDM2* par une cassette contenant des sites *loxP*, un promoteur *PAX3* ainsi que des marqueurs *RFP* et *kanamycine*. Cette lignée nous a été offerte par Michael Buszczak (University of Texas Southwestern Medical Center, Dallas) (Shalaby *et al*, 2017). Les homozygotes *JHDM2<sup>KO</sup>* sont stériles, l'allèle est donc maintenu grâce à un balancier *hu*. Dans l'article de référence, la lignée à l'état homozygote est considérée comme viable et fertile car les individus pondent des œufs mais il apparaît que les auteurs n'ont pas vérifiés si ces œufs étaient capables de donner des larves (Michael Buszczak, communication personnelle).
- **Canton** est une lignée de référence dépourvue d'élément *P* et de subtélomère sur le chromosome *X* (Asif-Laidin *et al*, 2017). La lignée du laboratoire (**Canton Y**) est dérivée de la lignée classique, elle porte une mutation qui affecte l'expression du gène *yellow* impliqué dans la pigmentation du corps.
- La lignée **P-1152** contient deux transgènes *P(lArB)* insérés dans les séquences *invader-4* du subtélomère du chromosome *X* (Site cytologique 1A) (Voir figure 25B). Cette lignée est mutante pour le gène *white* impliqué dans la pigmentation de l'œil. La lignée est viable et fertile à l'état homozygote.
- Les lignées RNAi (aussi appelées lignées TRiP pour Transgenic RNAi Project) que nous utilisons viennent pour la plupart du BDSC : *white* (#33466), *hfp* (#34785), *JHDM2* (#32975) et *Rhino* (#34071). Deux lignées RNAi de *moon* nous ont été offertes par Julius Brennecke (Institute of Molecular Biotechnology, Vienne) (Andersen *et al*, 2017). Pour toutes ces lignées RNAi, les transgènes permettant l'expression de shRNA sont sous le contrôle de séquences UAS et nécessitent la production de la protéine GAL4 pour être transcrit. Ces transgènes sont insérés soit sur le chromosome 2, soit sur le chromosome 3 ; ainsi dans certaines lignées, ces

transgènes peuvent être balancés par des chromosomes balanceurs des chromosomes 2 et 3.

### **3. Croisement et maintien des stocks**

Les croisements se font à une température de 25°C. Les stocks peuvent être maintenus à 18°C, à 25°C ou à température ambiante. Dans le cas particulier des lignées contenant le cluster *BX2*, les stocks sont maintenus à 25°C car à cette température l'état épigénétique du transgène est stable au cours des générations.

## **III. Histochimie**

### **1. Coloration $\beta$ -galactosidase**

Les femelles sont mises à nourrir deux à trois jours sur un milieu de ponte puis disséquées manuellement dans du PBS 1X froid. Les ovaires disséqués sont ensuite fixés 10 min (PBS 1X et glutaraldéhyde 0.5%), rincés trois fois dans du PBS 1X froid et incubés dans une solution de coloration (PBS 1X, Ferri-Ferro cyanide 3.5 mM et X-gal 0.15%) à 37°C sur la nuit. Le lendemain, la solution de coloration est retirée et les ovaires sont conservés à 4°C dans une solution de glycérol/éthanol (50/50) avant montage.

### **2. Analyse des colorations**

Pour analyser les colorations  $\beta$ -galactosidase, les ovaires sont montés entre lame et lamelle dans une solution de glycérol/éthanol (50/50). Pour chaque ovaire, on calcule le pourcentage de répression du gène *lacZ* dans les cellules germinales en divisant le nombre de chambres ovariennes présentant de la répression (pas de coloration des noyaux des cellules germinales) par le nombre total de chambres ovariennes. Dans un ovaire, on compte entre 16 et 18 ovarioles chacun constitué de 3 à 4 chambres ovariennes facilement analysables (pour lesquelles on peut voir nettement la coloration ou non des noyaux), ce qui représente un nombre moyen de 60 chambres ovariennes analysables par ovaires. Les photos sont prises au microscope Axio-ApoTome et traitées grâce au logiciel ZEN2.

## IV. Biologie Moléculaire

### 1. Extraction d'ARN

Pour chaque échantillon, 10 paires d'ovaires provenant de femelles bien nourries (2 à 3 jours) sont disséquées manuellement dans du PBS 1X froid, congelées dans de l'azote liquide et conservées à -80°C jusqu'à l'extraction d'ARN. Les ovaires sont broyés dans du TRIzol puis les ARN totaux sont séparés avec du chloroforme, précipités dans de l'éthanol 100%, rincés trois fois dans de l'éthanol 80% et resuspendus dans de l'eau bidistillée. Pour chaque génotype, entre trois et six échantillons biologiques sont utilisés.

### 2. RT-qPCR et RT-PCR

Pour chaque échantillon (entre 3 et 6 échantillons biologiques), 2 µg d'ARN totaux sont traités à la DNase (fermentas) et 1 µg d'ARN DNasé est reverse transcrit en utilisant des random hexamer primers (fermentas). Les échantillons sont analysés par PCR classique ou par PCR quantitative (qPCR).

Les PCR sont réalisées avec les échantillons non dilués et analysées par électrophorèse sur gel d'agarose. *RpL32* est utilisé comme gène contrôle.

Les qPCR sont réalisées en triplicat pour chaque échantillon et le gène *RpL32* est utilisé comme gène de référence, sauf indication contraire. Une série de dilutions obtenues à partir d'un mix des différentes RT est utilisée pour normaliser la quantité de transcrits pour chaque échantillon permettant de générer des valeurs de quantité standard (Sq pour standard quantity). Les variations entre les triplicats techniques sont négligeables comparées aux variations entre les répliquats biologiques. La moyenne des triplicats techniques (meanSq) est donc utilisée. Pour chaque échantillon biologique, le ratio  $\text{meanSq}(\text{gene})/\text{meanSq}(\text{RpL32})$  est calculé pour normaliser les quantités standards. Enfin, la moyenne des ratios est calculée pour comparer les différentes conditions.

### **3. Small RNA-sequencing & RNA-sequencing**

Pour chaque échantillon, 20 paires d'ovaires provenant de femelles bien nourries (2 à 3 jours) sont disséquées manuellement dans du PBS 1X froid, congelées dans de l'azote liquide et conservées à -80°C. Les ovaires sont broyés dans du TRIzol puis les ARN totaux sont séparés avec du chloroforme et précipités dans de l'éthanol 100%. Pour les small RNA-seq, les échantillons sont envoyés dans l'éthanol 100% à Fasteris qui réalise les bibliothèques et le séquençage (ce volume représente environ 10 µg d'ARN totaux). En revanche, pour les RNA-seq, les ARN sont culottés et rincés 3 fois dans l'éthanol 80% et resuspendus dans de l'eau bidistillée. Pour chaque échantillon, 2 µg d'ARN totaux sont traités à la DNase et 1 µg d'ARN DNase est utilisé pour le séquençage.

Pour la préparation des bibliothèques, les ARN totaux sont séparés en fonction de leur taille par migration sur gel de polyacrylamide dénaturant, les fragments inférieurs à 30 nucléotides sont gardés pour le small RNA-seq tandis que les ARN de plus de 100 nucléotides sont conservés pour le RNA-seq. Les données brutes sont récupérées au format FastQ et analysées grâce à la plateforme Galaxy. Les étapes de préparation des bibliothèques, de séquençage et d'analyse sont décrites de manière détaillée dans l'article I.

### **4. RNA-IP**

Pour chaque échantillon, 80 paires d'ovaires provenant de femelles bien nourries (2 à 3 jours) sont disséquées manuellement dans du PBS 1X froid, congelées dans de l'azote liquide et conservées à -80°C. Les ovaires sont broyés dans un tampon de lyse (KCl 150 mM, Tris pH7.4 25 mM, EDTA 5 mM, DTT 0.5 mM, Nodipet P-40 0.5%), supplémenté avec un inhibiteur de protéase (Roche #11697498001, 1 comprimé pour 10 mL de tampon de lyse) et un inhibiteur de RNase (40U pour 1 mL de tampon de lyse). Les lysats sont soniqués grâce au sonicateur Bioruptor (Diagenode) pendant 7 minutes (15 sec ON, 60 sec OFF) à intensité moyenne et le surnageant est récupéré après une centrifugation de 15 minutes à 12000 rpm (4°C). 10% du lysat est réservé pour servir d'input et le reste est incubé à 4°C avec 3 µg d'anticorps anti-GFP (Roche #11814460001) pendant trois heures sous agitation. Le lysat immunoprécipité est ensuite transféré dans un tube contenant 30 µl de billes magnétiques couplées à la protéine G (Invitrogen), lavées préalablement deux fois dans du tampon de lyse, et le tout est incubé pendant une heure à 4°C sous agitation. Les billes sont lavées 5 fois



pendant 10 minutes à 4°C dans du tampon de lyse pour éliminer les complexes protéines-anticorps non fixés. Les ARN totaux sont extraits comme décrit dans la partie ‘extraction d’ARN’ à la différence que l’extraction ne nécessite pas de broyage, le TRIzol est directement ajouté sur le lysat (pour les inputs) ou sur les billes lavées associées aux complexes protéine-anticorps. Pour chaque input, 2 µg d’ARN totaux sont traités à la DNase et 1 µg d’ARN DNasé est reverse transcrit tandis que pour les immunoprécipitations, tous les ARN totaux sont traités à la DNase et reverse transcrits en utilisant des random hexamer primers. Les qPCR sont réalisées en duplicat pour chaque échantillon biologique pour obtenir les valeurs de Ct (Cycle threshold). Les variations entre les duplicats techniques sont négligeables comparées aux variations entre les répliquats biologiques, de ce fait les moyennes des répliquats techniques sont utilisées pour l’analyse des données. Celles-ci sont analysées comme décrit par Miriam Gagliardi and Maria Matarazzo (Gagliardi & Matarazzo, 2016). Pour chaque échantillon, la fraction immunoprécipitée (IP) est normalisée par rapport à son Input pour tenir compte des différences de préparation des échantillons :  $\Delta Ct \text{ [normalized RIP]} = (Ct \text{ [RIP]} - (Ct \text{ [Input]} - \text{Log}_2 (\text{Input Dilution Factor})))$  où ‘Ct [RIP]’ correspond à la valeur de Ct mesurée pour l’IP, ‘Ct [Input]’ correspond à la valeur de Ct mesurée pour l’Input et ‘Input Dilution Factor’ correspond à la fraction du lysat mis de côté pour l’Input (ici, l’Input représente 10% du lysat, par conséquent l’Input Dilution Factor est de 10). A partir des moyennes de  $\Delta Ct$ , le pourcentage d’Input est calculé pour chaque IP :  $\% \text{Input} = 2^{(-\Delta Ct \text{ [normalized RIP]})}$ . Le pourcentage d’Input ne tient pas compte des contrôles négatifs d’IP qui rendent compte des interactions non spécifiques. Il est donc possible d’ajuster ces valeurs de  $\Delta Ct$  par rapport au(x) contrôle(s) négatif(s) :  $\Delta \Delta Ct \text{ [RIP/NS]} = \Delta Ct \text{ [normalized RIP]} - \Delta Ct \text{ [normalized NS]}$  où ‘RIP’ correspond à l’IP et ‘NS’ (pour Non Spécifique) aux interactions aspécifiques estimées grâce au(x) contrôle(s) négatif(s). Enfin, l’enrichissement par rapport à l’Input est calculé à partir des valeurs :  $\text{Fold Enrichment} = 2^{(-\Delta \Delta Ct \text{ [RIP/NS]})}$ . Dans les calculs de %Input et d’enrichissement, la valeur 2 correspond à une efficacité des amorces de qPCR égale ou supérieure à 100%.

## 5. Western Blot

Les expériences de western blot sont réalisées en complément des expériences de RNA-IP afin de vérifier que les protéines de fusion sont immunoprécipitées. Des extraits protéiques avant et après immunoprécipitation sont conservés à -80°C. Les protéines sont dénaturées par

ajout de Laemmli 2X. Les échantillons sont chargés sur un gel de polyacrylamide 10% (Biorad) dans un tampon de migration (25 mM Tris HCl pH 8.0, 192 mM Glycine, 0.1% SDS) et migrent pendant 1h30 à 80V. Le gel est récupéré et les protéines sont transférées sur une membrane PVDF pendant une heure à 80V. Une coloration au rouge Ponceau permet de valider l'étape de transfert en visualisant les protéines sur la membrane. La membrane est saturée dans une solution de lait à 5% dans du Tris-Buffered Saline Tween (TBST: 500 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Tween) pendant une heure puis lavée trois fois avec du tampon TBST. La membrane est ensuite incubée avec un anticorps primaire anti-GFP de souris (Roche #11814460001, dilution 1:10000) pendant la nuit à 4°C. La membrane est lavée trois fois avec un tampon TBST (environ 5 min) avant d'être incubée une heure à température ambiante avec un anticorps secondaire anti-souris (Promega #W4021, dilution 1:10000). L'anticorps secondaire est couplé à une phosphatase alcaline qui réagit avec le NBT (Nitro Blue Tetrazolium Chloride) et de BCIP (5-brom-4-chloro-3'-indolyphosphate p-toluidine salt), utilisés pour la révélation de la membrane.

## **6. ChIP-qPCR**

Pour chaque échantillon, 50 paires d'ovaires provenant de femelles bien nourries (2 à 3 jours) sont disséquées manuellement dans du PBS 1X froid. Les ovaires sont ensuite fixés dans une solution de formaldéhyde 1.8% (formaldéhyde 37%, PBS 1X). La fixation est arrêtée par ajout de glycine et les tissus rincés deux fois dans du PBS 1X froid avant d'être congelées dans de l'azote liquide et conservées à -80°C. Les ovaires sont broyés puis lysés dans un tampon de lyse (ChIP lysis buffer : NaCl 100 mM, Tris pH8 50 mM, EDTA 5 mM, 1% SDS). Les lysats sont soniqués à l'aide d'un sonicateur Bioruptor (Bioruptor Standard Diagenode) trois fois pendant 15 minutes (30 sec ON, 30 sec OFF) à forte intensité et la chromatine non soniquée est éliminée par centrifugation. 5% du lysat est mis de côté pour servir d'input et le reste est utilisé pour l'immunoprécipitation. Le lysat est incubé avec 3 µg d'anticorps anti-H3K9me3 (Active Motif #39161) ou un contrôle pour mesurer les interactions aspécifiques (IgG de lapin, Diagenode Low Cell ChIP-kit Cat#803-015) sous rotation à 4°C pendant la nuit. Des billes magnétiques couplées à la protéine G (Invitrogen) sont lavées deux fois dans un tampon de lavage (Low Salt Wash Buffer : 0.1% SDS, 1% Triton X-100, EDTA 2 mM, Tris pH8 20 mM, NaCl 150 mM), transférées dans les lysats immunoprécipités et incubées 1h à 4°C sous rotation. Les billes sont ensuite lavées deux fois dans du Low Salt Wash Buffer,

une fois dans du High Salt Wash Buffer (0.1% SDS, 1% Triton X-100, EDTA 2 mM, Tris pH8 20 mM, NaCl 500 mM), deux fois dans du LiCl Wash Buffer (LiCl 0.25 M, 1% NP-40, 1% Na deoxycholate, EDTA 1 mM, Tris pH8 10 mM) et deux fois dans du TE Buffer (Tris pH8 10 mM, EDTA 1 mM). Le dépontage de la chromatine, l'élution et la purification de l'ADN se font grâce au kit IPure (Diagenode Cat# C03010015) comme décrit dans le manuel fourni par le fabricant ([https://www.diagenode.com/files/products/kits/ipure\\_kit\\_v2\\_manual.pdf](https://www.diagenode.com/files/products/kits/ipure_kit_v2_manual.pdf)).

Les qPCR sont réalisées en duplicat pour chaque échantillon biologique pour obtenir les valeurs de Ct (Cycle threshold). Les variations entre les duplicats techniques sont négligeables comparées aux variations entre les réplicats biologiques, de ce fait les moyennes des réplicats techniques sont utilisées pour l'analyse des données (meanCt). Pour chaque échantillon, la fraction immunoprécipitée (IP) est normalisée par rapport à son Input pour tenir compte des différences de préparation des échantillons :  $\Delta Ct [\text{normalized ChIP}] = (Ct [\text{ChIP}] - (Ct [\text{Input}] - \text{Log}_2 (\text{Input Dilution Factor})))$  où 'Ct [ChIP]' correspond à la valeur de Ct mesurée pour l'IP, 'Ct [Input]' correspond à la valeur de Ct mesurée pour l'Input et 'Input Dilution Factor' correspond à la fraction du lysat mis de côté pour l'Input (ici, l'Input représente 5% du lysat, par conséquent l'Input Dilution Factor est de 20). A partir des moyennes de  $\Delta Ct$ , le pourcentage d'Input est calculé pour chaque IP :  $\% \text{Input} = 2^{(-\Delta Ct [\text{normalized ChIP}])}$ . Le pourcentage d'Input ne tient pas compte du contrôle négatif d'IP qui rendent compte des interactions non spécifiques. Il est donc possible d'ajuster ces valeurs de  $\Delta Ct$  par rapport au(x) contrôle(s) négatif(s) :  $\Delta \Delta Ct [\text{ChIP /NS}] = \Delta Ct [\text{normalized ChIP}] - \Delta Ct [\text{normalized NS}]$  où 'ChIP' correspond à l'IP et 'NS' (pour Non Spécifique) aux interactions aspécifiques estimées grâce au contrôle négatif. Enfin, l'enrichissement par rapport à l'Input est calculé à partir des valeurs :  $\text{Fold Enrichment} = 2^{(-\Delta \Delta Ct [\text{ChIP/NS}])}$ . Dans les calculs de %Input et d'enrichissement, la valeur 2 correspond à une efficacité des amorces de qPCR égale ou supérieure à 100%.

## 7. Liste des amorces

Nom du gène ou du transgène	Amorces 5' → 3'	Expérience(s)	Utilisée dans la(les) partie(s)...
<i>RpL32 / rp49</i>	for CCGCTTCAAGGGACAGTATCTG rev ATCTCGCCGCAGTAAACGC	RT-qPCR	I, II
	for GACCACGTTACAAGAACTCTCA rev GGTATCGACAACAGAGTGCG	RT-PCR	II
<i>eEF5</i>	for TAACATGGATGTGCCCAATG rev AACGCAATTGTTACCCAAT	RT-qPCR RIP-qPCR	I, II
<i>Otu</i>	Otu1 for TGGGTGTAGAGAACGACGTG rev ATACATGCCATTCGGTGGTT	RT-qPCR RIP-qPCR	II
	Otu2 for TCCCGCCTTGTGGTAGAATA rev AAGCTCGAATGGCAGTCAGT	RT-qPCR	III
<i>LacZ</i>	LacZ_Sens AATGCGCTCAGGTCAAATTC	RT	I
	LacZ_Antisens AGTACGAAATGCGTCGTTTAGAGC		
	LacZ1 for GAGAATCCGACGGGTTGTTA rev AAATTCAGACGGCAAACGAC  LacZ2 for ACTATCCCGACCGCCTTACT rev GTGGGCCATAATTCAATTCG	RT-qPCR RIP-qPCR ChIP-qPCR	I, II, III
<i>Ago1</i>	Ago1-RA-RC-RD for GGATCTCCAGATGACCTCCA rev GGACACTTGTCCGGCTGTAT	RT-qPCR	I
	Ago-RB for ATGAGCCGGTCATCTTTTTG rev GGCAATCGATGGTTTCTTGT		
<i>Tra2</i>	for TCCTTGTGTCCGGATGAAGT rev AAAAGCCAATGGATCGGGAG	RT-qPCR RIP-qPCR	II
<i>Hfp</i>	Hfp for GTTGGTAATGCAGCGACTGA rev ACTGACGGTGCCGAATTTAC	RT-qPCR	I, II
	Hfp-3'UTR for GCCAGTCCGAAAGCCAAC rev GACAATTGCTTCCCTCGCTG		

<i>JHDM2</i>	JHDM2_RA_Spliced_PCR for C G A C T G A A G T T T A C A C G G G C rev T G C A T C A G A A T A C G G A C G G A	RT-PCR	II
	JHDM2_RB_Spliced_PCR for C G C T T G C C C A A A C T T A C T C A rev C G A C G C C T T C A G T T C A A T G A		
	JHDM2_Unspliced_PCR for G A G A A T C C G A C G G G T T G T T A rev A A A T T C A G A C G G C A A A C G A C		
	JHDM2_CDS for A G A G A T T T G C G A C G T G T G T G rev A A T C C C T C C T T G C G A T C T T T	RT-qPCR RIP-qPCR	
	JHDM2_RA_Spliced for T C A A G C C A T T A C A C A A C A T G C A rev T G A G G T C G A G G A T A T G G T G C	RT-qPCR	II
	JHDM2_RB_Spliced for T G C A C A T G T T T A A A T T T C C C G A rev C A A C C A G A G C C G C C A A T T C		
	JHDM2_RA_Unspliced for C G A C T G A A G T T T A C A C G G G C rev C A G T G C A G C G C G T T G T T T T		
	JHDM2_RA+RB_Unspliced for A C C A C C T C A A A C C A C G A T C C rev A C C A G A G C C G C C A A T T C T T T		
<i>Moon</i>	for T C G A T C A T C A C A T C C T C C G G rev A G A A C C A G T C A T T T C A A C C A C A	RT-qPCR	II, III
<i>Rhino</i>	for T C T T G G T C T G G T C G A T G C A C rev A A A A C C G C T C C A C T T C A C C A	RT-qPCR	II
<i>Labial</i>	for T G A G G T T C A G G T T A G G A G G C rev G A G C G A C T T G G G G T A G A T G A	RT-qPCR	II
<i>opa</i>	for C A T G A A C C C G C T G A A C C A T T rev T C T C C C A C T C T C A A T A C G C C	RT-qPCR	II
<i>Oc</i>	for A T C A C C A C C G G C A T C C T T T A rev C C T C A A T T T G G G C G T C G G	RT-qPCR	II
<i>42AB</i>	42AB_Région 1 rev T G G A G T T T G G T G C A G A A G C for A G C C G T G C T T T A T G C T T T A C T	RT-qPCR ChIP-qPCR	I, III
	42AB_Région 2 for A A G A C C C A A T T T T T G C G T C G C rev C A A G G A T A G G G A T T T G G T C C		

<i>T3</i>	for CGTATGCGAGAGGAGTGTCA rev TCCAACGCCCTAAAGAGAGA	RT-qPCR ChIP-qPCR	III
<i>P(lArB)</i>	RyAd for GGGATCTGCGCCACATCG rev CGGGCTGCAGGAATTCGATA  Lash for CTTTGGCCTTAGTCGACGGA rev TCCTGGAGCCCGTCAGTATC	RT-qPCR ChIP-qPCR	III



# **BIBLIOGRAPHIE**



- Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, et al (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195
- Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, Sikkink K & Chandler VL (2006) An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**: 295–298
- Andersen PR, Tirian L, Vunjak M & Brennecke J (2017) A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549**: 54–59
- Andress A, Bei Y, Fonslow BR, Giri R, Wu Y, Yates JR & Carthew RW (2016) Spindle-E cycling between nuage and cytoplasm is controlled by Qin and PIWI proteins. *J. Cell Biol.* **213**: 201–211
- Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M & Tuschl T (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**: 203–207
- Arteaga-Vazquez M, Sidorenko L, Rabanal FA, Shrivistava R, Nobuta K, Green PJ, Meyers BC & Chandler VL (2010) RNA-mediated trans-communication can establish paramutation at the b1 locus in maize. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 12986–12991
- Asif-Laidin A, Delmarre V, Laurentie J, Miller WJ, Ronsseray S & Teyssset L (2017) Short and long-term evolutionary dynamics of subtelomeric piRNA clusters in *Drosophila*. *DNA Res.* **24**: 459–472
- Audibert A & Simonelig M (1999) The suppressor of forked gene of *Drosophila*, which encodes a homologue of human CstF-77K involved in mRNA 3'-end processing, is required for progression through mitosis. *Mech. Dev.* **82**: 41–50
- Barckmann B, Pierson S, Dufourt J, Papin C, Armenise C, Port F, Grentzinger T, Chambeyron S, Baronian G, Desvignes J-P, Curk T & Simonelig M (2015) Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo. *Cell Rep* **12**: 1205–1216
- Bastock R & St Johnston D (2008) *Drosophila* oogenesis. *Curr. Biol.* **18**: R1082-1087
- Batki J, Schnabl J, Wang J, Handler D, Andreev VI, Stieger CE, Novatchkova M, Lampersberger L, Kauneckaitė K, Xie W, Mechtler K, Patel DJ & Brennecke J (2019) The nascent RNA binding complex SFiNX licenses piRNA-guided heterochromatin formation. *Nat. Struct. Mol. Biol.* **26**: 720–731
- Baudry C, Malinsky S, Restituto M, Kapusta A, Rosa S, Meyer E & Bétermier M (2009) PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate *Paramecium tetraurelia*. *Genes Dev.* **23**: 2478–2483

- Belicard T, Jareosettasin P & Sarkies P (2018) The piRNA pathway responds to environmental signals to establish intergenerational adaptation to stress. *BMC Biol.* **16**: 103
- Belyayev A (2014) Bursts of transposable elements as an evolutionary driving force. *J. Evol. Biol.* **27**: 2573–2584
- Biémont C (2010) A brief history of the status of transposable elements: from junk DNA to major players in evolution. *Genetics* **186**: 1085–1093
- Biémont C & Vieira C (2006) Genetics: junk DNA as an evolutionary force. *Nature* **443**: 521–524
- Bingham PM, Kidwell MG & Rubin GM (1982) The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P-strain-specific transposon family. *Cell* **29**: 995–1004
- Boeke JD & Chapman KB (1991) Retrotransposition mechanisms. *Curr. Opin. Cell Biol.* **3**: 502–507
- Borchert GM, Lanier W & Davidson BL (2006) RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13**: 1097–1101
- Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew J-L, Ruan Y, Wei C-L, Ng HH & Liu ET (2008) Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res.* **18**: 1752–1762
- 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
- Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A & Hannon GJ (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**: 1387–1392
- Brink RA (1956) A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics* **41**: 872–889
- Britten R (2006) Transposable elements have contributed to thousands of human proteins. *Proc. Natl. Acad. Sci. U.S.A.* **103**: 1798–1803
- 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
- Camacho J, Truong L, Kurt Z, Chen Y-W, Morselli M, Gutierrez G, Pellegrini M, Yang X & Allard P (2018) The Memory of Environmental Chemical Exposure in *C. elegans* Is Dependent on the Jumonji Demethylases *jmjd-2* and *jmjd-3/utx-1*. *Cell Rep* **23**: 2392–2404
- Campbell CL, Black WC, Hess AM & Foy BD (2008) Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. *BMC Genomics* **9**: 425

- Casier K, Boivin A, Carré C & Teyssset L (2019a) Environmentally-Induced Transgenerational Epigenetic Inheritance: Implication of PIWI Interacting RNAs. *Cells* **8**:
- Casier K, Delmarre V, Gueguen N, Hermant C, Viodé E, Vaury C, Ronsseray S, Brasset E, Teyssset L & Boivin A (2019b) Environmentally-induced epigenetic conversion of a piRNA cluster. *Elife* **8**:
- Cen H, Zheng S, Fang Y-M, Tang X-P & Dong Q (2004) Induction of HSF1 expression is associated with sporadic colorectal cancer. *World J. Gastroenterol.* **10**: 3122–3126
- Chang TH, Mattei E, Gainetdinov I, Colpan C, Weng Z & Zamore PD (2019) Maelstrom Represses Canonical Polymerase II Transcription within Bi-directional piRNA Clusters in *Drosophila melanogaster*. *Mol. Cell* **73**: 291-303.e6
- Charlesworth B & Charlesworth D (1983) The population dynamics of transposable elements. *Genetics Research* **42**: 1–27
- Chen Y-CA, Stuwe E, Luo Y, Ninova M, Le Thomas A, Rozhavskaia E, Li S, Vempati S, Laver JD, Patel DJ, Smibert CA, Lipshitz HD, Toth KF & Aravin AA (2016) Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Mol. Cell* **63**: 97–109
- Cheng C-Y, Vogt A, Mochizuki K & Yao M-C (2010) A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in *Tetrahymena thermophila*. *Mol. Biol. Cell* **21**: 1753–1762
- Chuong EB, Elde NC & Feschotte C (2016) Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science* **351**: 1083–1087
- Clark JP, Rahman R, Yang N, Yang LH & Lau NC (2017) *Drosophila* PAF1 Modulates PIWI/piRNA Silencing Capacity. *Curr. Biol.* **27**: 2718-2726.e4
- Coe EH (1966) The properties, origin, and mechanism of conversion-type inheritance at the B locus in maize. *Genetics* **53**: 1035–1063
- Cordaux R & Batzer MA (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **10**: 691–703
- 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
- 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
- Czech B, Munafò M, Ciabrelli F, Eastwood EL, Fabry MH, Kneuss E & Hannon GJ (2018) piRNA-Guided Genome Defense: From Biogenesis to Silencing. *Annu. Rev. Genet.* **52**: 131–157

- 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. *Mol. Cell* **50**: 749–761
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG & Chovnick A (1990) Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* **124**: 339–355
- Dennis C, Brasslet E, Sarkar A & Vaury C (2016) Export of piRNA precursors by EJC triggers assembly of cytoplasmic Yb-body in *Drosophila*. *Nat Commun* **7**: 13739
- Dennis C, Zanni V, Brasslet E, Eymery A, Zhang L, Mteirek R, Jensen S, Rong YS & Vaury C (2013) ‘Dot COM’, a nuclear transit center for the primary piRNA pathway in *Drosophila*. *PLoS ONE* **8**: e72752
- 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
- Díaz-González J, Vázquez JF, Albornoz J & Domínguez A (2011) Long-term evolution of the roo transposable element copy number in mutation accumulation lines of *Drosophila melanogaster*. *Genet Res (Camb)* **93**: 181–187
- 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
- Doolittle WF & Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603
- Dorer DR & Henikoff S (1994) Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**: 993–1002
- Dorer DR & Henikoff S (1997) Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* **147**: 1181–1190
- Dubin MJ, Mittelsten Scheid O & Becker C (2018) Transposons: a blessing curse. *Curr. Opin. Plant Biol.* **42**: 23–29
- Duc C, Yoth M, Jensen S, Mounié N, Bergman CM, Vaury C & Brasslet E (2019) Trapping a somatic endogenous retrovirus into a germline piRNA cluster immunizes the germline against further invasion. *Genome Biol.* **20**: 127
- ElMaghraby MF, Andersen PR, Pühringer F, Hohmann U, Meixner K, Lendl T, Tirian L & Brennecke J (2019) A Heterochromatin-Specific RNA Export Pathway Facilitates piRNA Production. *Cell* **178**: 964-979.e20
- Engels WR & Preston CR (1979) Hybrid dysgenesis in *Drosophila melanogaster*: the biology of female and male sterility. *Genetics* **92**: 161–174
- Esnault C, Lee M, Ham C & Levin HL (2019) Transposable element insertions in fission yeast drive adaptation to environmental stress. *Genome Res.* **29**: 85–95

- Fabry MH, Ciabrelli F, Munafò M, Eastwood EL, Kneuss E, Falciatori I, Falconio FA, Hannon GJ & Czech B (2019) piRNA-guided co-transcriptional silencing coopts nuclear export factors. *Elife* **8**:
- Fast I, Hewel C, Wester L, Schumacher J, Gebert D, Zischler H, Berger C & Rosenkranz D (2017) Temperature-responsive miRNAs in *Drosophila* orchestrate adaptation to different ambient temperatures. *RNA* **23**: 1352–1364
- Fedoroff NV (2012) Presidential address. Transposable elements, epigenetics, and genome evolution. *Science* **338**: 758–767
- Feschotte C & Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* **41**: 331–368
- Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet.* **5**: 103–107
- Finnegan DJ (1992) Transposable elements. *Curr. Opin. Genet. Dev.* **2**: 861–867
- Gagliardi M & Matarazzo MR (2016) RIP: RNA Immunoprecipitation. *Methods Mol. Biol.* **1480**: 73–86
- Gainetdinov I, Colpan C, Arif A, Cecchini K & Zamore PD (2018) A Single Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA Production in Most Animals. *Mol. Cell* **71**: 775-790.e5
- Ganguly A, Dunbar T, Chen P, Godmilow L & Ganguly T (2003) Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum. Genet.* **113**: 348–352
- García Guerreiro MP (2012) What makes transposable elements move in the *Drosophila* genome? *Heredity (Edinb)* **108**: 461–468
- Gdula DA, Gerasimova TI & Corces VG (1996) Genetic and molecular analysis of the gypsy chromatin insulator of *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 9378–9383
- George P, Jensen S, Pogorelcnik R, Lee J, Xing Y, Brasslet E, Vaury C & Sharakhov IV (2015) Increased production of piRNAs from euchromatic clusters and genes in *Anopheles gambiae* compared with *Drosophila melanogaster*. *Epigenetics Chromatin* **8**: 50
- Geyer PK & Corces VG (1992) DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* **6**: 1865–1873
- 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
- Gilbert C & Cordaux R (2017) Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Curr Opin Virol* **25**: 16–22

- Girard A, Sachidanandam R, Hannon GJ & Carmell MA (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**: 199–202
- Goriaux C, Desset S, Renaud Y, Vaury C & Brassat E (2014) Transcriptional properties and splicing of the flamenco piRNA cluster. *EMBO Rep.* **15**: 411–418
- Gou L-T, Dai P, Yang J-H, Xue Y, Hu Y-P, Zhou Y, Kang J-Y, Wang X, Li H, Hua M-M, Zhao S, Hu S-D, Wu L-G, Shi H-J, Li Y, Fu X-D, Qu L-H, Wang E-D & Liu M-F (2014) Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* **24**: 680–700
- Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degan BM, Rokhsar DS & Bartel DP (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**: 1193–1197
- Guffanti G, Gaudi S, Fallon JH, Sobell J, Potkin SG, Pato C & Macciardi F (2014) Transposable elements and psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **165B**: 201–216
- Guio L, Vieira C & González J (2018) Stress affects the epigenetic marks added by natural transposable element insertions in *Drosophila melanogaster*. *Sci Rep* **8**: 12197
- 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
- Guo C, Jeong H-H, Hsieh Y-C, Klein H-U, Bennett DA, De Jager PL, Liu Z & Shulman JM (2018) Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep* **23**: 2874–2880
- Ha H, Song J, Wang S, Kapusta A, Feschotte C, Chen KC & Xing J (2014) A comprehensive analysis of piRNAs from adult human testis and their relationship with genes and mobile elements. *BMC Genomics* **15**: 545
- Hancock RL, Masson N, Dunne K, Flashman E & Kawamura A (2017) The Activity of JmjC Histone Lysine Demethylase KDM4A is Highly Sensitive to Oxygen Concentrations. *ACS Chem. Biol.* **12**: 1011–1019
- Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS & Brennecke J (2013) The genetic makeup of the *Drosophila* piRNA pathway. *Mol. Cell* **50**: 762–777
- Handler D, Olivieri D, Novatchkova M, Gruber FS, Meixner K, Mechtler K, Stark A, Sachidanandam R & Brennecke J (2011) A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *EMBO J.* **30**: 3977–3993
- Hayashi R, Schnabl J, Handler D, Mohn F, Ameres SL & Brennecke J (2016) Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature* **539**: 588–592
- He L, Wang X & Montell DJ (2011) Shining light on *Drosophila* oogenesis: live imaging of egg development. *Curr. Opin. Genet. Dev.* **21**: 612–619

- Heard E & Martienssen RA (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* **157**: 95–109
- Hermant C, Boivin A, Teyssset L, Delmarre V, Asif-Laidin A, van den Beek M, Antoniewski C & Ronsseray S (2015) Paramutation in *Drosophila* Requires Both Nuclear and Cytoplasmic Actors of the piRNA Pathway and Induces Cis-spreading of piRNA Production. *Genetics* **201**: 1381–1396
- Herz H-M, Morgan M, Gao X, Jackson J, Rickels R, Swanson SK, Florens L, Washburn MP, Eissenberg JC & Shilatifard A (2014) Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* **345**: 1065–1070
- Hickok JR, Vasudevan D, Antholine WE & Thomas DD (2013) Nitric oxide modifies global histone methylation by inhibiting Jumonji C domain-containing demethylases. *J. Biol. Chem.* **288**: 16004–16015
- Higuet D, Anxolabéhère D & Nouaud D (1992) A particular P-element insertion is correlated to the P-induced hybrid dysgenesis repression in *Drosophila melanogaster*. *Genet. Res.* **60**: 15–24
- Hirakata S & Siomi MC (2016) piRNA biogenesis in the germline: From transcription of piRNA genomic sources to piRNA maturation. *Biochim. Biophys. Acta* **1859**: 82–92
- Hirano T, Iwasaki YW, Lin ZY-C, Imamura M, Seki NM, Sasaki E, Saito K, Okano H, Siomi MC & Siomi H (2014) Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate. *RNA* **20**: 1223–1237
- Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P & Zamore PD (2007) The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.* **17**: 1265–1272
- Hoskins RA, Carlson JW, Kennedy C, Acevedo D, Evans-Holm M, Frise E, Wan KH, Park S, Mendez-Lago M, Rossi F, Villasante A, Dimitri P, Karpen GH & Celniker SE (2007) Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science* **316**: 1625–1628
- Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, Kaminker JS, Kennedy C, Mungall CJ, Sullivan BA, Sutton GG, Yasuhara JC, Wakimoto BT, Myers EW, Celniker SE, Rubin GM & Karpen GH (2002) Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol.* **3**: RESEARCH0085
- Hur JK, Luo Y, Moon S, Ninova M, Marinov GK, Chung YD & Aravin AA (2016) Splicing-independent loading of TREX on nascent RNA is required for efficient expression of dual-strand piRNA clusters in *Drosophila*. *Genes Dev.* **30**: 840–855
- 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
- Ishizu H, Kinoshita T, Hirakata S, Komatsuzaki C & Siomi MC (2019) Distinct and Collaborative Functions of Yb and Armitage in Transposon-Targeting piRNA Biogenesis. *Cell Rep* **27**: 1822-1835.e8

- Iwasaki YW, Siomi MC & Siomi H (2015) PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu. Rev. Biochem.* **84**: 405–433
- Jang HS, Shah NM, Du AY, Dailey ZZ, Pehrsson EC, Godoy PM, Zhang D, Li D, Xing X, Kim S, O'Donnell D, Gordon JI & Wang T (2019) Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.* **51**: 611–617
- Jehn J, Gebert D, Pipilescu F, Stern S, Kiefer JST, Hewel C & Rosenkranz D (2018) PIWI genes and piRNAs are ubiquitously expressed in mollusks and show patterns of lineage-specific adaptation. *Commun Biol* **1**: 137
- Josse T, Teyssset L, Todeschini A-L, Sidor CM, Anxolabéhère D & Ronsseray S (2007) Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet.* **3**: 1633–1643
- Karpen GH & Spradling AC (1992) Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome Dp1187 by single P element insertional mutagenesis. *Genetics* **132**: 737–753
- Keam SP, Young PE, McCorkindale AL, Dang THY, Clancy JL, Humphreys DT, Preiss T, Hutvagner G, Martin DIK, Copley JE & Suter CM (2014) The human Piwi protein Hiwi2 associates with tRNA-derived piRNAs in somatic cells. *Nucleic Acids Res.* **42**: 8984–8995
- Kelleher ES, Azevedo RBR & Zheng Y (2018) The Evolution of Small-RNA-Mediated Silencing of an Invading Transposable Element. *Genome Biol Evol* **10**: 3038–3057
- 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
- Kidwell MG, Kidwell JF & Sved JA (1977) Hybrid Dysgenesis in *DROSOPHILA MELANOGASTER*: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination. *Genetics* **86**: 813–833
- Kitano S, Kurasawa H & Aizawa Y (2018) Transposable elements shape the human proteome landscape via formation of cis-acting upstream open reading frames. *Genes Cells* **23**: 274–284
- Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S, Sugano S, Shimada T, Suzuki Y, Suzuki MG & Katsuma S (2014) A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* **509**: 633–636
- Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, Weng Z & Theurkauf WE (2009) The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **138**: 1137–1149
- Klenov MS, Lavrov SA, Korbut AP, Stolyarenko AD, Yakushev EY, Reuter M, Pillai RS & Gvozdev VA (2014) Impact of nuclear Piwi elimination on chromatin state in *Drosophila melanogaster* ovaries. *Nucleic Acids Res.* **42**: 6208–6218



- Klenov MS, Sokolova OA, Yakushev EY, Stolyarenko AD, Mikhaleva EA, Lavrov SA & Gvozdev VA (2011) Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proc. Natl. Acad. Sci. U.S.A.* **108**: 18760–18765
- Knebelmann B, Forestier L, Drouot L, Quinones S, Chuet C, Benessy F, Saus J & Antignac C (1995) Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome. *Hum. Mol. Genet.* **4**: 675–679
- Kneuss E, Munafò M, Eastwood EL, Deumer U-S, Preall JB, Hannon GJ & Czech B (2019) Specialization of the Drosophila nuclear export family protein Nxf3 for piRNA precursor export. *Genes Dev.* **33**: 1208–1220
- Koch CM, Honemann-Capito M, Egger-Adam D & Wodarz A (2009) Windei, the Drosophila homolog of mAM/MCAF1, is an essential cofactor of the H3K9 methyl transferase dSETDB1/Eggless in germ line development. *PLoS Genet.* **5**: e1000644
- Kofler R (2019) Dynamics of Transposable Element Invasions with piRNA Clusters. *Mol. Biol. Evol.* **36**: 1457–1472
- Kofler R (2020) piRNA Clusters Need a Minimum Size to Control Transposable Element Invasions. *Genome Biol Evol* **12**: 736–749
- Kofler R, Hill T, Nolte V, Betancourt AJ & Schlötterer C (2015a) The recent invasion of natural Drosophila simulans populations by the P-element. *Proc. Natl. Acad. Sci. U.S.A.* **112**: 6659–6663
- Kofler R, Nolte V & Schlötterer C (2015b) Tempo and Mode of Transposable Element Activity in Drosophila. *PLoS Genet.* **11**: e1005406
- Kofler R, Senti K-A, Nolte V, Tobler R & Schlötterer C (2018) Molecular dissection of a natural transposable element invasion. *Genome Res.* **28**: 824–835
- Komarov PA, Sokolova O, Akulenko N, Brasslet E, Jensen S & Kalmykova A (2020) Epigenetic Requirements for Triggering Heterochromatinization and Piwi-Interacting RNA Production from Transgenes in the Drosophila Germline. *Cells* **9**:
- Kondo-Iida E, Kobayashi K, Watanabe M, Sasaki J, Kumagai T, Koide H, Saito K, Osawa M, Nakamura Y & Toda T (1999) Novel mutations and genotype-phenotype relationships in 107 families with Fukuyama-type congenital muscular dystrophy (FCMD). *Hum. Mol. Genet.* **8**: 2303–2309
- Kotov AA, Adashev VE, Godneeva BK, Ninova M, Shatskikh AS, Bazylev SS, Aravin AA & Olenina LV (2019) piRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in Drosophila melanogaster. *Nucleic Acids Res.* **47**: 4255–4271
- Laski FA, Rio DC & Rubin GM (1986) Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. *Cell* **44**: 7–19
- Law PT-Y, Qin H, Ching AK-K, Lai KP, Co NN, He M, Lung RW-M, Chan AW-H, Chan T-F & Wong N (2013) Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. *J. Hepatol.* **58**: 1165–1173

- 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
- Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen Y-CA, Luo Y, Sachidanandam R, Toth KF, Patel D & Aravin AA (2014) Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev.* **28**: 1667–1680
- Lemaitre B, Ronsseray S & Coen D (1993) Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: a model for the P cytotype. *Genetics* **135**: 149–160
- Levin HL & Moran JV (2011) Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12**: 615–627
- 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
- Li XZ, Roy CK, Dong X, Bolcun-Filas E, Wang J, Han BW, Xu J, Moore MJ, Schimenti JC, Weng Z & Zamore PD (2013) An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol. Cell* **50**: 67–81
- Lim AK & Kai T (2007) Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **104**: 6714–6719
- Lim JP & Brunet A (2013) Bridging the transgenerational gap with epigenetic memory. *Trends Genet.* **29**: 176–186
- Liu G, NISC Comparative Sequencing Program, Zhao S, Bailey JA, Sahinalp SC, Alkan C, Tuzun E, Green ED & Eichler EE (2003) Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome. *Genome Res.* **13**: 358–368
- Loreto ELS, Carareto CMA & Capy P (2008) Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity (Edinb)* **100**: 545–554
- Lu Y, Zhang K, Li C, Yao Y, Tao D, Liu Y, Zhang S & Ma Y (2012) Piwil2 suppresses p53 by inducing phosphorylation of signal transducer and activator of transcription 3 in tumor cells. *PLoS ONE* **7**: e30999
- Luo S, Zhang H, Duan Y, Yao X, Clark AG & Lu J (2020) The evolutionary arms race between transposable elements and piRNAs in *Drosophila melanogaster*. *BMC Evol. Biol.* **20**: 14
- Lutz B, Lu HC, Eichele G, Miller D & Kaufman TC (1996) Rescue of *Drosophila* labial null mutant by the chicken ortholog Hoxb-1 demonstrates that the function of Hox genes is phylogenetically conserved. *Genes Dev.* **10**: 176–184

- Marie PP, Ronsseray S & Boivin A (2017) From Embryo to Adult: piRNA-Mediated Silencing throughout Germline Development in *Drosophila*. *G3 (Bethesda)* **7**: 505–516
- Marin L, Lehmann M, Nouaud D, Izaabel H, Anxolabéhère D & Ronsseray S (2000) P-Element repression in *Drosophila melanogaster* by a naturally occurring defective telomeric P copy. *Genetics* **155**: 1841–1854
- Martinez VD, Vucic EA, Thu KL, Hubaux R, Enfield KSS, Pikor LA, Becker-Santos DD, Brown CJ, Lam S & Lam WL (2015) Unique somatic and malignant expression patterns implicate PIWI-interacting RNAs in cancer-type specific biology. *Sci Rep* **5**: 10423
- Mason JM, Frydrychova RC & Biessmann H (2008) *Drosophila* telomeres: an exception providing new insights. *Bioessays* **30**: 25–37
- McCLINTOCK B (1950) The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U.S.A.* **36**: 344–355
- McDonald JF (1993) Evolution and consequences of transposable elements. *Curr. Opin. Genet. Dev.* **3**: 855–864
- McDonald JF (1995) Transposable elements: possible catalysts of organismic evolution. *Trends Ecol. Evol. (Amst.)* **10**: 123–126
- Mei Y, Clark D & Mao L (2013) Novel dimensions of piRNAs in cancer. *Cancer Lett.* **336**: 46–52
- 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
- Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, Vogelstein B & Nakamura Y (1992) Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* **52**: 643–645
- Miousse IR, Chalbot M-CG, Lumen A, Ferguson A, Kavouras IG & Koturbash I (2015) Response of transposable elements to environmental stressors. *Mutat Res Rev Mutat Res* **765**: 19–39
- Mohn F, Sienski G, Handler D & Brennecke J (2014) The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* **157**: 1364–1379
- Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD & Kazazian HH (1996) High frequency retrotransposition in cultured mammalian cells. *Cell* **87**: 917–927
- 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**: e1002470

- Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R & Hannon GJ (2013) A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol. Cell* **50**: 736–748
- Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ & Aravin AA (2012) Production of artificial piRNAs in flies and mice. *RNA* **18**: 42–52
- Mugat B, Nicot S, Varela-Chavez C, Jourdan C, Sato K, Basyuk E, Juge F, Siomi MC, Péliesson A & Chambeyron S (2020) The Mi-2 nucleosome remodeler and the Rpd3 histone deacetylase are involved in piRNA-guided heterochromatin formation. *Nat Commun* **11**: 2818
- Mukherjee S, Mukhopadhyay A, Banerjee D, Chandak GR & Ray K (2004) Molecular pathology of haemophilia B: identification of five novel mutations including a LINE 1 insertion in Indian patients. *Haemophilia* **10**: 259–263
- Murano K, Iwasaki YW, Ishizu H, Mashiko A, Shibuya A, Kondo S, Adachi S, Suzuki S, Saito K, Natsume T, Siomi MC & Siomi H (2019) Nuclear RNA export factor variant initiates piRNA-guided co-transcriptional silencing. *EMBO J.* **38**: e102870
- Murota Y, Ishizu H, Nakagawa S, Iwasaki YW, Shibata S, Kamatani MK, Saito K, Okano H, Siomi H & Siomi MC (2014) Yb integrates piRNA intermediates and processing factors into perinuclear bodies to enhance piRISC assembly. *Cell Rep* **8**: 103–113
- Narita N, Nishio H, Kitoh Y, Ishikawa Y, Ishikawa Y, Minami R, Nakamura H & Matsuo M (1993) Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J. Clin. Invest.* **91**: 1862–1867
- Ng KW, Anderson C, Marshall EA, Minatel BC, Enfield KSS, Saprunoff HL, Lam WL & Martinez VD (2016) Piwi-interacting RNAs in cancer: emerging functions and clinical utility. *Mol. Cancer* **15**: 5
- Ninova M, Chen Y-CA, Godneeva B, Rogers AK, Luo Y, Fejes Tóth K & Aravin AA (2020) Su(var)2-10 and the SUMO Pathway Link piRNA-Guided Target Recognition to Chromatin Silencing. *Mol. Cell* **77**: 556-570.e6
- Nishida KM, Iwasaki YW, Murota Y, Nagao A, Mannen T, Kato Y, Siomi H & Siomi MC (2015) Respective functions of two distinct Siwi complexes assembled during PIWI-interacting RNA biogenesis in *Bombyx* germ cells. *Cell Rep* **10**: 193–203
- Nishida KM, Okada TN, Kawamura T, Mituyama T, Kawamura Y, Inagaki S, Huang H, Chen D, Kodama T, Siomi H & Siomi MC (2009) Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J.* **28**: 3820–3831
- 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

- Niu Y, DesMarais TL, Tong Z, Yao Y & Costa M (2015) Oxidative stress alters global histone modification and DNA methylation. *Free Radic. Biol. Med.* **82**: 22–28
- O’Hare K & Rubin GM (1983) Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* **34**: 25–35
- Ohno S (1972) So much ‘junk’ DNA in our genome. *Brookhaven Symp. Biol.* **23**: 366–370
- 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
- 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
- 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
- Orgel LE & Crick FH (1980) Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607
- 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
- Pane A, Wehr K & Schüpbach T (2007) zucchini and squash encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. *Dev. Cell* **12**: 851–862
- Parhad SS, Tu S, Weng Z & Theurkauf WE (2017) Adaptive Evolution Leads to Cross-Species Incompatibility in the piRNA Transposon Silencing Machinery. *Dev. Cell* **43**: 60-70.e5
- Patil VS, Anand A, Chakrabarti A & Kai T (2014) The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of *Drosophila melanogaster*. *BMC Biol.* **12**: 61
- Patil VS & Kai T (2010) Repression of retroelements in *Drosophila* germline via piRNA pathway by the Tudor domain protein Tejas. *Curr. Biol.* **20**: 724–730
- Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D & Knowles BB (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell* **7**: 597–606
- Péligsson 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
- 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**: E4218-4227

- Piacentini L, Fanti L, Specchia V, Bozzetti MP, Berloco M, Palumbo G & Pimpinelli S (2014) Transposons, environmental changes, and heritable induced phenotypic variability. *Chromosoma* **123**: 345–354
- Praher D, Zimmermann B, Genikhovich G, Columbus-Shenkar Y, Modepalli V, Aharoni R, Moran Y & Technau U (2017) Characterization of the piRNA pathway during development of the sea anemone *Nematostella vectensis*. *RNA Biol* **14**: 1727–1741
- Preall JB, Czech B, Guzzardo PM, Muerdter F & Hannon GJ (2012) shutdown is a component of the *Drosophila* piRNA biogenesis machinery. *RNA* **18**: 1446–1457
- Pritham EJ (2009) Transposable elements and factors influencing their success in eukaryotes. *J. Hered.* **100**: 648–655
- Radion E, Morgunova V, Ryazansky S, Akulenko N, Lavrov S, Abramov Y, Komarov PA, Glukhov SI, Olovnikov I & Kalmykova A (2018) Key role of piRNAs in telomeric chromatin maintenance and telomere nuclear positioning in *Drosophila* germline. *Epigenetics Chromatin* **11**: 40
- Radion E, Sokolova O, Ryazansky S, Komarov PA, Abramov Y & Kalmykova A (2019) The Integrity of piRNA Clusters is Abolished by Insulators in the *Drosophila* Germline. *Genes (Basel)* **10**:
- Rajasethupathy P, Antonov I, Sheridan R, Frey S, Sander C, Tuschl T & Kandel ER (2012) A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell* **149**: 693–707
- Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ & Lehmann R (2011) piRNA production requires heterochromatin formation in *Drosophila*. *Curr. Biol.* **21**: 1373–1379
- Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I & Cuzin F (2006) RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441**: 469–474
- Rebollo R, Farivar S & Mager DL (2012a) C-GATE - catalogue of genes affected by transposable elements. *Mob DNA* **3**: 9
- Rebollo R, Horard B, Hubert B & Vieira C (2010) Jumping genes and epigenetics: Towards new species. *Gene* **454**: 1–7
- Rebollo R, Romanish MT & Mager DL (2012b) Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* **46**: 21–42
- 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
- Roche SE & Rio DC (1998) Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the *Drosophila* Polycomb group gene, Enhancer of zeste. *Genetics* **149**: 1839–1855

- Ronsseray S, Anxolabéhère D & Périquet G (1984) Hybrid dysgenesis in *Drosophila melanogaster*: influence of temperature on cytotype determination in the P-M system. *Mol. Gen. Genet.* **196**: 17–23
- Ronsseray S, Boivin A & Anxolabéhère D (2001) P-Element repression in *Drosophila melanogaster* by variegating clusters of P-lacZ-white transgenes. *Genetics* **159**: 1631–1642
- Ronsseray S, Lehmann M & Anxolabéhère D (1989) Copy number and distribution of P and I mobile elements in *Drosophila melanogaster* populations. *Chromosoma* **98**: 207–214
- Ronsseray S, Lehmann M, Nouaud D & Anxolabéhère D (1996) The regulatory properties of autonomous subtelomeric P elements are sensitive to a Suppressor of variegation in *Drosophila melanogaster*. *Genetics* **143**: 1663–1674
- Rouget C, Papin C, Boureux A, Meunier A-C, Franco B, Robine N, Lai EC, Pelisson A & Simonelig M (2010) Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* **467**: 1128–1132
- Roy J, Sarkar A, Parida S, Ghosh Z & Mallick B (2017) Small RNA sequencing revealed dysregulated piRNAs in Alzheimer’s disease and their probable role in pathogenesis. *Mol Biosyst* **13**: 565–576
- 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
- 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
- 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
- Saito K, Sakaguchi Y, Suzuki T, Suzuki T, 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
- Sato K, Iwasaki YW, Shibuya A, Carninci P, Tsuchizawa Y, Ishizu H, Siomi MC & Siomi H (2015) Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the *Drosophila* Germline. *Mol. Cell* **59**: 553–563
- Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, Kutter C, Brown GD, Marshall A, Flicek P & Odom DT (2012) Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* **148**: 335–348
- Schnettler E, Donald CL, Human S, Watson M, Siu RWC, McFarlane M, Fazakerley JK, Kohl A & Fragkoudis R (2013) Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *J. Gen. Virol.* **94**: 1680–1689

- Seong K-H, Li D, Shimizu H, Nakamura R & Ishii S (2011) Inheritance of stress-induced, ATF-2-dependent epigenetic change. *Cell* **145**: 1049–1061
- Serrato-Capuchina A & Matute DR (2018) The Role of Transposable Elements in Speciation. *Genes (Basel)* **9**:
- Shalaby NA, Sayed R, Zhang Q, Scoggin S, Eliazer S, Rothenfluh A & Buszczak M (2017) Systematic discovery of genetic modulation by Jumonji histone demethylases in *Drosophila*. *Sci Rep* **7**: 5240
- Shirayama M, Seth M, Lee H-C, Gu W, Ishidate T, Conte D & Mello CC (2012) piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **150**: 65–77
- Shmakova A, Batie M, Druker J & Rocha S (2014) Chromatin and oxygen sensing in the context of JmjC histone demethylases. *Biochem. J.* **462**: 385–395
- 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
- Sienski G, Batki J, Senti K-A, Dönertas D, Tirian L, Meixner K & Brennecke J (2015) Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes Dev.* **29**: 2258–2271
- 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
- Skinner MK (2008) What is an epigenetic transgenerational phenotype? F3 or F2. *Reprod. Toxicol.* **25**: 2–6
- Soshnev AA, Baxley RM, Manak JR, Tan K & Geyer PK (2013) The insulator protein Suppressor of Hairy-wing is an essential transcriptional repressor in the *Drosophila* ovary. *Development* **140**: 3613–3623
- Sousa A, Bourgard C, Wahl LM & Gordo I (2013) Rates of transposition in *Escherichia coli*. *Biol. Lett.* **9**: 20130838
- Stam M, Belele C, Dorweiler JE & Chandler VL (2002) Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev.* **16**: 1906–1918
- Su-Keene EJ, Bonilla MM, Padua MV, Zeh DW & Zeh JA (2018) Simulated climate warming and mitochondrial haplogroup modulate testicular small non-coding RNA expression in the neotropical pseudoscorpion, *Cordylochernes scorpioides*. *Environ Epigenet* **4**: dvy027
- Symington LS & Gautier J (2011) Double-strand break end resection and repair pathway choice. *Annu. Rev. Genet.* **45**: 247–271



- Tighe PJ, Stevens SE, Dempsey S, Le Deist F, Rieux-Laucat F & Edgar JDM (2002) Inactivation of the Fas gene by Alu insertion: retrotransposition in an intron causing splicing variation and autoimmune lymphoproliferative syndrome. *Genes Immun.* **3 Suppl 1**: S66-70
- Todeschini A-L, Teyssset 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
- Tóth KF, Pezic D, Stuwe E & Webster A (2016) The piRNA Pathway Guards the Germline Genome Against Transposable Elements. *Adv. Exp. Med. Biol.* **886**: 51–77
- 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
- Vagin VV, Wohlschlegel J, Qu J, Jonsson Z, Huang X, Chuma S, Girard A, Sachidanandam R, Hannon GJ & Aravin AA (2009) Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* **23**: 1749–1762
- Vagin VV, Yu Y, Jankowska A, Luo Y, Wasik KA, Malone CD, Harrison E, Rosebrock A, Wakimoto BT, Fagegaltier D, Muerdter F & Hannon GJ (2013) Minotaur is critical for primary piRNA biogenesis. *RNA* **19**: 1064–1077
- Van Buskirk C & Schüpbach T (2002) Half pint regulates alternative splice site selection in *Drosophila*. *Dev. Cell* **2**: 343–353
- de Vanssay A, Bougé A-L, Boivin A, Hermant C, Teyssset L, Delmarre V, Antoniewski C & Ronsseray S (2012) Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* **490**: 112–115
- Wang H, Ma Z, Niu K, Xiao Y, Wu X, Pan C, Zhao Y, Wang K, Zhang Y & Liu N (2016) Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila*. *Development* **143**: 530–539
- Wang J, Xie G, Singh M, Ghanbarian AT, Raskó T, Szvetnik A, Cai H, Besser D, Prigione A, Fuchs NV, Schumann GG, Chen W, Lorincz MC, Ivics Z, Hurst LD & Izsvák Z (2014a) Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**: 405–409
- Wang S, Wagner EJ & Mattox W (2013) Half pint/Puf68 is required for negative regulation of splicing by the SR splicing factor Transformer2. *RNA Biol* **10**: 1396–1406
- 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
- Wang W, Yoshikawa M, Han BW, Izumi N, Tomari Y, Weng Z & Zamore PD (2014b) The initial uridine of primary piRNAs does not create the tenth adenine that is the hallmark of secondary piRNAs. *Mol. Cell* **56**: 708–716

- Webster A, Li S, Hur JK, Wachsmuth M, Bois JS, Perkins EM, Patel DJ & Aravin AA (2015) Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Mol. Cell* **59**: 564–575
- Wieschaus E & Nüsslein-Volhard C (2016) The Heidelberg Screen for Pattern Mutants of *Drosophila*: A Personal Account. *Annu. Rev. Cell Dev. Biol.* **32**: 1–46
- Wilson AC, Sarich VM & Maxson LR (1974) The importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein, and anatomical evolution. *Proc. Natl. Acad. Sci. U.S.A.* **71**: 3028–3030
- 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. *Proc. Natl. Acad. Sci. U.S.A.* **107**: 1606–1611
- Wylie A, Jones AE & Abrams JM (2016a) p53 in the game of transposons. *Bioessays* **38**: 1111–1116
- Wylie A, Jones AE, D’Brot A, Lu W-J, Kurtz P, Moran JV, Rakheja D, Chen KS, Hammer RE, Comerford SA, Amatruda JF & Abrams JM (2016b) p53 genes function to restrain mobile elements. *Genes Dev.* **30**: 64–77
- Xiol J, Spinelli P, Laussmann MA, Homolka D, Yang Z, Cora E, Couté Y, Conn S, Kadlec J, Sachidanandam R, Kaksonen M, Cusack S, Ephrussi A & Pillai RS (2014) RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. *Cell* **157**: 1698–1711
- Yamane K, Toumazou C, Tsukada Y, Erdjument-Bromage H, Tempst P, Wong J & Zhang Y (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **125**: 483–495
- Yamashiro H & Siomi MC (2018) PIWI-Interacting RNA in *Drosophila*: Biogenesis, Transposon Regulation, and Beyond. *Chem. Rev.* **118**: 4404–4421
- Yao Y, Li C, Zhou X, Zhang Y, Lu Y, Chen J, Zheng X, Tao D, Liu Y & Ma Y (2014) PIWIL2 induces c-Myc expression by interacting with NME2 and regulates c-Myc-mediated tumor cell proliferation. *Oncotarget* **5**: 8466–8477
- Yin J, Jiang X-Y, Qi W, Ji C-G, Xie X-L, Zhang D-X, Cui Z-J, Wang C-K, Bai Y, Wang J & Jiang H-Q (2017) piR-823 contributes to colorectal tumorigenesis by enhancing the transcriptional activity of HSF1. *Cancer Sci.* **108**: 1746–1756
- Yohn NL, Bartolomei MS & Blendy JA (2015) Multigenerational and transgenerational inheritance of drug exposure: The effects of alcohol, opiates, cocaine, marijuana, and nicotine. *Prog. Biophys. Mol. Biol.* **118**: 21–33
- Yoon J, Lee K-S, Park JS, Yu K, Paik S-G & Kang Y-K (2008) dSETDB1 and SU(VAR)3-9 sequentially function during germline-stem cell differentiation in *Drosophila melanogaster*. *PLoS ONE* **3**: e2234
- 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

- Zambon RA, Vakharia VN & Wu LP (2006) RNAi is an antiviral immune response against a dsRNA virus in *Drosophila melanogaster*. *Cell. Microbiol.* **8**: 880–889
- Zamparini AL, Davis MY, Malone CD, Vieira E, Zavadil J, Sachidanandam R, Hannon GJ & Lehmann R (2011) Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in *Drosophila*. *Development* **138**: 4039–4050
- 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
- Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, Schultz N, Vreven T, Meignin C, Davis I, Zamore PD, Weng Z & Theurkauf WE (2012) UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **151**: 871–884
- Zhang G, Tu S, Yu T, Zhang X-O, Parhad SS, Weng Z & Theurkauf WE (2018) Co-dependent Assembly of *Drosophila* piRNA Precursor Complexes and piRNA Cluster Heterochromatin. *Cell Rep* **24**: 3413-3422.e4
- Zhang Y, Cheng TC, Huang G, Lu Q, Surleac MD, Mandell JD, Pontarotti P, Petrescu AJ, Xu A, Xiong Y & Schatz DG (2019) Transposon molecular domestication and the evolution of the RAG recombinase. *Nature* **569**: 79–84
- Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, Tu S, Vreven T, Zamore PD, Weng Z & Theurkauf WE (2014) The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* **157**: 1353–1363
- Zhao K, Cheng S, Miao N, Xu P, Lu X, Zhang Y, Wang M, Ouyang X, Yuan X, Liu W, Lu X, Zhou P, Gu J, Zhang Y, Qiu D, Jin Z, Su C, Peng C, Wang J-H, Dong M-Q, et al (2019) A Pandas complex adapted for piRNA-guided transcriptional silencing and heterochromatin formation. *Nat. Cell Biol.* **21**: 1261–1272

# **ANNEXES**

	<b><i>BX2<sup>OFF</sup></i></b>			<b><i>BX2<sup>ON</sup></i></b>						
<b>Sous-lignées</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>G1</b>	0/720	0/720	0/480	720/720	600/600	480/480	360/360	480/480	360/360	480/480
<b>G6</b>	0/600	0/840	0/960	720/720	840/840	240/240	480/480	420/420	480/480	480/480
<b>G7</b>	0/720	0/840	0/600	960/960	960/960	840/840	960/960	960/960	720/720	960/960
<b>G8</b>	0/960	0/780	0/420	960/960	720/720	nt	nt	nt	nt	nt
<b>G10</b>	0/720	0/120	0/600	nt	nt	nt	360/360	nt	600/600	600/600
<b>G11</b>	0/840	0/720	0/720	720/720	720/720	720/720	720/720	720/720	720/720	720/720
<b>G14</b>	nt	nt	nt	nt	nt	240/240	960/960	960/960	120/120	360/360
<b>G15</b>	0/480	0/1200	0/1200	960/960	1200/1200	nt	nt	nt	nt	nt
<b>G16</b>	0/1200	0/1200	0/1080	1080/1080	1440/1440	600/600	660/660	720/720	720/720	720/720
<b>G25</b>	0/960	0/960	0/960	1080/1080	960/960	360/360	720/720	720/720	960/960	nt
<b>G26</b>	0/1200	0/1200	0/1080	960/960	720/720	1200/1200	600/600	1200/1200	720/720	600/600
<b>G34</b>	0/840	0/840	0/960	960/960	960/960	960/960	960/960	960/960	960/960	960/960
<b>G35</b>	0/1440	0/1320	0/1440	1440/1440	1440/1440	1440/1440	1440/1440	1440/1440	1440/1440	1440/1440
<b>G36</b>	nt	0/1440	0/360	840/840	960/960	nt	nt	nt	nt	nt
<b>G37</b>	nt	0/1800	0/1800	1200/1200	1200/1200	nt	nt	nt	nt	nt
<b>G38</b>	nt	0/1440	0/1440	1440/1440	1440/1440	nt	nt	nt	nt	nt
<b>G39</b>	nt	0/720	0/960	960/960	960/960	nt	nt	nt	nt	nt

<b>G40</b>	nt	0/960	0/960	960/960	960/960	nt	nt	nt	nt	nt
<b>G41</b>	nt	0/1200	0/960	720/720	960/960	nt	nt	nt	nt	nt
<b>G43</b>	nt	0/960	0/960	960/960	960/960	960/960	840/840	960/960	960/960	960/960
<b>G44</b>	nt	0/960	0/960	960/960	960/960	nt	nt	nt	nt	nt
<b>G45</b>	nt	0/960	0/960	960/960	960/960	nt	nt	nt	nt	nt
<b>G46</b>	nt	0/1200	0/240	720/720	600/600	nt	nt	nt	nt	nt
<b>G47</b>	nt	0/720	0/1200	600/600	600/600	nt	nt	nt	nt	nt
<b>G48</b>	nt	0/1440	0/480	960/960	960/960	nt	nt	nt	nt	nt
<b>G49</b>	nt	0/1440	0/1320	960/960	600/600	nt	nt	nt	nt	nt
<b>G50</b>	nt	0/960	0/1440	600/600	840/840	960/960	960/960	960/960	960/960	960/960
<b>G51</b>	nt	0/1320	0/1440	960/960	960/960	nt	nt	nt	nt	nt
<b>G52</b>	nt	0/1320	0/1440	720/720	960/960	nt	nt	nt	nt	nt
<b>G53</b>	nt	0/1440	0/360	720/720	360/360	nt	nt	nt	nt	nt
<b>G54</b>	nt	0/1440	0/1440	960/960	960/960	nt	nt	nt	nt	nt
<b>G55</b>	0/1440	0/1440	0/1440	960/960	960/960	nt	nt	nt	nt	nt
<b>G56</b>	nt	0/480	0/1440	600/600	480/480	nt	nt	nt	nt	nt
<b>G62</b>	nt	0/1200	0/1680	600/600	720/720	nt	nt	nt	nt	nt
<b>Total</b>	0/12600	0/35580	0/33780	28920/28920	28920/28920	9000/9000	10020/10020	10500/10500	9720/9720	9240/9240
<b>%R</b>	0% n=81960			100% n=106320						

**Annexe 1 : Pourcentages de répression des sous-lignées  $BX2^{OFF}$   $P(TARGET)$  et  $BX2^{ON}$   $P(TARGET)$  à 25°C au cours des générations.** Les sous-lignées  $BX2^{OFF}$  et  $BX2^{ON}$  sont recombinées et portent le cluster  $BX2$  et un transgène  $PlacZ$  cible (transgène  $P(PZ)$ ) sur le même chromosome (voir Matériel et Méthodes). Parmi les sept sous-lignées  $BX2^{ON}$ , les lignées 1 et 2 sont issues des expériences de paramutation de  $BX2$  par le cluster  $TI$  (voir Introduction – le cluster  $BX2$ ) tandis que les sous-lignées 3 à 7 sont issues des expériences d'activation de  $BX2$  à 29°C décrites dans l'article I. Notons que les lignées  $BX2^{ON}$  issues de la paramutation et des expériences de thermo-activation possèdent les mêmes propriétés fonctionnelles et moléculaires (production de piRNAs, enrichissement en marques H3K9me3 et en Rhino). Les sous-lignées  $BX2^{OFF}$  ne présentent pas de répression du transgène cible (0%) à l'inverse des sous-lignées  $BX2^{ON}$  pour lesquelles il y a une répression complète du transgène cible (100%), ce qui indique que l'état épigénétique de  $BX2$ , qu'il soit inactif ou actif, est extrêmement stable au cours des générations à 25°C. Les valeurs indiquées représentent le nombre de chambres ovariennes réprimées sur le nombre total de chambres ovariennes à chaque génération. nt=non testé, %R=pourcentage de répression.

Sous-lignées	<i>BX2<sup>ON</sup></i>								<i>BX2<sup>OFF</sup></i>	
	1	2	3	4	5	6a	6b	7	1	2
<b>G1</b>	420/420	245/245	530/530	580/580	500/500	380/380	540/540	510/510	0/665	0/600
<b>G2</b>	649/650	385/385	600/600	695/700	787/800	975/975	699/700	555/555	0/350	0/300
<b>G3</b>	415/415	275/275	565/565	550/550	120/500	445/450	450/450	398/400	0/500	0/500
<b>G5</b>	720/960	1080/1080	960/960	1078/1080	1080/1080	1054/1080	1080/1080	1080/1080	0/1200	0/1080
<b>G6</b>	1680/1800	1800/1800	1680/1860	1800/1800	1669/1800	1800/1800	1800/1800	1200/1200	0/1680	0/1560
<b>G8</b>	1797/1800	1800/1800	1560/1560	1800/1800	1800/1800	1800/1800	1800/1800	1800/1800	0/1200	nt
<b>G10</b>	1680/1680	1800/1800	1800/1800	1800/1800	1680/1800	1800/1800	1800/1800	1800/1800	0/1200	0/1200
<b>G15</b>	1315/1320	840/840	1800/1800	1200/1200	1799/1800	2160/2160	1800/1800	1560/1560	0/840	0/480
<b>G20</b>	2999/3000	2880/2880	2880/2880	3000/3000	2370/3000	2960/3000	3000/3000	2760/2760	0/1200	0/1200
<b>Total</b>	11675/12045	11105/11105	12555/12555	12503/12510	11805/13080	13374/13445	12969/12970	11663/11665	0/8835	0/6920
<b>%R</b>	98.26% n=99375								0% n=15755	

**Annexe 2:** Pourcentages de répression de sous-lignées *BX2<sup>OFF</sup>* *P(TARGET)* et *BX2<sup>ON</sup>* *P(TARGET)* à 18°C au cours des générations, annexe à la table 1. Les valeurs représentent le nombre de chambres ovariennes réprimées sur le nombre total de chambres ovariennes à chaque génération. nt=non testé, %R=pourcentage de répression.



<b>Gene name</b>	<b>Gene symbol</b>	<b>CG number</b>	<b>Bloomington number</b>	<b>Gonads phenotype</b>	<b>Fertility</b>
absent, small, or homeotic discs 1	ash1	CG8887	33705	Normal	Fertile
			36803	Normal	Fertile
absent, small, or homeotic discs 2	ash2	CG6677	35388	Normal	Sterile
Ada2a-containing complex			32890	Normal	Fertile
component 2	Atac2	CG10414	53918	Atrophy	Sterile
Additional sex combs	Asx	CG8787	51677	Normal	Fertile
adrift	aft	CG5032	63639	Normal	Fertile
			33996	Normal	Fertile
alan shepard	shep	CG32423	38218	Normal	Fertile
			43545	Normal	Fertile
Andropin	Anp	CG1361	55385	Normal	Fertile
archipelago	ago	CG15010	34802	Normal	Fertile
Argonaute 2	AGO2	CG7439	34799	Normal	Fertile
			35232	Normal	Sterile
Argonaute 3	AGO3	CG40300	44543	Normal	Sterile
Argonaute-1	AGO1	CG6671	33727	Normal	Sterile
			35394	Atrophy	Sterile
arrest/bruno 1	aret/bru1	CG31762	38983	Atrophy	Sterile
			44483	Normal	Sterile
ATP-dependent chromatin assembly					
factor large subunit	Acf1	CG1966	35575	Normal	Fertile
Autophagy-related 1	Atg1	CG10967	44034	Normal	Fertile
Autophagy-related 17	Atg17	CG1347	36918	Normal	Fertile
Autophagy-related 18a	Atg18a	CG7986	34714	Normal	Fertile
Autophagy-related 2	Atg2	CG1241	35177	Normal	Fertile
Autophagy-related 8a	Atg8a	CG32672	34340	Normal	Fertile
B52	B52	CG10851	37519	Normal	Fertile
			32977	Normal	Fertile
basket	bsk	CG5680	35594	Normal	Fertile

			36643	Normal	Fertile
			53310	Normal	Fertile
belle	bel	CG9748	35185	Atrophy	Sterile
			35302	Normal	Fertile
blanks	blanks	CG10630	33667	Normal	ND
			34520	Normal	Fertile
brahma	brm	CG5942	35210	Normal	Fertile
			35211	Normal	Fertile
Brahma associated protein 60kD	Bap60	CG4303	32503	Normal	Fertile
			33954	Normal	Fertile
brivido-3	brv3	CG13762	33763	Normal	Fertile
			36774	Normal	Fertile
cap-n-collar	cnc	CG43286	32863	Normal	Fertile
			40854	Normal	Fertile
CG10418	CG10418	CG10418	44458	Atrophy	Sterile
CG10445	CG10445	CG10445	43137	Normal	Fertile
CG11266	CG11266	CG11266	44431	Normal	Fertile
CG11447	CG11447	CG11447	43207	Normal	Fertile
			42825	Normal	Fertile
			42912	Atrophy	Sterile
CG1239	CG1239	CG1239	44023	Atrophy	Sterile
			44506	Atrophy	Sterile
			55400	Normal	Fertile
CG12493	CG12493	CG12493	42791	Normal	Fertile
CG13472	CG13472	CG13472	36819	Normal	Fertile
CG13902	CG13902	CG13902	33702	Normal	Fertile
CG16817	CG16817	CG16817	41862	Normal	Fertile
CG16941	CG16941	CG16941	34840	Atrophy	Sterile
CG17454	CG17454	CG17454	43199	Atrophy	Sterile
CG17724	CG17724	CG17724	35676	Normal	Fertile
CG18537	CG18537	CG18537	53960	Normal	Fertile
CG2926	CG2926	CG2926	34941	Normal	ND

CG2982	CG2982	CG2982	33596	Normal	Fertile
CG3542	CG3542	CG3542	33711	Normal	Fertile
CG4078	CG4078	CG4078	32973	Normal	Fertile
CG4461	CG4461	CG4461	53298	Normal	Fertile
CG5001	CG5001	CG5001	32392	Normal	Fertile
CG5728	CG5728	CG5728	36592	Normal	Fertile
CG6283	CG6283	CG6283	51498	Normal	Fertile
CG6415	CG6415	CG6415	51867	Normal	Fertile
CG6506	CG6506	CG6506	42552	Normal	Fertile
CG6841	CG6841	CG6841	51909	Atrophy	Sterile
CG7185	CG7185	CG7185	34804	Atrophy	Sterile
CG7409	CG7409	CG7409	51433	Normal	Fertile
CG7878	CG7878	CG7878	35229	Normal	Fertile
CG8273	CG8273	CG8273	34805	Normal	Fertile
CG8778	CG8778	CG8778	36793	Normal	Fertile
CG8920	CG8920	CG8920	34738	Normal	Fertile
CG9684	CG9684	CG9684	36880	Normal	Fertile
CG9925	CG9925	CG9925	35811	Normal	Fertile
Chromatin assembly factor 1, p55 subunit	Caf1-55	CG4236	34069	Atrophy	Sterile
Chromosome associated protein D3	Cap-D3	CG31989	36615	Normal	Fertile
CMP-sialic acid synthase	Csas	CG32220	54843	Normal	Fertile
Complex I intermediate-associated protein, 30 kDa	CIA30	CG7598	55660	Normal	Fertile
cup	cup	CG11181	35406	Normal	Sterile
Cyclin T	CycT	CG6292	32976	Atrophy	Sterile
			35168	Normal	Fertile
			34982	Atrophy	Sterile
Cyclin-dependent kinase 9	Cdk9	CG5179	35323	Small ovaries	Fertile
			41932	Normal	Fertile
Darkener of apricot	Doa	CG42320	50903	Normal	Sterile

Death-associated inhibitor of apoptosis 1 / thread	Diap1 / th	CG12284	33597	Normal	Fertile
			33618	Normal	Fertile
decapentaplegic	dpp	CG9885	33767	Normal	Fertile
			35214	Normal	Fertile
Desaturase 1	Desat1	CG5887	35591	Normal	Fertile
			37512	Normal	Fertile
diaphanous	dia	CG1768	33424	Normal	Sterile
			35479	Atrophy	Sterile
Dicer-2	Dcr2	CG6493	33656	Normal	Fertile
DISCO Interacting Protein 1	DIP1	CG17686	35226	Normal	Fertile
			35333	Normal	Fertile
DISCO Interacting Protein 2	DIP2	CG7020	34918	Normal	Fertile
			42598	Normal	Fertile
DnaJ-like-1	Dnaj-1	CG10578	32899	Normal	Fertile
			32978	Normal	Fertile
Dodeca-satellite-binding protein 1	Dp1	CG5170	32872	Normal	Fertile
			34827	Atrophy	Sterile
			38385	Atrophy	Sterile
domino	dom	CG9696	38941	Normal	Fertile
			40914	Atrophy	Sterile
			41674	Atrophy	Sterile
			32934	Normal	Sterile
dorsal	dl	CG6667	36650	Normal	Sterile
			38905	Normal	Sterile
DP transcription factor	Dp	CG4654	33372	Normal	Fertile
			33657	Normal	ND
droscha	droscha	CG8730	35233	Normal	Fertile
			32444	Atrophy	Sterile
eIF4AIII	eIF4AIII	CG7483	32907	Atrophy	Sterile
			38202	Atrophy	Sterile
Ejaculatory bulb protein III	PebIII	CG11390	55933	Normal	Fertile

encore	enc	CG10847	42797	Normal	Fertile
Enhancer of bithorax	E(bx)	CG32346	33658	Normal	Substerile
Enhancer of Polycomb	E(Pc)	CG7776	35271	Atrophy	Sterile
Enhancer of zeste	E(z)	CG6502	33659	Normal	ND
			36068	Normal	Sterile
enoki mushroom	enok	CG11290	40917	Normal	Fertile
			41664	Normal	Fertile
			42941	Normal	Fertile
Ets at 97D	Ets97D	CG6338	35749	Normal	Fertile
			36635	Normal	Fertile
female sterile (1) Yb	fs(1)Yb	CG2706	35181	Normal	Fertile
			35301	Normal	Fertile
flower	fwe	CG6151	34157	Normal	Fertile
Fmr1	Fmr1	CG6203	35200	Normal	Fertile
Frost	Fst	CG9434	33376	Normal	Fertile
G9a	G9a	CG2995	34817	Normal	Fertile
			33981	Normal	Fertile
Gcn5 acetyltransferase / Pcaf	Gcn5 / Pcaf	CG4107	35601	Normal	Sterile
Glutathione S transferase E1 half pint / pUf68	GstE1 hfp / pUf68	CG5164 CG12085	36878	Normal	Fertile
			34785	Normal	Fertile
hangover	hang	CG32575	35674	Normal	Fertile
			41870	Normal	Sterile
haywire	hay	CG8019	53345	Atrophy	Sterile
Heat shock 70-kDa protein cognate 3	Hsc70-3	CG4147	32402	Normal	Fertile
Heat shock factor	Hsf	CG5748	41581	Normal	Fertile
Heat shock gene 67Ba	Hsp67Ba	CG4167	41962	Normal	Fertile
			53007	Normal	Fertile
Heat shock gene 67Bc	Hsp67Bc	CG4190	35452	Normal	Fertile
			42607	Normal	Fertile
Heat shock protein 22	Hsp22	CG4460	41709	Normal	Fertile
			51397	Normal	Fertile

Heat shock protein 23	Hsp23	CG4463	44029	Normal	Fertile
Heat shock protein 26	Hsp26	CG4183	35408	Normal	Fertile
			42610	Normal	Fertile
Heat shock protein 27	Hsp27	CG4466	33007	Normal	Fertile
			33922	Normal	Fertile
Heat shock protein 60A	Hsp60A	CG12101	34729	Normal	Fertile
Heat shock protein 60B	Hsp60B	CG2830	34729	Normal	Fertile
			41709	Normal	Fertile
Heat shock protein 67Bb	Hsp67Bb	CG4456	51397	Normal	Fertile
Heat shock protein 68	Hsp68	CG5436	50637	Normal	Fertile
Heat shock protein 70Aa	Hsp70Aa	CG31366	42639	Normal	Fertile
Heat shock protein 70Ab	Hsp70Ab	CG18743	35663	Normal	Fertile
			32997	Normal	Fertile
Heat shock protein 70Ba	Hsp70Ba	CG31449	43289	Normal	Fertile
			35672	Normal	Fertile
			32997	Normal	Fertile
Heat shock protein 70Bb	Hsp70Bb	CG31359	33948	Normal	Fertile
			32997	Normal	Fertile
Heat shock protein 70Bc	Hsp70Bc	CG6489	35697	Normal	Fertile
			42626	Normal	Fertile
			32996	Normal	Sterile
Heat shock protein 83	Hsp83	CG1242	33947	Atrophy	Sterile
Heat shock protein cognate 1	Hsc70-1	CG8937	34527	Normal	Fertile
			42014	Normal	Fertile
Heat shock protein cognate 2	Hsc70-2	CG7756	44485	Normal	Fertile
			32997	Normal	Fertile
			34836	Atrophy	Sterile
Heat shock protein cognate 4	Hsc70-4	CG4264	35684	Atrophy	Sterile
Helicase at 25E	Hel25E	CG7269	33666	Atrophy	Sterile
hemipterous	hep	CG4353	35210	Normal	Sterile
Heterochromatin Protein 1b	HP1b	CG7041	32401	Normal	Fertile

Heterogeneous nuclear ribonucleoprotein at 27C	Hrb27C	CG10377	33716	Normal	Sterile
Heterogeneous nuclear ribonucleoprotein at 87F	Hrb87F	CG12749	52937	Normal	Fertile
Heterogeneous nuclear ribonucleoprotein at 98DE	Hrb98DE	CG9983	32351	Atrophy	Sterile
Heterogeneous nuclear ribonucleoprotein K/bancal	HnRNP-K/bl	CG13425	42540	Normal	Fertile
Host cell factor	Hcf	CG1710	32453	Normal	Fertile
			36799	Normal	Fertile
			32997	Normal	Fertile
Hsp70Bbb	Hsp70Bbb	CG5834	33000	Normal	Fertile
			33916	Atrophy	Sterile
			32845	Normal	ND
Imitation SWI	Iswi	CG8625	51931	Normal	Fertile
Insulin-like peptide 5 janus A	Ilp5 janA	CG33273 CG7933	33683	Normal	Fertile
			41846	Normal	Fertile
			41592	Normal	Fertile
JIL-1 kinase	JIL-1	CG6297	42571	ND	ND
			55875	Normal	Fertile
Jumonji, AT rich interactive domain 2	jarid2	CG3654	32891	Normal	Fertile
Keap1	Keap1	CG3962	40932	Normal	Sterile
			34908	Normal	Fertile
kismet	kis	CG3696	36597	Normal	Sterile
			44542	Normal	Sterile
kohtalo	kto	CG8491	34588	Normal	Sterile
			35230	Normal	Sterile
krimper	krimp	CG15707	35231	Normal	Sterile
			37511	Normal	Sterile
Lamin lethal (2) essential for life	Lam l(2)efl	CG6944 CG4533	36617	Normal	Fertile
			41724	Normal	Fertile

			51816	Normal	Fertile
lethal (3) 72Ab	l(3)72Ab	CG5931	34024	Atrophy	Sterile
			50716	Atrophy	Sterile
Ligase4	lig4	CG12176	51933	Normal	Fertile
little imaginal discs	lid	CG9088	36652	Normal	Fertile
			35706	Normal	Fertile
locomotion defects	loco	CG5248	32456	Normal	Fertile
loki	lok	CG10895	35152	Normal	Sterile
			32955	Normal	Fertile
			33427	Normal	Fertile
			34779	Normal	Fertile
loquacious	loqs	CG6866	34780	Normal	Fertile
			34781	Normal	Fertile
			34782	Normal	Fertile
			34851	Normal	Fertile
Lysine (K)-specific demethylase 2	Kdm2	CG11033	33699	Normal	Fertile
Lysine (K)-specific demethylase 4A	Kdm4A	CG15835	34629	Normal	Fertile
Lysine (K)-specific demethylase 4B	Kdm4B	CG33182	35676	Normal	Fertile
mago nashi	mago	CG9401	35453	Normal	Fertile
maleless	mle	CG11680	34864	Normal	ND
			33675	Atrophy	Sterile
maternal expression at 31B	Me31b	CG4916	38923	Normal	Fertile
Max	Max	CG9648	40851	Normal	Fertile
meiotic 41	mei-41	CG4252	35371	Normal	Fertile
			41934	Normal	Fertile
meiotic 9	mei-9	CG3697	55313	Normal	Fertile
Mekk1	Mekk1	CG7717	35402	Normal	Fertile
menage a trois	metro	CG30021	35810	Normal	Fertile
Menin 1	Mnn1	CG13778	35150	Atrophy	Sterile
			51862	Normal	Fertile
methuselah	mth	CG6936	36823	Normal	Fertile
Methyltransferase 2	Mt2	CG10692	38224	Normal	Fertile



			42906	Normal	Fertile
			35398	Atrophy	Sterile
Mi-2	Mi-2	CG8103	33419	Atrophy	Sterile
			51774	Atrophy	Sterile
			34919	Normal	Fertile
moira	mor	CG18740	35630	Normal	Fertile
			35662	Normal	Sterile
mushroom-body expressed	mub	CG7437	34870	Normal	Fertile
nejire	nej	CG15319	36682	Normal	Fertile
			37489	Normal	Sterile
Neurofibromin 1	Nf1	CG8318	53322	Normal	Fertile
Nucleosome remodeling factor - 38kD	Nurf-38	CG4634	35444	Normal	Sterile
			35447	Normal	Fertile
osa	osa	CG7467	38285	Normal	Fertile
			34744	Normal	Fertile
p38a MAP kinase	p38a	CG5475	35244	Normal	Fertile
			35252	Normal	Fertile
p38b MAP kinase	p38b	CG7393	36814	Normal	Fertile
			41638	Normal	Fertile
p53	p53	CG10895	41720	Normal	Fertile
			34690	Normal	Fertile
pacman	pcm	CG3291	51835	Normal	Fertile
painless	pain	CG15860	35162	Normal	Fertile
			42499	Normal	Fertile
pancreatic eIF-2 $\alpha$ kinase	PEK	CG2087	34932	Normal	Fertile
			35450	Normal	Fertile
			37513	Normal	Fertile
			38216	Normal	Fertile
partner of drosha	pasha	CG1800	33972	Normal	ND
peanuts	pea	CG8241	32838	Atrophy	Sterile
P-element induced wimpy testis	piwi	CG6122	33724	Atrophy	Sterile

			34866	Atrophy	Sterile
			37483	Normal	Sterile
P-element somatic inhibitor	Psi	CG8912	34825	Normal	Fertile
Phosphoglycerate mutase 5	Pgam5	CG14816	33346	Normal	Fertile
Phosphatidylinositol 3 kinase 59F	Pi3K59F	CG5373	33384	Normal	Fertile
Poly-(ADP-ribose) polymerase	Parp	CG40411	36056	Normal	Fertile
			35792	Normal	Fertile
			33622	Normal	Fertile
Polycomb	Pc	CG32443	33964	Normal	Fertile
			36070	Normal	Fertile
			33945	Normal	Fertile
Polycomblike	Pcl	CG5109	33946	Normal	Fertile
polyhomeotic proximal	ph-p	CG18412	33669	Normal	Fertile
pontin	pont	CG4003	50972	Atrophy	Sterile
protein partner of snf	pps	CG6525	38529	Normal	Fertile
			38912	Normal	Fertile
			53019	Normal	Fertile
puckered	puc	CG7850	34392	Normal	Fertile
			36085	Normal	Sterile
			44038	Normal	Fertile
			35195	Normal	Sterile
punt	put	CG7904	35701	Normal	Fertile
			39025	Normal	Fertile
pyrexia	pyx	CG17142	51836	Normal	Fertile
			37475	Normal	Fertile
qin	qin	CG43726	41662	Normal	ND
r2d2	r2d2	CG7138	34784	Normal	Fertile
Rab5	Rab5	CG3664	34832	Normal	Sterile
			51847	Normal	Fertile
raspberry	ras	CG1799	51717	Small ovaries	Fertile
Ras-related protein interacting with calmodulin	Ric	CG8418	41819	Normal	Fertile

Recombination repair protein 1	Rrp1	CG3178	35420	Normal	Fertile
refractory to sigma P	ref(2)P	CG10360	33978	Normal	Fertile
			36111	Normal	Fertile
Rev1	Rev1	CG12189	36654	Normal	Fertile
Ribosomal protein L8	Rpl8	CG1263	50610	Normal	Fertile
Rm62	Rm62	CG10279	34829	Normal	Fertile
RNA polymerase II subunit Rpb4	Rpb4	CG43662	50905	Atrophy	Sterile
Rox8	Rox8	CG5422	32472	Normal	Fertile
			33725	Normal	Fertile
Rpd3	Rpd3	CG7471	34846	Normal	Fertile
			34809	Letal	ND
Rrp6	Rrp6	CG7292	42064	Normal	Sterile
runt	run	CG1849	34707	Normal	Fertile
sans fille	snf	CG4528	34593	Normal	Fertile
			51459	Atrophy	Sterile
Scaffold attachment factor B	Saf-B	CG6995	51759	Normal	Sterile
Scamp	Scamp	CG9195	38277	Normal	Fertile
scute	sc	CG3827	41594	Normal	Fertile
Secreted protein, acidic, cysteine-rich	SPARC	CG6378	40885	Normal	Fertile
Serine palmitoyltransferase subunit I	Spt-I	CG4016	55685	Normal	Fertile
SET domain binding factor	Sbf	CG6939	32419	Normal	Fertile
			44004	Normal	Fertile
			34393	Atrophy	Sterile
Sex lethal	Sxl	CG43770	38195	Atrophy	Sterile
shibire	shi	CG18102	36921	Normal	Sterile
			32904	Normal	Fertile
shotgun	shg	CG3722	38207	Normal	Sterile
			37507	Normal	Fertile
shriveled	shv	CG4164	54797	Normal	Fertile
shutdown	shu	CG4735	35454	Normal	Sterile
sisterless A	sisA	CG1641	55181	Normal	Fertile

skuld	skd	CG9936	34630	Normal	Fertile
slipper	slpr	CG2272	32948	Normal	Fertile
			41605	Normal	Fertile
small ribonucleoprotein particle U1 subunit 70K	snRNP-U1-70K	CG8749	33396	Atrophy	Sterile
small ribonucleoprotein particle U1 subunit C	snRNP-U1-C	CG5454	34822	Atrophy	Sterile
smaug	smg	CG5263	35477	Normal	Fertile
Snf5-related 1	Snr1	CG1064	32372	Normal	Sterile
			32439	ND	ND
spaghetti squash	sqh	CG3595	33892	Normal	Fertile
			38222	Normal	Fertile
Spf45	Spf45	CG17540	41954	Normal	Fertile
spindle A	Spn-A	CG7948	38898	Normal	Fertile
			51936	Normal	Fertile
spindle E	spn-E	CG3158	35303	Normal	Sterile
Splicing factor 2	SF2	CG6987	32367	Normal	Fertile
Spt3	Spt3	CG3169	35148	Normal	Fertile
Spt4	Spt4	CG12372	32896	Atrophy	Sterile
Spt5	Spt5	CG7626	34837	Atrophy	Sterile
Spt6	Spt6	CG12225	32373	Atrophy	Sterile
squid	sqd	CG16901	35627	Normal	Fertile
			53891	Normal	Fertile
staufen	stau	CG5753	35690	Normal	Fertile
			43187	Normal	Sterile
stonewall	stw1	CG3836	35415	Atrophy	Sterile
Su(var)2-HP2	Su(var)2-HP2	CG12864	38255	Atrophy	Sterile
			33906	Normal	Fertile
suppressor of Hairy wing	su(Hw)	CG8573	34006	Normal	Fertile
Suppressor of sable	su(sable)	CG6222	33982	Normal	Fertile
Suppressor of Under-Replication	SuUR	CG7869	36893	Normal	Fertile
Suppressor of variegation 205	Su(var)205	CG8409	33400	Atrophy	Sterile

			36792	Atrophy	Sterile
			32853	Normal	Fertile
Suppressor of variegation 3-3	Su(var)3-3	CG17149	33726	Normal	Fertile
			36867	Normal	Fertile
			32914	Atrophy	Sterile
Suppressor of variegation 3-9	Su(var)3-9	CG43664	33401	Atrophy	Sterile
			43661	Atrophy	Sterile
Suppressor of zeste 2 survival motor neuron	Su(z)2 Smn	CG3905 CG16725	33403	Normal	Fertile
			36621	Normal	Fertile
TBP-associated factor 1	Taf1	CG17603	32421	Atrophy	Sterile
			35314	Atrophy	Sterile
TBP-associated factor 10	Taf10	CG2859	35239	Normal	Fertile
TBP-associated factor 12	Taf12	CG17358	34852	Normal	Fertile
			35427	Normal	Sterile
TBP-associated factor 4	Taf4	CG5444	50985	Atrophy	Sterile
TBP-associated factor 5	Taf5	CG7704	35367	Atrophy	Sterile
TBP-associated factor 7	Taf7	CG2670	55216	Normal	Fertile
			36879	Normal	Sterile
tejas	tej	CG8589	41929	Normal	Sterile
			35166	Normal	Fertile
			35653	Normal	Fertile
thickveins	tkv	CG14026	40937	Atrophy	Sterile
			41904	Normal	Fertile
Thioredoxin peroxidase 2	Jafrac2	CG1274	56043	Normal	Fertile
transcriptional Adaptor 2a	Ada2a	CG43663	50905	Atrophy	Sterile
transcriptional Adaptor 2b	Ada2b	CG9638	35334	Normal	Fertile
transcriptional Adaptor 3 / diskette	Ada3 / dik	CG7098	32451	Normal	Fertile
Transient receptor potential cation channel A1	TrpA1	CG5751	36780	Normal	Fertile
trithorax	trx	CG8651	33703	Normal	Fertile
tsunagi	tsu	CG8781	36585	Atrophy	Sterile
tudor	tud	CG9450	42800	Normal	Fertile

Tudor staphylococcal nuclease	Tudor-SN	CG7008	34865	Normal	ND
			53244	Normal	Fertile
Turandot A	TotA	CG31509	55378	Normal	Fertile
			58357	Normal	Fertile
Turandot C	TotC	CG31508	51407	Normal	Fertile
U2 small nuclear riboprotein auxiliary factor 38	U2af38	CG3582	50561	Atrophy	Sterile
U2 small nuclear riboprotein auxiliary factor 50	U2af50	CG9998	50521	Normal	Sterile
Upf1	Upf1	CG1559	43144	Normal	Sterile
Upf3	Upf3	CG11184	44565	Normal	ND
upSET	upSET	CG9007	51447	Normal	Fertile
Utx histone demethylase	Utx	CG5640	34076	Normal	Fertile
vasa	vas	CG46283	32434	Normal	Sterile
			35183	Normal	Sterile
vasa intronic gene	vig	CG4170	35184	Normal	Fertile
Victoria	Victoria	CG33117	55953	Normal	Fertile
walrus	wal	CG8996	34915	Normal	Fertile
XNP	XNP	CG4548	32894	Normal	Fertile

**Annexe 3 : Liste complète des gènes testés dans le crible génétique ‘perte de fonction’.** Pour chaque gène, le nom complet, le symbole et le numéro CG (correspondant à un nombre de référence unique permettant d’identifier un gène) sont indiqués. Pour certains gènes, il peut exister plusieurs lignées (référencées par un ‘numéro Bloomington’) pour lesquelles le transgène *UAS-gene<sup>KD</sup>*, et en conséquence le shRNA produit, sont différents. Ainsi, pour 272 gènes testés il y a 402 shRNA différents. Enfin, pour toutes les lignées Bloomington testées, le phénotype des gonades et la fertilité (ponte + développement des œufs) sont analysés chez les descendants. ND : non défini.

Parental cross <i>mat x pat</i>	Progeny genotype	Ovariole number	Non repressed egg chamber number	Repressed egg chamber number	<i>BX2 conversion</i> (%)
<i>hfp<sup>13/+</sup> x hfp<sup>9/+</sup></i>	<i>hfp<sup>9/+</sup></i>	95	475	0	0
	<i>hfp<sup>13/+</sup></i>	476	2380	0	0
	<i>hfp<sup>13/9</sup></i>	232	750	257	25.5
<i>hfp<sup>9/+</sup> x hfp<sup>13/+</sup></i>	<i>hfp<sup>9/+</sup></i>	401	2005	0	0
	<i>hfp<sup>13/+</sup></i>	264	1320	0	0
	<i>hfp<sup>9/13</sup></i>	141	442	248	35.9
<i>hfp<sup>13/+</sup> x hfp<sup>13/+</sup></i>	<i>hfp<sup>13/+</sup></i>	1369	6845	0	0
	<i>hfp<sup>13/13</sup></i>	651	3185	70	2.2
<i>hfp<sup>9/+</sup> x hfp<sup>9/+</sup></i>	<i>hfp<sup>9/+</sup></i>	557	2779	6	0.2
	<i>hfp<sup>9/9</sup></i>	196	16	954	98.4

**Annexe 4: Combinaison de différents allèles mutants d'*hfp* confirmant l'effet observé en contexte *hfp GLKD* sur la conversion de *BX2*.** *hfp<sup>9</sup>* et *hfp<sup>13</sup>* sont deux allèles hypomorphes permettant la survie de l'adulte (Van Buskirk & Schüpbach, 2002). Dans le croisement parental, des femelles *w*; *hfp<sup>mat</sup> / hu* sont croisées par des mâles *w<sup>1118</sup>*; *BX2<sup>OFF</sup>*, *P(TARGET) / Cy* ; *hfp<sup>pat</sup> / Ubx*. Pour plus de clarté, seuls les allèles maternels et paternels de *hfp* sont représentés dans la table. A la génération suivante, les descendants des différents génotypes sont analysés par coloration β-Galactosidase, permettant d'estimer le pourcentage de répression du transgène *P(TARGET)* pour chaque chambre ovarienne. Le pourcentage de conversion de *BX2* est ainsi calculé en divisant le nombre de chambres réprimées par le nombre total de chambres ovariennes. Les chromosomes balanciers de type *hu*, *Cy* et *Ubx* permettant de sélectionner les différents génotypes grâce à des marqueurs phénotypiques dominants (voir 'Balanceurs et transgènes' dans le matériels et méthodes). *BX2* est activé pour la production de piRNAs aussi bien dans le cas de combinaisons homoalléliques ou hétéroalléliques. De manière intéressante, le pourcentage d'activation de *BX2* varie d'une combinaison à l'autre. L'allèle hypomorphe le plus fort (*hfp<sup>9</sup>*) affecte davantage le pourcentage de conversion de *BX2* (environ 100% chez les homozygotes *hfp<sup>9</sup> / hfp<sup>9</sup>*) que l'allèle hypomorphe le plus faible (*hfp<sup>13</sup>*, environ 2% pour les homozygotes *hfp<sup>13</sup> / hfp<sup>13</sup>*). Les combinaisons hétérozygotes sont associées à des pourcentages de conversion de *BX2* intermédiaires.

<b>Gene name</b>	<b>Gene symbol</b>	<b>CG number</b>	<b>Bloomington number</b>	<b>Gonads phenotype</b>	<b>Fertility</b>
antimeros	atms (Paf1)	CG2503	12114	Normal	Fertile
antisense RNA:CR45485	asRNA:CR45485	CR45485	62273	Normal	Fertile
ATP-dependent chromatin assembly factor large subunit	Acf1	CG1966	35575	Normal	Fertile
Carbonic anhydrase 2	CAH2	CG6906	41836	Normal	Fertile
			65081	Normal	Fertile
CG12054	CG12054	CG12054	50511	Normal	Fertile
			50910	Normal	Fertile
CG13397	CG13397	CG13397	51808	Normal	Fertile
CG14131	CG14131	CG14131	63716	Normal	Fertile
CG1673	CG1673	CG1673	38363	Normal	Fertile
CG17544	CG17544	CG17544	64632	Normal	Fertile
CG3036	CG3036	CG3036	43179	Normal	Fertile
CG31075	CG31075	CG31075	50654	Normal	Fertile
			62535	Normal	Fertile
CG31262	CG31262	CG31262	53970	Normal	Fertile
CG40006	CG40006	CG40006	34691	Normal	Fertile
CG40160	CG40160	CG40160	44268	Normal	Fertile
CG40178	CG40178	CG40178	44578	Normal	Fertile
CG42542	CG42542	CG42542	35736	Normal	Fertile
CG42588	CG42588	CG42588	63698	Normal	Fertile
			32332	Normal	Fertile
CG4267	CG4267	CG4267	56900	Normal	Fertile
CG43373	CG43373	CG43373	64884	Atrophy	Sterile
CG4404	CG4404	CG4404	62415	Normal	Fertile
CG5694	CG5694	CG5694	36655	Normal	Fertile
			62485	Normal	Fertile
CG6175	CG6175	CG6175	62516	Normal	Fertile



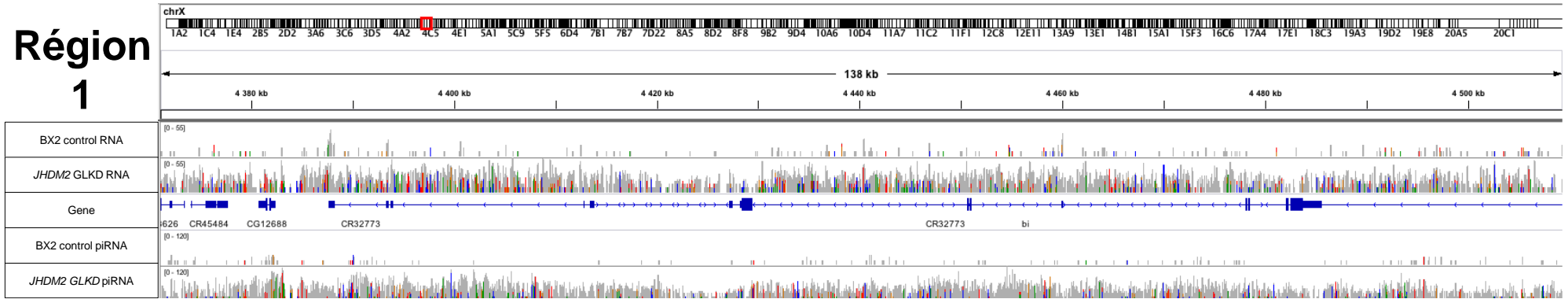
CG7255	CG7255	CG7255	58352	Normal	Fertile
CG9914	CG9914	CG9914	62273	Normal	Fertile
Chitinase 2	Cht2	CG2054	35717	Normal	Sterile
			60369	Normal	Fertile
			33619	Normal	Fertile
corkscrew	csw	CG3954	35215	Normal	Sterile
			35638	Normal	Sterile
			60448	Normal	Fertile
crossover suppressor on 3 of Gowen	c(3)G	CG17604	62969	Normal	Fertile
C-terminal Binding Protein	CtBP	CG7583	32889	Normal	Sterile
Cyclic-AMP response element binding protein A	CrebA	CG7450	42562	Normal	Fertile
Cyclin-dependent kinase 4	Cdk4	CG5072	36060	Normal	Fertile
			57031	Normal	Fertile
Cystathionine beta-synthase	Cbs	CG1753	36767	Normal	Fertile
			41877	Normal	Fertile
Cytochrome P450-4d8	Cyp4d8	CG4321	58019	Atrophy	Sterile
D-2-hydroxyglutaric acid dehydrogenase	D2hgdh	CG3835	53355	Normal	Fertile
Dead box protein 80 deadhead	Dbp80 dhd	CG17023 CG4193	34682	Normal	Fertile
			41857	Normal	Fertile
DnaJ-like-2	Droj2	CG8863	36089	Normal	Fertile
			32903	Normal	Fertile
			33975	Normal	Fertile
			38907	Normal	Fertile
Dual oxidase	Duox	CG3131	38916	Normal	Fertile
			36876	Normal	ND
			43990	Normal	Fertile
eukaryotic translation initiation factor 4E homologous protein	eIF4EHP	CG33100	43157	Normal	Fertile
flower	fwe	CG6151	43157	Normal	Fertile
Glutathione S transferase	GstE1	CG5164	36878	Normal	Fertile

E1 Glutathione S transferase E6	GstE6	CG17530	65197	Normal	Fertile
hephaestus	heph	CG31000	35669	Normal	Sterile
HSPB1 associated protein 1	HSPBAP1	CG43320	55655	Normal	Fertile
Imaginal disc growth factor 2	Idgf2	CG4475	34606	Normal	Fertile
Inwardly rectifying potassium channel 1	Irk1	CG44159	55935	Normal	Fertile
Isocitrate dehydrogenase Jabba	Idh Jabba	CG7176 CG42351	42644	Normal	Fertile
jing interacting gene regulatory 1	jjgr1	CG17383	41708	Normal	Sterile
JmjC domain-containing histone demethylase 2	JHDM2	CG8165	36852	Normal	Fertile
Juvenile hormone epoxide hydrolase 3	Jheh3	CG15106	58173	Normal	Fertile
Kank	Kank	CG10249	32975	Normal	Sterile
kayak	kay	CG33956	60021	Normal	Fertile
Leucine zipper and EF- hand containing transmembrane protein 1	Letm1	CG4589	33379	Normal	Sterile
longitudinals lacking Major Facilitator Superfamily Transporter 17	lola MFS17	CG12052 CG40263	37502	Atrophy	Sterile
Mucin related 18B	Mur18B	CG7874	35721	Normal	Fertile
ovarian tumor p24-related-2	otu p24-2	CG12743 CG33105	44033	Normal	Fertile
Prip	Prip	CG7777	56957	Normal	Fertile
rhino	rhi	CG10683	34065	Atrophy	Sterile
			40839	Normal	Fertile
			44464	Normal	Fertile
			50695	Normal	Fertile
			34071	Normal	Fertile

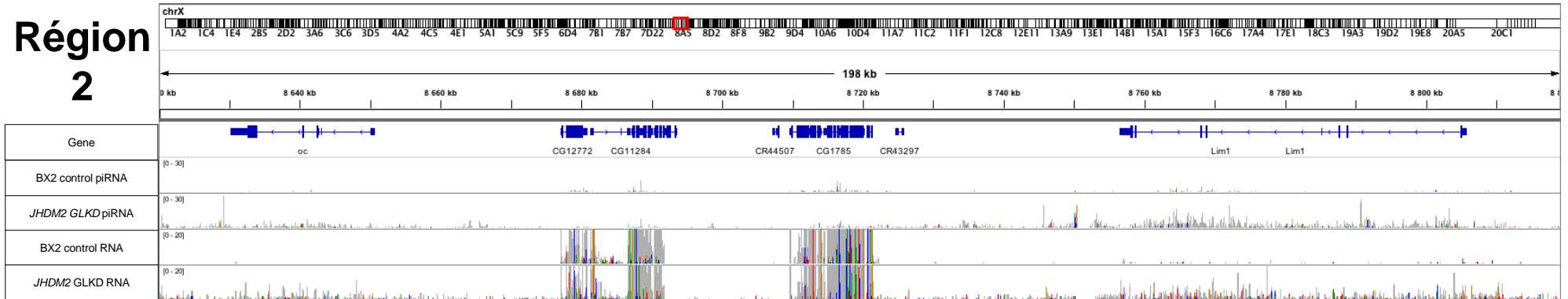
RhoGAP71E	RhoGAP71E	CG32149	35171	Normal	Sterile
rho-type guanine exchange factor	rtGEF	CG10043	32417	Normal	Fertile
rotated abdomen	rt	CG6097	32947	Normal	Fertile
Sin3A	Sin3A	CG8815	51805	Normal	Fertile
Sorbitol dehydrogenase-2	Sodh-2	CG4649	32368	Normal	Fertile
Thor	Thor	CG8846	53353	Normal	Few fertile
Uncoordinated 115a	Unc-115a	CG31352	36667	Normal	Fertile
Uncoordinated 115b	Unc-115b	CG31332	36815	Normal	Fertile
$\beta$ -Mannosidase	$\beta$ -Man	CG12582	40839	Normal	Fertile
			40839	Normal	Fertile
			53272	Normal	Fertile

**Annexe 5 : Liste des gènes cibles de *hfp* testés avec le crible génétique ‘perte de fonction’.** Pour chaque gène, le nom complet, le symbole et le numéro CG (correspondant à un nombre de référence unique permettant d’identifier un gène) sont indiqués. Pour certains gènes, il peut exister plusieurs lignées (référencées par un ‘numéro Bloomington’) pour lesquelles le transgène *UAS-gene<sup>KD</sup>*, et en conséquence le shRNA produit, sont différents. Ainsi, pour 69 gènes testés il y a 85 shRNA différents. Enfin, pour toutes les lignées Bloomington testées, le phénotype des gonades et la fertilité (ponte + développement des œufs) sont analysés chez les descendants. ND : non défini.

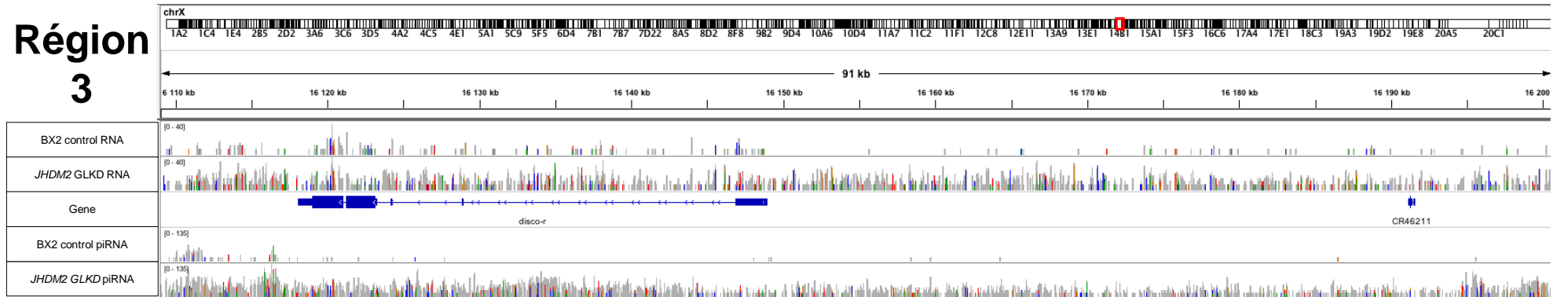
# Région 1



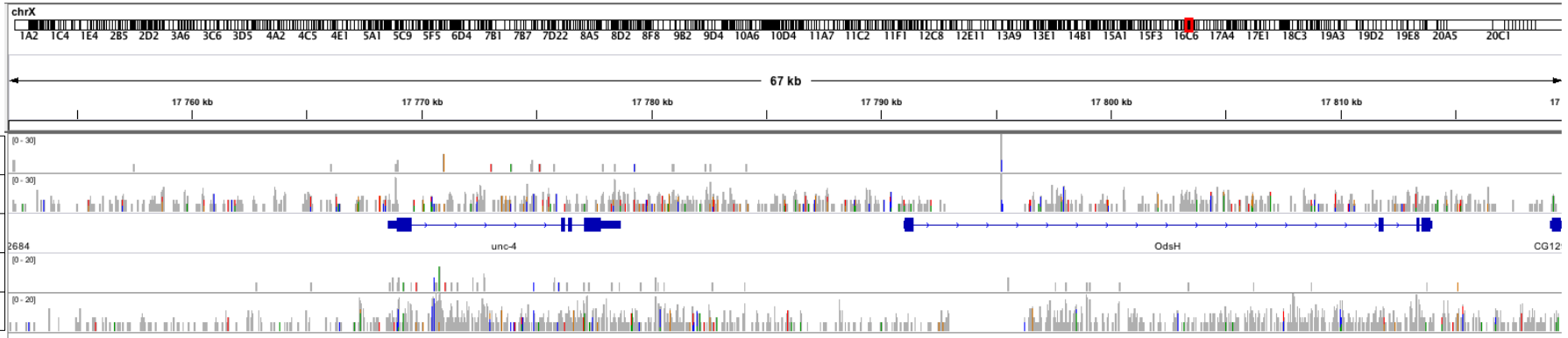
# Région 2



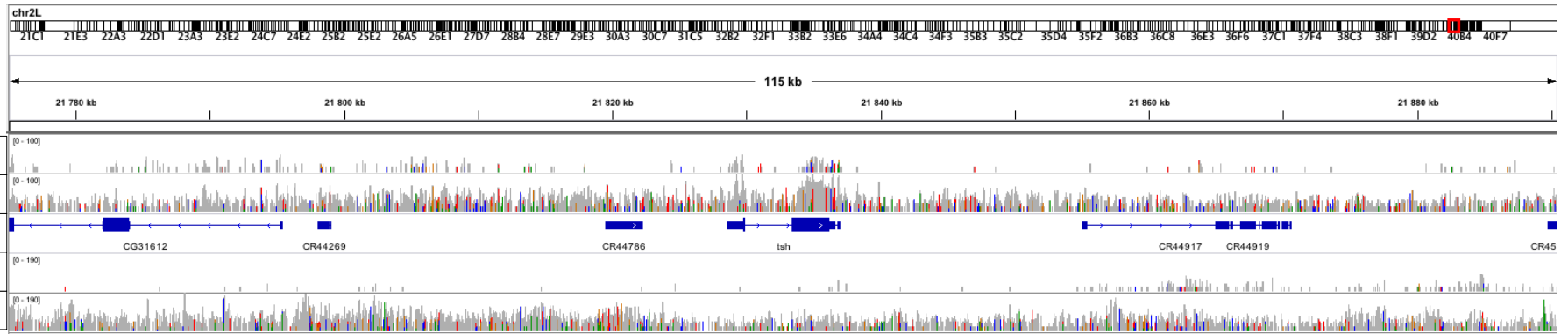
# Région 3



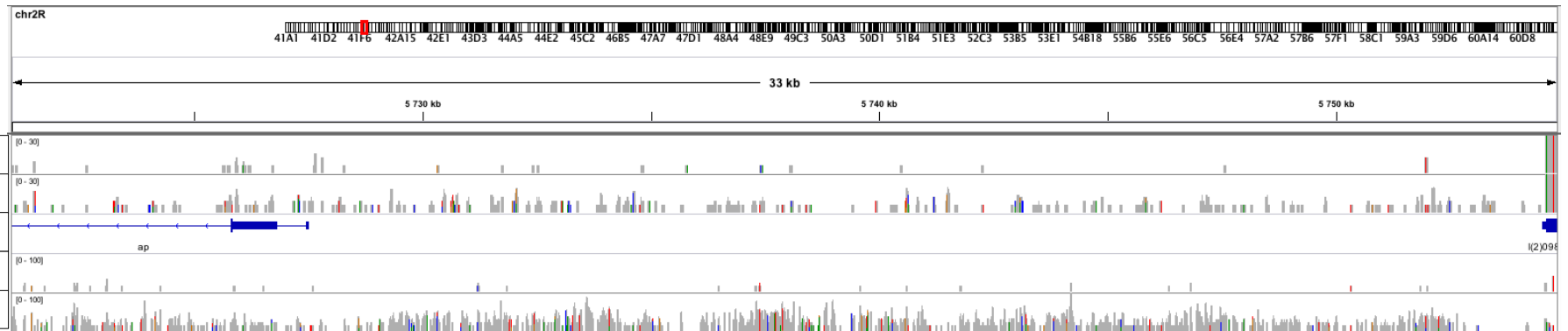
# Région 4



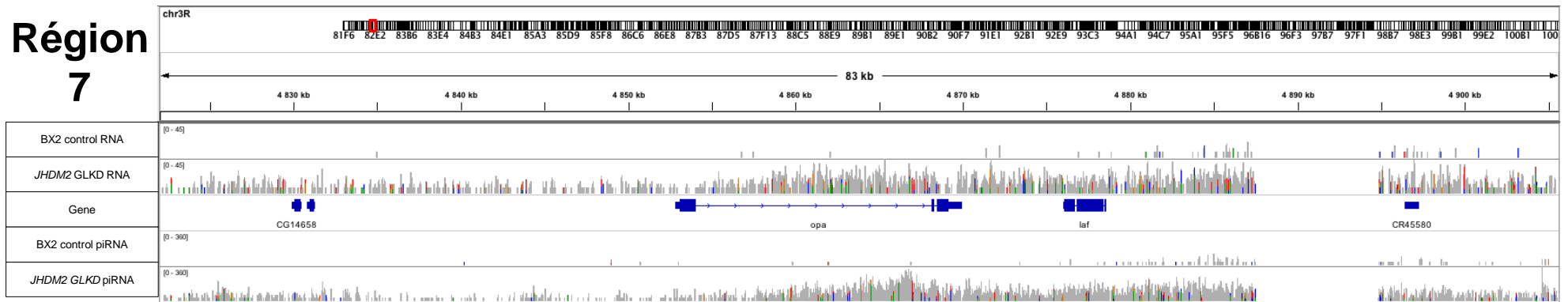
# Région 5



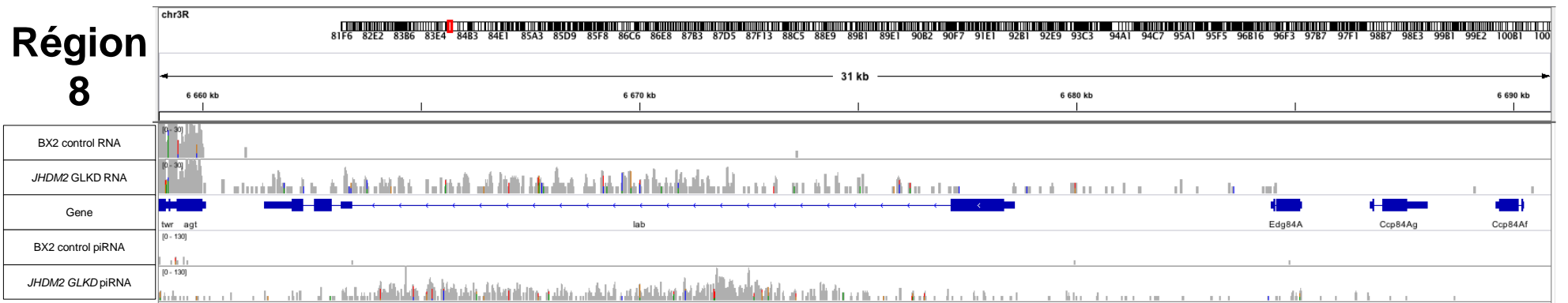
# Région 6



## Région 7



## Région 8



**Annexe 6: Identification de régions génomiques pour lesquelles le contexte *JHDM2 GLKD* induit une transcription et une production de piRNA aberrantes.** Les analyses de RNA-seq et de small RNA-seq ont permis d'identifier huit régions génomiques particulières: ces régions ne sont pas (ou très peu) transcrites et ne produisent pas (ou très peu) de piRNA dans le contrôle (*BX2* control RNA et *BX2* control piRNA) mais elles sont transcrites et produisent des piRNA en contexte *JHDM2 GLKD* (*JHDM2 GLKD* RNA et *JHDM2 GLKD* piRNA). Ces régions sont localisées sur les différents chromosomes et mesurent de 30 à 130 kb. On peut remarquer que ces régions contiennent des gènes (Gene) codants principalement des facteurs de transcription impliqués dans des processus développementaux. Les données ont été analysées avec la plateforme Galaxy et représentées grâce à IGV. Pour l'analyse des données de small RNA-seq, seuls les ARN de 23 à 29 nucléotides « unique mappers » sont gardés et alignés sur le génome de référence de *Drosophila melanogaster* (Flybase, release 6.26, 0 mismatch). Les données d'alignement des deux conditions sont comparées sur des fenêtres de 1kb et filtrées (ratio  $^{WT/JHDM2} > 2$  ou ratio  $^{WT/JHDM2} < 0,5$  avec un minimum de 5 piRNA par kb). Parmi les régions identifiées (un peu plus de 60 à cette étape), seules les régions présentant une distribution des piRNA homogène ont été conservées. Pour le RNA-seq, les données ont été alignées sur le génome de référence de *Drosophila melanogaster* (Flybase, release 6.26) pour les deux conditions.

Région n°...	Chromosome : Coordonnées	Gènes présent dans la région	Type de facteur codé	Fonction(s) biologique(s) connue(s)	Expression dans les ovaires en contexte sauvage
1	X: 4 372 000 – 4 509 000	<i>Bifid</i>	T-Box FT	Prolifération et migration cellulaire, développement (cerveau, œil, ailes)	Aucune
		<i>CG12688</i>	Inconnu	Inconnue	Aucune
2	X: 8 600 000 – 8 820 000	<i>Lim1</i>	Homeobox FT	Développement (œil, pattes)	Aucune
		<i>Oc</i>	Homeobox FT	Développement, organisation cellulaire	Aucune
3	X: 16 110 000 – 16 200 000	<i>Disco-r</i>	Zinc-finger FT	Développement (pattes, antennes, pièces buccales chez la larve et l'adulte)	Aucune
4	X: 17 752 000 – 17 820 000	<i>Unc-4</i>	Homeobox FT	Développement (système nerveux central)	Aucune
		<i>OdsH</i>	Homeobox FT	Reproduction (mâle)	Aucune
5	2L: 21 776 000 – 21 890 000	<i>Tsh</i>	Zinc-finger FT	Développement, prolifération cellulaire	Faible
		<i>CG31612</i>	Zinc-finger FT	Prolifération cellulaire	Faible
6	2R: 5 720 000 – 5 754 000	<i>Ap</i>	Homeobox FT	Développement (ailes, muscles), organisation cellulaire	Aucune
7	3R: 4 823 000 – 4 905 000	<i>Opa</i>	Zinc-finger FT	Développement (embryogenèse, intestin, tête, système nerveux central), rythme circadien	Aucune
		<i>Laf</i>	Inconnu	Inconnue	Aucune
		<i>CG14658</i>	Inconnu	Inconnue	Aucune
		<i>CG14659</i>	Inconnu	Inconnue	Aucune
8	3R: 6 660 000 – 6 690 000	<i>Labial</i>	Homeobox FT	Développement (tête, thorax)	Aucune

**Annexe 7: Informations complémentaires à l'identification de régions génomiques pour lesquelles le contexte *JHDM2 GLKD* induit une transcription et une production de piRNA aberrantes.** Pour chaque région, la localisation (incluant le chromosome et les coordonnées génomiques), les gènes présents ainsi que le type de facteur codé, la ou les fonction(s) biologique(s) connue(s) et son niveau d'expression dans les ovaires sont indiqués. Les informations données ici sont issues de la base de données FlyBase. Le numéro associé à chaque région est arbitraire, il correspond au numéro donné dans l'Annexe 6. FT : Facteur de Transcription.





# Table des figures

Figure 1 : Exemples de la proportion d'ET présents dans le génome de différents organismes.....	15
Figure 2 : Classification et structures des éléments transposables.....	16
Figure 3 : Les différents mécanismes de transposition.....	18
Figure 4 : Impact des ET sur la structure des génomes.....	20
Figure 5 : Séquences régulatrices des éléments transposables.....	21
Figure 6 : Effets de l'insertion des ET sur l'expression des gènes.....	22
Figure 7 : Rôle de <i>p53</i> dans le contrôle des rétrotransposons.....	24
Figure 8 : Rôle des ET dans la réponse au stress environnemental.....	27
Figure 9 : Schéma représentant la structure de l'élément <i>P</i> .....	28
Figure 10 : Phénomène de dysgénésie des hybrides.....	30
Figure 11 : Principales familles de petits ARN non-codants chez <i>Drosophila melanogaster</i> .....	33
Figure 12 : Lignée germinale de la femelle drosophile.....	35
Figure 13 : Transcription d'un piRNA cluster uni-strand – exemple du locus <i>flamenco</i> .....	37
Figure 14 : Signatures transcriptionnelles des piRNA clusters.....	39
Figure 15 : Transcription d'un piRNA cluster dual-strand – exemple de <i>42AB</i> .....	41
Figure 16 : Régions produisant le plus de piRNA dérivés de leur 3'UTR.....	42
Figure 17 : Rôle des piRNA dérivés du 3'UTR de <i>tj</i> dans la régulation de <i>FasIII</i> .....	43
Figure 18 : Export des ARN précurseurs vers le cytoplasme.....	46
Figure 19 : Structure des protéines PIWI.....	47
Figure 20 : Biogénèse des piRNA dans les ovaires de drosophile.....	48
Figure 21 : Protéine Piwi et piRNA associés à Piwi.....	50
Figure 22 : Régulation transcriptionnelle des ET.....	53
Figure 23 : Protéines Aub et Ago3 et piRNA associés à Aub et Ago3.....	54
Figure 24 : Mécanisme d'amplification Ping-Pong.....	56
Figure 25 : Structure, localisation et propriétés du transgène <i>P(lArB)</i> .....	65
Figure 26 : Structure, localisation et propriétés du cluster <i>BX2</i> .....	68

Figure 27 : Estimation des effets transgénérationnels, quelle génération ?.....	72
Figure 28 : Mesure des quantités relatives de transcrits par RT-qPCR dans les ovaires de différents acteurs moléculaires à 10°C (en bleu) et à 25°C (en vert).....	160
Figure 29 : La perte de fonction de <i>hfp</i> entraîne l'activation de <i>BX2<sup>OFF</sup></i> pour la production de piRNA.....	168
Figure 30 : La conversion du cluster <i>BX2<sup>OFF</sup></i> en contexte <i>hfp GLKD</i> est stable et ne requiert pas la présence du transgène cible <i>P(TARGET)</i> .....	170
Figure 31 : La perte de fonction de <i>JHDM2</i> entraîne l'activation de <i>BX2<sup>OFF</sup></i> pour la production de piRNA.....	172
Figure 32 : La conversion de <i>BX2<sup>OFF</sup></i> dépend de l'épissage de <i>JHDM2</i> par <i>hfp</i> .....	174
Figure 33 : Quantité de transcrits relative de <i>hfp</i> et <i>JHDM2</i> à 25°C et à 29°C.....	176
Figure 34 : Quantité relative de transcrits <i>BX2</i> et <i>hfp</i> dans les contextes <i>hfp GLKD</i> et <i>JHDM2 GLKD</i> .....	177
Figure 35 : Effet des contextes <i>hfp GLKD</i> et <i>JHDM2 GLKD</i> sur l'état épigénétique de <i>P(lArB)</i> .....	179
Figure 36 : Identification de régions génomiques pour lesquelles le contexte <i>JHDM2 GLKD</i> induit une transcription et une production de piRNA aberrantes.....	180
Figure 37 : <i>moon</i> est impliqué dans la transcription de <i>labial</i> , <i>opa</i> et <i>Oc</i> , localisés dans trois des huit régions identifiées.....	181
Figure 38 : Les descendants des <i>JHDM2 GLKD</i> ne dépassent pas le stade L1.....	182
Figure 39 : Phénotypes des descendants des <i>JHDM2 GLKD</i> chez les L1-like.....	184
Figure 40 : Modèle permettant d'expliquer les phénotypes des L1-like descendants des <i>JHDM2 GLKD</i> .....	186
Figure 41 : Modèle sur le rôle de <i>JHDM2</i> .....	258
Figure 42 : Structures des transgènes « cibles » (aussi appelés <i>P(TARGET)</i> ).....	274

# Table des tables

Table 1 : Pourcentages de répression de sous-lignées $BX2^{OFF} P(TARGET)$ et $BX2^{ON} P(TARGET)$ à 18°C au cours des générations.....	157
Table 2 : Pourcentages de répression de sous-lignées $BX2^{OFF} P(TARGET)$ et $BX2^{ON} P(TARGET)$ après un traitement de 8 semaines à 10°C (=Génération 0) et de leur descendance, développée à 25°C (=Génération 1).....	159
Table 3 : Effet d'une exposition au NaCl au cours du développement.....	162
Table 4 : Effet d'une exposition au CuSO <sub>4</sub> au cours du développement.....	163
Table 5 : Effet d'une exposition au bisphénol A (BPA) au cours du développement.....	164
Table 6 : Répartition de la répression dans les contextes <i>hfp GLKD</i> et <i>JHDM2 GLKD</i> .....	177



## **Mécanismes d'activation d'un piRNA cluster par des facteurs environnementaux et génétiques chez *Drosophila melanogaster***

Les éléments transposables (ET) sont des séquences d'ADN mobiles dont l'activité peut être délétère pour l'intégrité du génome. Dans la lignée germinale, l'activité des ET est régulée par les PIWI-Interacting RNA (piRNA), des petits ARN non-codants produits par des séquences spécifiques appelées piRNA clusters. Pour étudier la biologie de ces piRNA clusters, nous utilisons une construction transgénique capable de mimer un piRNA cluster : le cluster *BX2*. Le laboratoire a montré que *BX2* pouvait être activé à chaque génération pour la production de piRNA grâce à l'héritage de piRNA maternels homologues, décrivant ainsi le premier cas de paramutation stable chez un animal. Ce résultat pose la question de l'initiation de l'activation d'un piRNA cluster en absence d'héritage maternel de piRNA. Au cours de ma thèse, j'ai étudié le rôle de deux facteurs indépendants que nous avons identifiés comme étant capables d'activer *BX2* en l'absence d'héritage de piRNA maternels : l'augmentation de la température au cours du développement (29°C au lieu de 25°C) et la perte de fonction de la déméthylase d'histone *JHDM2*. J'ai montré que l'augmentation de la température entraînait la conversion de *BX2* dans 2% des cas et nécessite la présence d'une séquence homologue transcrite dans le génome, suggérant un mécanisme d'activation ARN-dépendant. J'ai également montré que la perte de fonction de *JHDM2*, en plus d'entraîner l'activation de *BX2*, est responsable de la conversion d'autres régions génomiques pour la production de piRNA. La transcription de ces régions dépend de la machinerie de transcription des piRNA clusters suggérant que *JHDM2* est impliqué dans la détermination des piRNA clusters.

Mots-clés : [Eléments transposables ; PIWI-Interacting RNA ; piRNA clusters ; hétérochromatine ; épigénétique transgénérationnelle ; drosophile]

## **piRNA cluster activation mechanisms by environmental and genetic factor in *Drosophila melanogaster***

Transposable elements (TE) are mobile DNA sequences that can be deleterious for genome integrity. In germ cells, TE activity is regulated by PIWI-Interacting RNAs (piRNAs), small non-coding RNA which are produced from specific sequences called piRNA clusters. To study piRNA cluster biology, we use a transgenic construct that mimics a piRNA cluster: the *BX2* cluster. In the lab, we have previously demonstrated that *BX2* can be activated at each generation for piRNA production by maternal piRNA inheritance, describing the first stable paramutation case in animal. Hence, is it possible to activate *BX2* without maternal piRNA inheritance? During my PhD, I focused on the role of two independent factors that we identified to be capable to activate *BX2* without maternal piRNA inheritance: the increase of developmental temperature (29°C instead of 25°C) and the loss-of-function of the histone demethylase *JHDM2*. I have shown that the temperature increase induces *BX2* conversion in 2% of the cases and that this conversion requires the presence of transcribed homologous sequences in the genome, suggesting an RNA-dependent mechanism. I have also discovered that *JHDM2* loss-of-function, besides *BX2* conversion, drives piRNA production from other genomic regions. The transcription of these regions depending on piRNA cluster transcription machinery, suggesting that *JHDM2* is implicated in piRNA cluster determination.

Keywords : [Transposable elements ; PIWI-Interacting RNA ; piRNA clusters ; heterochromatin ; transgenerational epigenetics ; drosophila]