



Emergence d'un locus producteur de piRNAs chez la drosophile : mise en place de l'épigénome

Catherine Hermant

► To cite this version:

Catherine Hermant. Emergence d'un locus producteur de piRNAs chez la drosophile : mise en place de l'épigénome. Génétique animale. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066097 . tel-01264981

HAL Id: tel-01264981

<https://theses.hal.science/tel-01264981>

Submitted on 30 Jan 2016

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.

Université Pierre et Marie Curie

ED 515 : Complexité du Vivant

IBPS - Laboratoire de biologie du développement - UMR7622 - CNRS

Université Pierre et Marie Curie

Equipe : Répression épigénétique et ADN mobile

Emergence d'un locus producteur de piRNAs chez la drosophile

Mise en place de l'épigénome

Par Catherine Hermant

Thèse de doctorat de Génétique

Dirigée par Stéphane Ronsseray

Présentée et soutenue publiquement le 28 janvier 2015

Devant un jury composé de :

Dr. Hervé Seitz	CR1 CNRS	Rapporteur
Dr. André Verdel	CR1 CNRS	Rapporteur
Dr. Jean René Huynh	DR2 CNRS	Examineur
Dr. Vincent Colot	DR1 CNRS	Examineur
Pr. Emmanuèle Mouchel-Vielh	Professeur	Examineur
Dr. Emilie Brassat	Maitre de Conférences	Examineur
Dr. Stéphane Ronsseray	DR2 CNRS	Directeur de thèse

Emergence d'un locus producteur de piRNAs chez la drosophile

Les éléments transposables d'ADN sont presque ubiquitaires dans le monde vivant et leur mobilité peut être délétère pour le génome. Leur régulation dans les tissus germinaux animaux passe par la voie de silencing des piRNAs (*PIWI-interacting RNAs*). Les piRNAs sont produits à partir de loci contenant des fragments d'éléments transposables insérés en clusters. Nous étudions l'émergence de ces clusters de piRNAs chez la drosophile.

Nous avons activé *de novo* un cluster de transgènes par héritage maternel de piRNAs homologues. Il s'agit d'un cas de paramutation, ou conversion épigénétique stable et récurrente. Nous avons montré que ce cluster paramuté produit *de novo* des piRNAs, et étonnement des siRNAs.

J'ai caractérisé de façon fonctionnelle et moléculaire ce phénomène de paramutation par l'utilisation de mutants. J'ai montré que les propriétés de silencing, ainsi que la production de piRNAs et de siRNAs, sont abolies en contexte mutant pour tous les gènes testés de la voie des piRNAs (voies primaire et secondaire). Parallèlement, j'ai étudié un cas de paramutation « partiellement homologue » dans laquelle le cluster reçoit des piRNAs homologues seulement à une partie de sa séquence. J'ai montré qu'il y a production de piRNAs par la totalité du cluster dès la 3^e génération.

J'ai montré, enfin, que des clusters activés *de novo* par la chaleur, présentent des propriétés fonctionnelles et moléculaires semblables aux clusters activés par les piRNAs maternels.

Ces travaux apportent des éléments clés pour la compréhension de la mise en place de l'épigénome, tant d'un point de vue mécanistique qu'évolutif.

Mots clés : *PIWI-interacting RNAs*, paramutation, épigénétique, éléments transposables, *Drosophila*

Emergence of a piRNA-producing locus in Drosophila

DNA transposable elements are almost ubiquitous in the living world and their mobility can be deleterious for the genome. Their regulation in germaria is mediated by the piRNA silencing pathway (*PIWI-interacting RNAs*). piRNAs are produced by loci formed by clusters of fragments of transposable elements. We are studying the emergence of these piRNA-producing clusters in *Drosophila*.

We have *de novo* activated a cluster of transgenes *via* maternal inheritance of homologous piRNAs. This is a case of paramutation i.e. a stable and recurrent epigenetic conversion process. We have shown that this paramutated cluster produces *de novo* piRNAs and, surprisingly, also siRNAs.

I have characterized this paramutation functionally and molecularly, by a mutant approach. I have shown that its silencing properties, as well as piRNA and siRNA production are abolished in mutant contexts for all the genes from the primary and secondary piRNA pathways I have tested. At the same time, I have studied the case of a partially homologous paramutation, in which piRNAs maternally inherited by the cluster are homologous to only a part of its sequence. I have shown that piRNAs are produced all along the cluster as early as the 3rd generation.

Finally, I have shown that a cluster activated *de novo* by an environmental stress shows the same functional and molecular properties as a cluster paramutated *via* maternal piRNA inheritance.

These studies provide key elements for understanding the emergence of the epigenome from a mechanistic and an evolutionary perspective.

Key words: *PIWI-interacting RNAs*, paramutation, epigenetics, transposable elements, *Drosophila*

Remerciements

Tout d'abord, je tiens à remercier les membres du jury d'avoir accepté d'évaluer mon travail de thèse. Un grand merci à Hervé et André qui ont accepté d'être Rapporteurs et qui ont lu mon manuscrit, relativement conséquent, en pleine période de Fêtes. J'ai beaucoup apprécié vos remarques et suggestions concernant mon travail qui annoncent une discussion passionnante.

Je tiens à remercier l'ensemble de l'équipe pour m'avoir aussi bien accueillie et soutenue durant ces dernières années. Vous avez été pour moi comme une seconde famille et c'est avec beaucoup de bonheur que j'ai partagé ces moments.

Je te remercie Stéphane pour tout ce que tu as fait pour moi. Je ne pense pas pouvoir en faire une liste exhaustive mais je suis très reconnaissante de tout ce que tu m'as apporté et te remercie d'avoir fait de moi une bonne scientifique.

Je te remercie Val, pour tout. Tu as été pour moi beaucoup plus qu'une collègue, tu es une amie, une confidente et même une seconde mère. Je ne sais pas comment je vais pouvoir continuer sans t'avoir à mes côtés. Tu ne veux pas partir avec moi ?

Un grand merci à toi Amna pour ton amitié et l'aide que tu m'apportes chaque jour, que ce soit pour galaxy ou même mes problèmes de PC... Nous étions de très bonnes amies mais tu es devenue une collègue géniale et je vais regretter de ne plus travailler avec toi. Mais on peut toujours y remédier dans l'avenir en montant notre boîte ou une équipe. J'espère que l'on va continuer à se voir encore très longtemps.

Merci à toi Laure pour m'avoir guidée et conseillée depuis que je suis entrée dans l'équipe. Ça a été un grand plaisir de travailler avec toi et d'écouter tes conseils.

Je tiens à remercier l'ensemble de l'équipe Péronnet : Frédérique, Camille, Valérie, Delphine, Emmanuèle, Jean Michel, Neel, Hélène, Jérôme, Sébastien et Anne, car vous avez tous à un moment été d'un grand secours pour moi tant sur le plan professionnel que personnel. Je voudrais remercier plus particulièrement Frédérique. Merci de m'avoir soutenue et encouragée durant ma thèse. Tu as été une très bonne Tata. Un grand merci à Valérie et Delphine pour toutes ces conversations, votre soutien et vos conseils. Un grand merci à Camille qui est devenue au fil du temps une très bonne amie et confidente. Même si on habite très loin l'une de l'autre, j'espère qu'on continuera à se voir. Merci Emmanuèle de me faire l'honneur de participer à mon jury de thèse. Merci à Jérôme pour toutes ces conversations d'un autre temps !! Je te souhaite de faire une belle carrière. Merci Hélène pour tous ces petits pots de verre mais pas que.

A notre petit club du midi, mille mercis !!! Le japonais n'est pas très loin et j'espère que nous continuerons à faire nos petits repas de temps en temps.

Un grand merci à Hafida, Isa, Hari pour l'aide et la bonne humeur qu'elles m'ont apportées. En plus d'une efficacité remarquable, vous êtes toujours accueillantes et joviales. C'est un plaisir de venir vous voir dans votre bureau et de partager quelques minutes avec vous.

Je tiens à remercier ma famille pour son soutien et de m'avoir supportée tout au long de cette aventure. Merci à toi Yann qui a été là pour moi à chaque seconde dans mes moments de joie, de stress, de doute qui sont récurrents chez moi. Tu es pour moi une source de vie et ma reconnaissance est sans limite. Tu m'as offert le plus beau des cadeaux, ma vie, ma joie, ma petite Colline qui n'a pas conscience du bonheur et de l'aide qu'elle m'a apportée durant ma thèse. C'est en partie grâce à vous deux si j'ai réussi à aller jusqu'au bout de ce travail. Cette thèse n'est pas ma thèse mais la nôtre.

Merci Maman et Adrien pour votre soutien indéfectible. J'espère vous avoir près de moi encore des années pour continuer à partager avec vous tous ces moments. Merci Papa d'être présent. Tu sais que je suis toujours là pour toi. Merci à toi ma sœur. Nous sommes toujours présentes l'une pour l'autre. Merci pour ton soutien.

Merci à tous ceux que je n'ai pas cité (UMR 7622, ED CDV, financement,...) et qui ont contribué à ma réussite. Vous avez toute ma reconnaissance.

Merci aux mouches pour leur sacrifice et qui m'ont offert un splendide sujet.

Merci aux bisounours qui me rappellent chaque jour que nous ne vivons pas « dans un monde de bisounours ». Chaque contrariété que j'ai pu rencontrer m'a touché mais n'a fait que me rendre plus forte.

Et enfin, je souhaite un bon courage et bonne continuation à l'équipe. Je vous souhaite de continuer à faire de la belle génétique et des papiers de plus en plus exceptionnels. Il s'agit d'une petite équipe mais avec un grand cœur. Ne l'oubliez pas s'il vous plaît !!

Que Dieu vous garde tous !!!

Cat la Param

*« Ils ne savaient pas que c'était
impossible, alors ils l'ont fait »*

De Mark Twain.

Sommaire

Préambule.....	1
Introduction générale	3
I/ Les éléments transposables	3
1/ La classification des éléments transposables	3
A/ Les ETs de classe I.....	4
B/ Les ETs de classe II.....	5
II/ Régulations des éléments transposables par les petits ARNs non-codants.....	7
1/ Régulations germinales par les piRNAs : petits ARNs non-codants de la lignée germinale chez la drosophile.....	7
A/ Présentation de la lignée germinale de la drosophile.....	7
a/ L'appareil génital de la drosophile femelle.....	7
b/ L'unité structurelle de développement pour l'ovocyte : le follicule ovarien.....	8
c/ Le développement de la chambre ovarienne.....	8
B/ Biogenèse des piRNAs germinaux.....	11
a/ Les clusters de piRNAs et la production des transcrits précurseurs de piRNAs.....	13
b/ La biogenèse des piRNAs primaires.....	24
c/ La biogenèse des piRNAs secondaires.....	28
C/ Répression des ETs euchromatiques par les piRNAs germinaux.....	31
a/ Régulation Post-transcriptionnelle : PTGS.....	31
b/ Régulation transcriptionnelle : TGS.....	32
c/ Maintien des clusters de piRNAs.....	37
2/ Régulations somatiques par les petits ARNs non-codants chez la drosophile...40	40
A/ Biogenèse des piRNAs somatiques folliculaires.....	40
a/ Les clusters de piRNAs et la production des transcrits précurseurs de piRNAs.....	41
b/ Biogenèse des piRNAs primaires somatiques.....	42
B/ Mécanismes de répression folliculaires des éléments Transposables.....	43

C/ Biogenèse des siRNAs : petits ARNs non codants dans les tissus somatiques.....	46
D/ Mécanisme de répression somatiques des éléments Transposables.....	49
E/ siRNAs produit par les clusters de piRNAs.....	50
F/ piRNAs like dans les tissus somatiques.....	50
 III/ Clusters de piRNAs transgéniques : outils pour l'étude de l'émergence de loci producteurs de piRNAs.....	51
1/ Mécanisme de répression de transgènes par les piRNAs : Le <i>Trans-Silencing Effect</i> (TSE).....	53
2/ Propriétés génétiques et moléculaires du TSE.....	54
 IV/ Phénomène de paramutation lié à l'effet maternel du TSE.....	58
1/ Définition de la paramutation.....	58
2/ La paramutation au locus <i>Booster B1</i> chez le maïs.....	59
3/ La paramutation au locus <i>Kit</i> chez la souris.....	62
 V/ Présentation du sujet.....	63
 Résultats.....	65
 VI/ Emergence d'un locus producteur de piRNAs par un phénomène de paramutation.....	67
1/ Présentation des lignées <i>T-1</i> et <i>BX2</i>	67
2/ Article: « Paramutation in <i>Drosophila</i> linked to emergence of piRNA-producing locus » (de Vanssay <i>et al.</i> 2012).....	68
3/ <i>Extra-View</i> : « piRNAs and epigenetic conversion in <i>Drosophila</i> . ».....	93

VII/ Etudes des propriétés fonctionnelles et moléculaires d'un locus producteur de piRNAs paramuté.....	102
1/ « Paramutation in <i>Drosophila</i> involves the primary and secondary piRNA pathways and can induce a rapid spreading within a locus for piRNA production capacity ».....	103
 VIII/ Emergence d'un locus producteur de piRNAs par un phénomène environnemental.....	158
1/ Activation du locus <i>BX2</i> par la chaleur.....	158
2/ Propriétés fonctionnelles et moléculaires d'un locus activé par la chaleur.....	159
 Discussion.....	165
 Conclusion.....	179
 Références Bibliographiques.....	181

Introduction générale

Préambule

Les éléments transposables (ETs) sont presque ubiquitaires dans le monde vivant. Le séquençage massif des génomes nous montre qu'il existe très peu d'espèces dépourvues d'éléments transposables (Pritham *et al.* 2009). La notion d'élément transposable a été proposée pour la première fois par Barbara Mac Clintock dans son ouvrage intitulé « *The Origin and Behavior of Mutable loci in maize* » publié le 8 avril 1950. Dans cet ouvrage, elle démontre que des séquences d'ADN sont capables de « transposer », c'est-à-dire de se déplacer d'un endroit à un autre dans le génome. Cette découverte a été très largement rejetée par la communauté scientifique de l'époque pour qui les chromosomes étaient considérés comme des structures stables. Ce ne sera que des années plus tard que l'importance de ses travaux sera reconnue et elle se verra décerner le prix Nobel de Médecine en 1983 pour la découverte des éléments transposables.

Ces séquences d'ADN mobiles codent pour la plupart pour des gènes catalysant leur mobilité. Cela en fait des séquences à haut effet mutagène, voir même délétère pour la cellule lorsqu'elles se déplacent dans le génome. Toutefois de nombreux cas de domestication d'éléments transposables ont été identifiés, comme par exemple celui des éléments transposables *hetT-A*, *TART* et *TAHRE* qui constituent les télomères de la drosophile et protègent l'extrémité des chromosomes de la dégradation (Pride *et al.* 1997 ; Mason *et al.* 2008). Néanmoins, la cellule a su développer des mécanismes de régulation de ces éléments transposables afin de préserver l'intégrité de son génome, aussi bien dans les tissus somatiques que germinaux afin de préparer au mieux sa future descendance. Il existe de nombreux mécanismes différents de répression des éléments transposables selon les espèces. Il ne s'agira pas ici de faire un catalogue exhaustif de toutes les voies permettant de réguler les éléments transposables. Notons toutefois que ces mécanismes peuvent être épigénétiques et requièrent souvent l'utilisation d'ARNs non-codants et des marques chromatiniennes spécifiques. L'épigénétique désigne les mécanismes qui permettent la transmission de changements dans l'expression des gènes qui sont héréditaires à travers la mitose et/ou la méiose, alors que l'agent ayant induit le changement a disparu, et qui ne résultent pas de modifications de la séquence d'ADN.

De plus en plus de cas d'interactions entre les ARNs non-codants et les marques chromatiniennes ont été relatés dans la littérature. Prenons l'exemple du RITS pour RNA-inducing Transcriptional Silencing chez *S. pombe*, où de petits ARNs non-codants issus de transcrits double-brins reconnaissent les transcrits naissants homologues du génome et induisent le recrutement de protéines chromatiniennes sur le locus transcrit (Verdel *et al.* 2004). Aussi je m'attacherai à vous présenter ici les différents mécanismes de silencing des éléments transposables par les petits ARNs non-codants chez la drosophile, plus particulièrement au sein de la lignée germinale. En effet, il est primordial pour les cellules de la lignée germinale de réguler la transposition de ces éléments car ces cellules sont à l'origine de la future descendance. Même si lors de la méiose, puis de la mitose, un certain degré de réarrangements est supportable afin de permettre au génome d'évoluer, l'intégrité du génome est indispensable à la survie de l'espèce. Cette régulation est réalisée à partir de loci contenant des copies dégénérées d'éléments transposables insérées en clusters, qui permettront la production de petits ARNs non-codants appelés piRNAs (Piwi-interacting RNAs) qui sont à l'origine de la répression en *trans* des copies insérées dans l'euchromatine. Simultanément, il a été montré la contribution essentielle de protéines chromatiniennes qui interviennent dans la répression des copies euchromatiques, mais également dans le maintien des capacités répressives des clusters de piRNAs chez la drosophile.

Introduction générale

I/ Les éléments transposables

Les éléments transposables (ETs) sont des séquences d'ADN qui sont capables de se déplacer de manière autonome dans un génome, par un mécanisme appelé transposition. Ils sont présents dans presque tous les génomes en nombre variable (Tableau 1). Ces séquences d'ADN mobiles constituent une part importante de ce qu'on appelle les séquences répétées.

espèce	% ET
<i>Zea mays</i>	80
<i>Arabidopsis thaliana</i>	10
<i>Homo sapiens</i>	45
<i>Mus musculus</i>	40
<i>Drosophila simulans</i>	5
<i>Drosophila melanogaster</i>	15
<i>Caenorhabditis elegans</i>	12
<i>Saccharomyces cerevisiae</i>	3-5

Tableau 1 : Proportion en éléments transposables des génomes eucaryotes de référence. Les proportions en ETs dans les articles de séquençage de référence : *Z. mays* (Schnable *et al.* 2009), *A. thaliana* (Arabidopsis Genome Initiative, 2000), *C. elegans* (*C. elegans* Sequencing Consortium, 1998), *H. sapiens* (Lander *et al.* 2001), *D. melanogaster* et *D. simulans* (Drosophila 12 genome consortium *et al.* 2007), *S. cerevisiae* (the Yeast genome directory 1997), *M. musculus* (Mouse genome sequencing consortium *et al.* 2002)).

1/ La classification des éléments transposables

Les ETs peuvent être classés selon plusieurs critères. Une première classification a été proposée par Finnegan en 1989 à partir de la nature de leur intermédiaire de transposition (Figure 1).

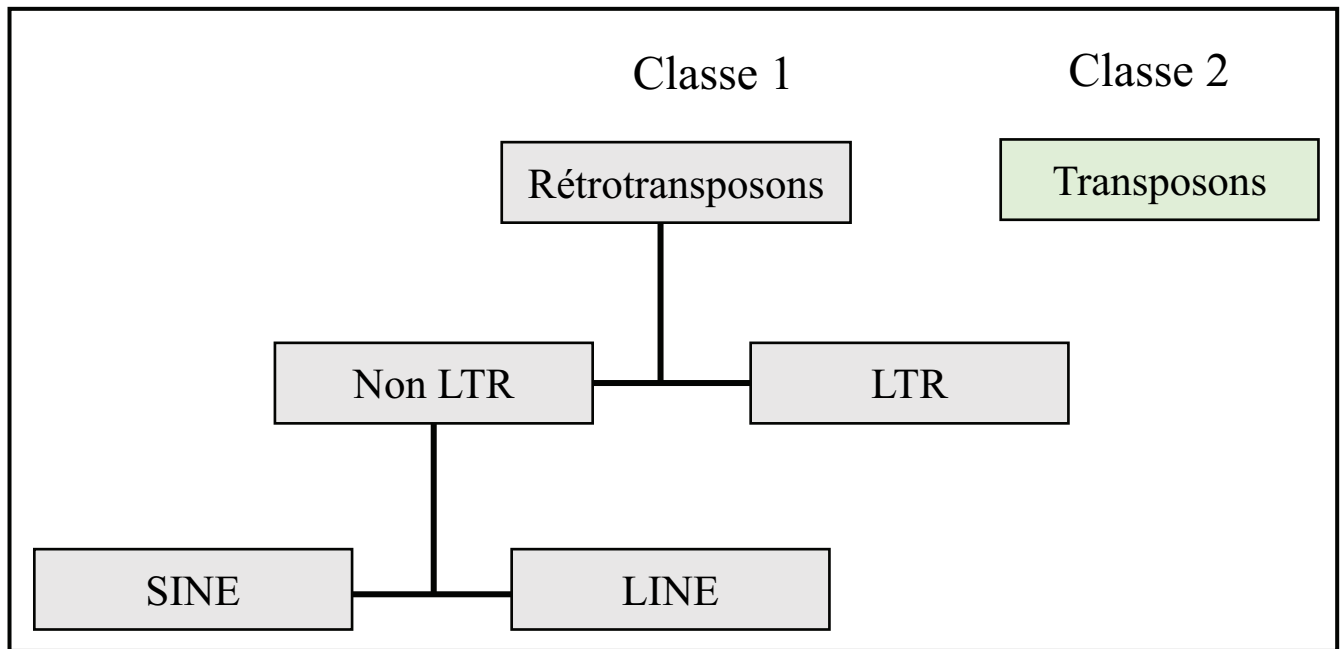


Figure 1 : Classification des ETs eucaryotes par Finnegan (1989).

A/ Les ETs de classe I

Les ETs de la classe I ou rétrotransposons (Feschotte et Pritham, 2007) utilisent une molécule d'ARN comme intermédiaire de transposition. Ces éléments sont d'abord transcrits en ARN, puis retrotranscrits *via* une Transcriptase Inverse (ou RT pour « *Reverse Transcriptase* ») en une molécule d'ADN qui sera ensuite intégrée à un nouveau site du génome par une Intégrase (**Figure 2**). La transposition se fait ainsi par un mode de transposition en « copier-coller ». Les rétrotransposons sont eux-mêmes divisés en deux sous-classes selon qu'ils disposent ou non à leurs extrémités de longues répétitions terminales (LTR, pour « *Long Terminal Repeat* »). Les rétrotransposons à LTR ont une structure très proche de celle des rétrovirus et comportent les régions nécessaires à leur transposition ainsi que des séquences régulatrices situées dans les séquences LTR (**Figure 2b**). Les rétrotransposons sans LTR sont eux-mêmes divisés en deux familles : les « *Long Interspersed Nuclear Elements* » (LINEs) qui codent leurs propres RT et sont donc autonomes et les « *Short Interspersed Nuclear Elements* » (SINEs) qui sont non-autonomes et utilisent la machinerie des LINEs pour pouvoir transposer (**Figure 2c**).

B/ Les ETs de classe II

Les ETs de la classe II ou transposons à ADN (Feschotte et Pritham, 2007) utilisent une molécule d'ADN comme intermédiaire de transposition (**Figure 2a**). Dans ce cas, la transposition consiste en l'excision d'une copie de l'ET puis en son insertion ailleurs dans le génome. Ce mode de transposition, appelé « couper-coller », est non réplcatif puisqu'il s'agit de la séquence de l'ET elle-même qui se déplace. Les transposons codent le plus souvent une Transposase qui est entourée des deux côtés par des répétitions terminales inversées (TIR, pour « *Terminal Inverted repeats* »).

C'est cette classification qui est la plus utilisée. Néanmoins, le développement des techniques de séquençage a permis d'accumuler une grande quantité de données génomiques et cette méthode de classification ne permet pas de positionner toutes les familles d'ETs.

Une classification, qui prend en compte l'enzymologie et la similarité de structure ainsi que les relations entre les séquences, a permis de mettre en évidence des nouvelles familles d'ETs tels que les MITEs pour « *Miniature Inverted repeat Transposable Elements* » qui transposent par un mécanisme de « copier-coller ». Les séquences consensus de la plupart des familles d'ETs eucaryotes décrites sont présentes dans une seule base de données appelée RepBase (Jurka et al. 2005) (<http://girinst.org/repbase/>) créée en 1990.

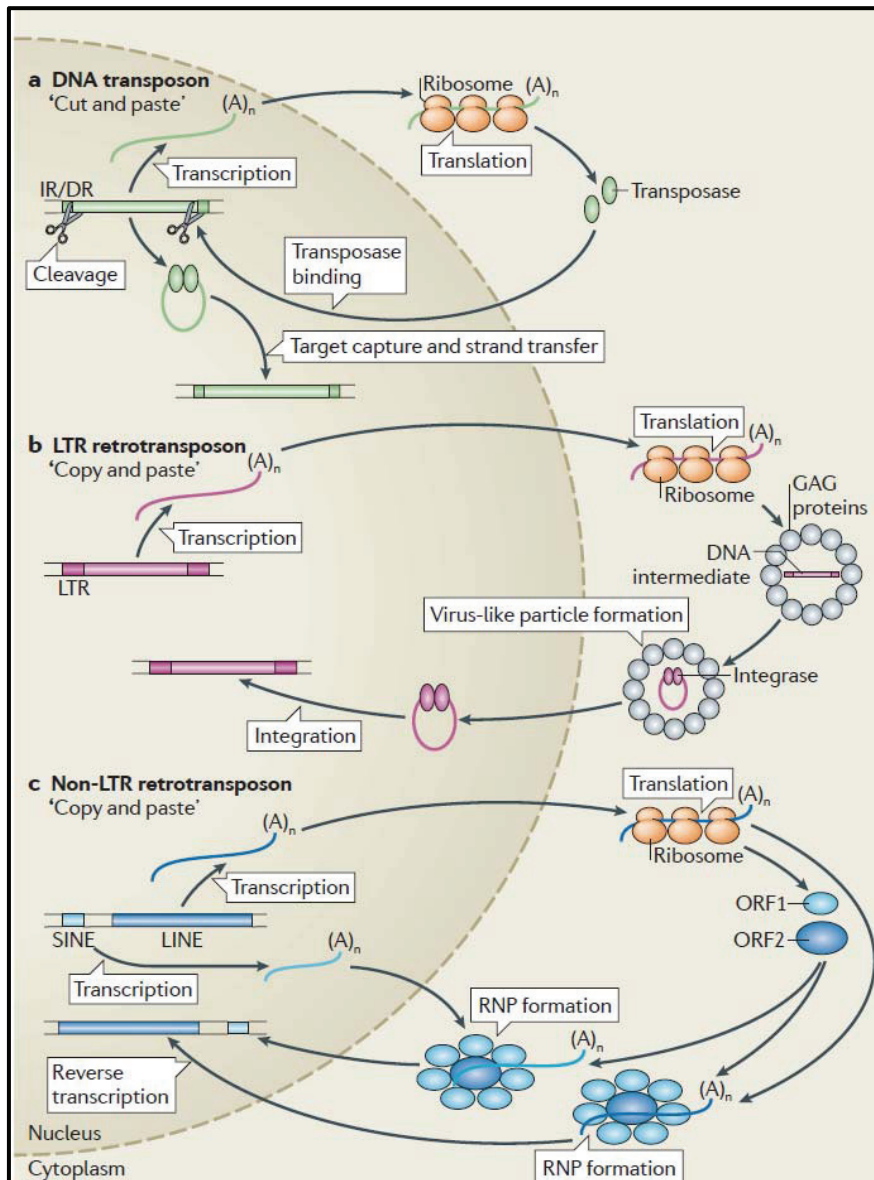


Figure 2 : Mécanismes de transposition des différentes classes d'éléments transposables.

(a) les transposons à ADN ou classe II transposent selon un mode « couper-coller » en vert. Ce type de transposon code pour sa transposase qui va être capable de cliver le transposon *via* les séquences IR et insérer cette séquence ailleurs dans le génome. (b) les rétrotransposons à LTR ou classe I transposent selon un mode « copier-coller » en rose. Les rétrotransposons produisent un transcrit qui va être traduit en différentes protéines : GAG, Intégrase. Les protéines GAG forment une particule virale qui va contenir le transcrit du rétrotransposon à LTR. Ce transcrit va être rétrotranscrit et pris en charge par une Intégrase qui permet l'insertion de ce rétrotransposon à nouveau site dans le génome. (c) les rétrotransposons non-LTR ou classe I transposent selon un mode « copier-coller » en bleu. Les rétrotransposons non-LTR : LINEs et SINEs vont être transcrits. Seuls les transcrits issus des LINEs vont être traduits en protéines ORF1 et ORF2. Les protéines ORF1 et ORF2 ont toutes les deux un domaine de reverse transcriptase. Ces deux protéines vont former un complexe RNP (pour « ribonucleoprotein ») qui va charger les transcrits issus des LINEs et SINEs pour les ramener dans le noyau et leur fonction reverse-transcriptase leur permet d'intégrer ces rétrotransposons à un nouveau site dans le génome. (Siomi *et al.* 2011)

Si la classification des ETs est complexe, la façon de les réguler l'est également. Chaque espèce a su développer des mécanismes de régulation, ceux-ci varient d'une espèce à l'autre et même pour une même espèce d'un tissu à l'autre. Nous verrons ici le cas de la régulation des ETs chez *Drosophila melanogaster*, qui est le modèle d'étude auquel je me suis intéressée durant ma thèse.

II/ Régulations des éléments transposables par les petits ARNs non-codants

1/ Régulations germinales par les piRNAs : petits ARNs non-codants de la lignée germinale chez la drosophile.

A/ Présentation de la lignée germinale de la drosophile

a/ L'appareil génital de la drosophile femelle

Une femelle drosophile possède une paire d'ovaires connectée au niveau de l'oviducte qui permet la connexion ensuite avec l'utérus (**Figure 3**). L'appareil génital de la drosophile femelle comprend une spermathèque qui permet de conserver les spermatozoïdes pour féconder les œufs (King, 1970). Chaque ovaire est composé d'une quinzaine d'ovarioles de type méroistique polytrophique, c'est-à-dire que le trophocyte constitué de cellules nourricières est enfermé avec l'ovocyte dans le follicule. Une ovariole est une chaîne linéaire et polarisée qui contient 6 à 7 follicules ovariens en cours de développement à des degrés de maturation croissante du pôle antérieur vers le pôle postérieur (**Figure 4**). Ces 6 ou 7 follicules ovariens sont séparés par des cellules inter-folliculaires.

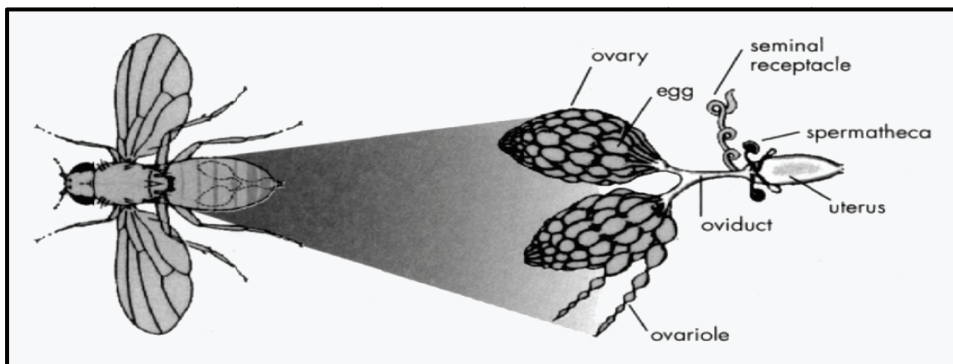


Figure 3 : Représentation de l'appareil génital de la drosophile femelle.

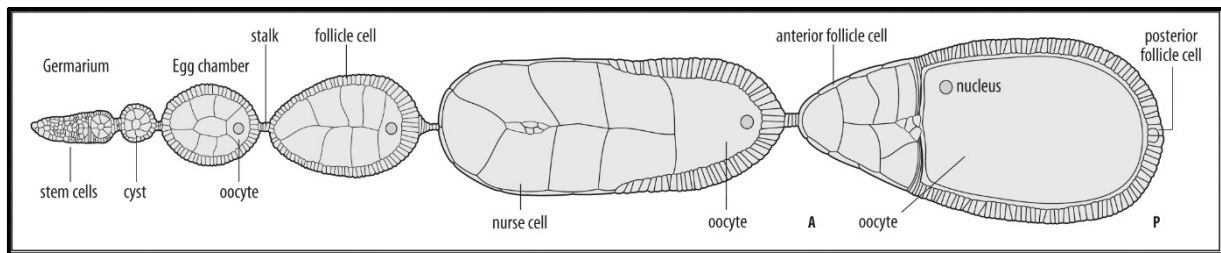


Figure 4 : Représentation schématisque d'une ovariole.

L'ovariole est composée de plusieurs follicules en cours de développement selon un axe antéro-postérieur où les stades les plus précoces sont dans la partie antérieure et les stades les plus tardifs dans la partie postérieure. Le stade le plus précoce est appelé germarium et contient les cellules souches permettant la production du cyste. Le cyste est constitué de 16 cellules germinales au centre qui sont entourées par les cellules somatiques. Au cours du développement, l'une des cellules germinales positionnées au pôle postérieur de la chambre ovarienne va devenir l'ovocyte et les 15 autres seront appelées cellules nourricières (ou « *nurse cells* »). Les cellules somatiques vont se différencier en cellules folliculaires (ou « *follicle cells* ») et en cellules inter-folliculaires (« *stalk* »). Dans la partie la plus postérieure, il y aura donc les follicules plus matures dans lesquels l'ovocyte est de taille plus importante. (Diapositive08.JPG)

b/ L'unité structurelle de développement pour l'ovocyte : le follicule ovarien

L'ovocyte se développe au sein du follicule ovarien, encore appelé chambre ovarienne (**Figure 4**). Celui-ci est constitué de deux populations distinctes de cellules. Les cellules germinales, qui sont les plus internes du follicule, comprennent 16 cellules qui sont connectées entre elles par des ponts cytoplasmiques qui forment un syncytium (ou cyste). 15 de ces cellules sont les cellules nourricières polypléïdes. La 16^e cellule, l'ovocyte, est quant à elle bloquée en métaphase de la première division de méiose et reprendra le cycle lors de la fécondation. Le deuxième type cellulaire correspond aux cellules folliculaires somatiques qui forment une monocouche de cellules épithéliales autour du cyste.

c/ Le développement de la chambre ovarienne

- Formation du premier follicule ovarien

La production des cellules somatiques et germinales a lieu tout au long de la vie des femelles et ceci grâce à l'activité continue de deux types de cellules souches présentes dans le germarium, qui est la partie la plus antérieure de l'ovariole (**Figure 5**) : les cellules souches germinales (GSC pour « *germ stem cells* ») et les cellules souches somatiques (SSC pour

« somatic stem cells »). Le germarium est divisé morphologiquement en 4 régions appelées 1, 2a, 2b et 3 de la plus antérieure à la plus postérieure. C'est dans la région 1 que nous trouvons les GSC. Celles-ci sont au nombre de 2 à 3 et se divisent de façon asymétrique, de manière à former une cellule fille, appelée cystoblaste, qui va s'engager dans la voie de différenciation germinale, et une cellule souche, ce qui permet d'entretenir le système (revue : Spradling *et al.* 2001 ; Gonzalez-Reyes, 2003). Dans cette partie du germarium les cellules germinales sont entourées par des cellules somatiques appelées cellules du manteau intérieur (ou « *inner-sheat* »). Ces cellules forment un manchon qui entoure les régions 1 et 2a. Dans la partie la plus antérieure de cette région se trouvent les cellules du filament terminal, puis viennent les cellules de la coiffe (ou « *cap cells* ») qui sont en contact direct avec les GSC.

Détaché de la niche permettant le maintien à l'état cellule souche, constituée par les cellules de la coiffe et le filament terminal, le cystoblaste va alors subir quatre mitoses synchrones, mais à cytokynèse incomplète. Cela aura pour effet de former, dans la région 2a, un cyste germinale de seize cellules interconnectées entre elles par des ponts cytoplasmiques (Deng and Lin, 2001). Les cellules constituant ce cyste sont appelées cystocytes.

A la frontière des régions 2a et 2b se trouvent les SSC (**Figure 5**). Elles sont au nombre de deux ou trois et se divisent également de façon asymétrique pour régénérer des SSC et donner une cellule fille qui va alors rentrer en différenciation pour donner les cellules pré-folliculaires qui entourent le cyste et permettent de l'individualiser. Cela donnera lieu, dans la région 3, à la formation du premier follicule ovarien. C'est dans cette région qu'il va y avoir migration de l'ovocyte au pôle postérieur du cyste et ce jusqu'au stade 10. L'expulsion hors du germarium (ou bourgeonnement) de la chambre de stade 1 est conditionnée par la différenciation des cellules pré-folliculaires en cellules polaires et inter-folliculaires d'une part, et en cellules folliculaires épithéliales d'autre part (Zhang and Kalderon, 2001).

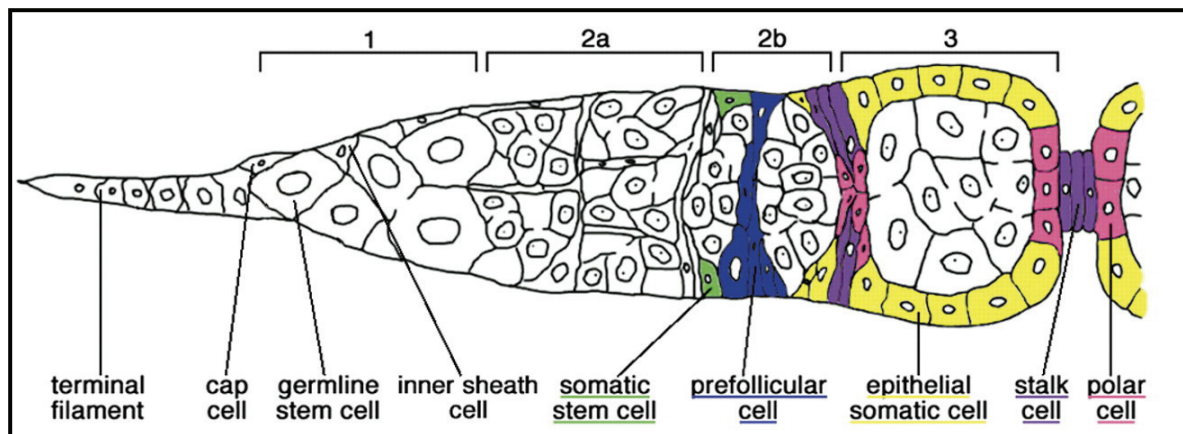


Figure 5 : Représentation schématisique du germarium.

Le germarium est le stade le plus précoce et contient les cellules souches nécessaires à la formation du follicule ovarien. Il a été compartimenté en quatre régions morphologiques distinctes de la plus précoce à la plus tardive respectivement appelées régions 1, 2a, 2b et 3.

La partie la plus antérieure est composée de cellules somatiques composant le filament terminal, puis viennent les cellules de la coiffe (ou « *cap cell* »). Au contact des cellules de la coiffe et des cellules du manteau interne (ou « *innersheat cell* ») se trouvent les cellules souches germinales (GSC). Dans la région 2b se trouvent les cellules souches somatiques (SSC) qui donneront après division les cellules pré-folliculaires (ou « *prefollicular cell* »). Elles seront différenciées ensuite en cellules folliculaires (ou « *epithelial somatic cell* »), cellules inter-folliculaires (ou « *stalk cell* ») et cellules polaires (ou « *polar cell* »). (Zhang and Kalderon 2001)

- Les différents stades du follicule ovarien :

Le développement des follicules a été divisé par King en 14 stades distincts sur des critères morphologiques (King, 1970). Le stade 1 correspond au follicule ovarien prêt à être expulsé du germarium, alors que le stade 14 correspond à l'ovocyte mature prêt à être fécondé. La production d'un follicule de stade 1 à partir des cellules souches germinales prend environ quatre jours à 25°C. A partir du stade 1, il faut deux jours de plus pour arriver au stade 7 et un jour de plus pour arriver au stade 14.

On distingue les stades pré-vitellogéniques : du stade 1 à 7 au cours desquels toutes les cellules germinales (cellules nourricières et ovocyte) ont à peu près la même taille et les stades vitellogéniques à partir du stade 8, au cours desquels la taille de l'ovocyte va augmenter considérablement par rapport aux cellules nourricières. Cette augmentation de volume est due à l'accumulation d'ARNm et protéines du stade 8 à 11. A partir du stade 11, l'augmentation de volume est fortement accentuée par le phénomène de « *dumping* ». Il s'agit d'un phénomène au cours duquel il y a le déversement du contenu des cellules nourricières dans l'ovocyte. Les cellules nourricières vont ensuite mourir par apoptose et l'ovocyte occupera l'intégralité du follicule.

Un ovocyte mature va donc être généré avec une polarisation antéro-postérieure et dorso-ventrale, en fonction de l'accumulation de protéines et d'ARNm spécifiques ainsi que de la présence de structures morphologiques spécifiques (Ephrussi and Lehmann 1992 ; Lehmann and Ephrussi 1994).

B/ Biogenèse des piRNAs germinaux

La régulation des ETs « germinaux » se fait dans les cellules nourricières et l'ovocyte qui contiennent un cytoplasme syncytial commun (Figure 4). Cette régulation se fait par une classe de petits ARNs récemment découverte, appelés piRNAs (pour « Piwi-interacting RNAs »). Les piRNAs sont générés entre autres par les protéines de la famille PIWI (Piwi Aub et AGO3) (**Figure 6**) et sont des petits ARNs de 23-28nt (Brennecke et al. 2007).

Cette régulation passe par la production de transcrits précurseurs de piRNAs issus d'une transcription atypique des clusters de piRNAs, régions génomiques riches en séquences d'ETs dégénérées (Senti and Brennecke 2010 ; Mohn *et al.* 2014 ; Zhang *et al.* 2014). Ces précurseurs de piRNAs seront ensuite clivés en piRNAs primaires et pris en charge par les protéines Piwi et Aub. Il a été observé une amplification des piRNAs dits secondaires dans un mécanisme appelé cycle Ping-Pong entre les protéines Aub et AGO3. Les piRNAs chargés par Piwi vont être capables de rentrer dans le noyau pour permettre d'une part la répression des ETs actifs et d'autre part l'entretien de clusters de piRNAs. Nous verrons ici plus en détails chacune de ces étapes ainsi que les diverses protéines qui interviennent à chaque étape.

a/ Les clusters de piRNAs et la production des transcrits précurseurs de piRNAs

La cartographie des piRNAs sur le génome montre qu'ils proviennent de deux types de régions génomiques. Les principales sont des régions spécifiques appelées clusters de piRNAs, qui contiennent de nombreux fragments d'ETs délétés et dégénérés. Ces régions sont définies par leur sens de transcription et leurs protéines associées. Il existe deux types de clusters de piRNAs : les clusters dits bidirectionnels et les clusters dits unidirectionnels (**Figure 7**). Les clusters de piRNAs sont préférentiellement localisés dans des régions hétérochromatiques péricentromériques et subtélomériques des chromosomes. Ils contiennent de nombreux fragments d'ETs et peuvent être considérés comme un catalogue des séquences d'ETs qu'il est nécessaire de réguler. Ceci signifie qu'en l'absence d'un fragment d'un ET donné dans l'un des clusters, celui-ci peut échapper à la répression (Khurana *et al.* 2011). Le second type de régions produisant des piRNAs provient de gènes codants (Brennecke *et al.* 2007 ; Robine *et al.* 2009 ; Saito *et al.* 2009) et de copies euchromatiques d'ETs. L'argument le plus direct provient de l'élément *I* de la famille des LINEs. En effet, l'élément *I* est présent dans le génome au niveau du cluster de piRNAs, ainsi que sous forme de copies euchromatiques actives. Les séquences de ces différentes copies de l'élément *I* sont très divergentes, ce qui a permis l'identification de piRNAs spécifiques des copies actives de l'élément *I* (Brennecke *et al.* 2008).

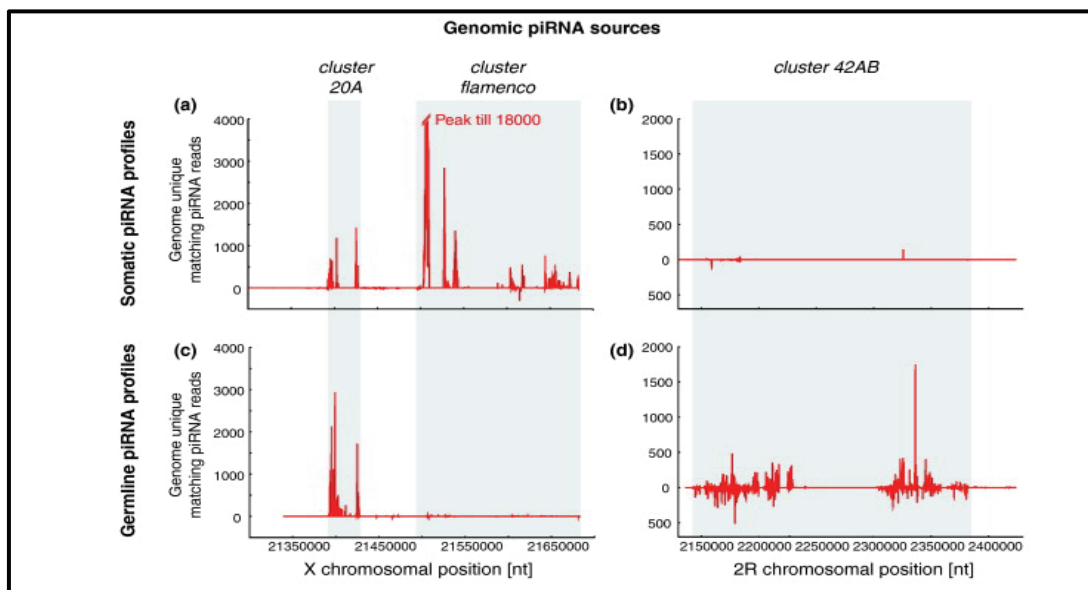


Figure 7 : Profil de piRNAs des 2 principaux clusters unidirectionnels : cluster 20A et *flamenco* et cluster bidirectionnel majeur (cluster 42AB).

Il s'agit du profil de piRNAs (23-30nt) issus du cluster 20A (cluster 2), cluster *flamenco* et du cluster 42AB représentés en gris. Pour faire une comparaison des profils, les banques ont été normalisées à 1 million de petits ARNs de 23-30nt issus de séquences répétées. Les banques des piRNAs somatiques ont été faites sur les cellules OSS qui sont des cellules en culture dérivées de cellules somatiques folliculaires (a et b) alors que les banques des piRNAs germinaux ont été faites sur des embryons très jeunes possédant un cytoplasme d'origine germinale (c et d). (a) Ce graphique montre que les clusters 20A et *flamenco* produisent des piRNAs uniquement sens dans les cellules somatiques alors que le cluster 42AB (b) ne produit pas de piRNAs. (c) Ce graphique montre que dans les embryons précoces, il y a production de piRNAs par le cluster 20A et pas par le cluster *flamenco*. Donc *flamenco* est un cluster spécifique des cellules folliculaires somatiques alors que le cluster 20A s'exprime dans les deux tissus et produit toujours des piRNAs uniquement sens. Par contre, le cluster 42AB produit des piRNAs sens et antisens uniquement dans les tissus germinaux. Cela montre que ce cluster est spécifique de la lignée germinale. (Senti and Brennecke 2010)

- Les clusters bidirectionnels germinaux:

La majorité des clusters dans la lignée germinale est bidirectionnelle, ce qui signifie qu'ils sont transcrits dans les deux sens (Figure 7 et 8). Le cluster principal est le cluster 42AB qui contient entre autres une séquence altérée de l'élément *I* (Brennecke *et al.* 2007). Il a été montré par des expériences de ChIP-seq que les clusters bidirectionnels sont des régions hétérochromatiques riches en marque H3K9me3 et pauvre en marque H3K4me2 (Mohn *et al.* 2014) (Figure 8). La marque H3K9me3 est établie par la protéine Eggless qui est une histone méthyltransférase (Rangan *et al.* 2011). Il a été montré, toujours par ChIP seq, que la protéine Rhino (protéine homologue de HP1 qui est une protéine chromatienne) est présente

spécifiquement sur les clusters bidirectionnels (Vermaek *et al.* 2005 ; Klattenhoff *et al.* 2009 ; Mohn *et al.* 2014 ; Zhang *et al.* 2014) (**Figure 8**).

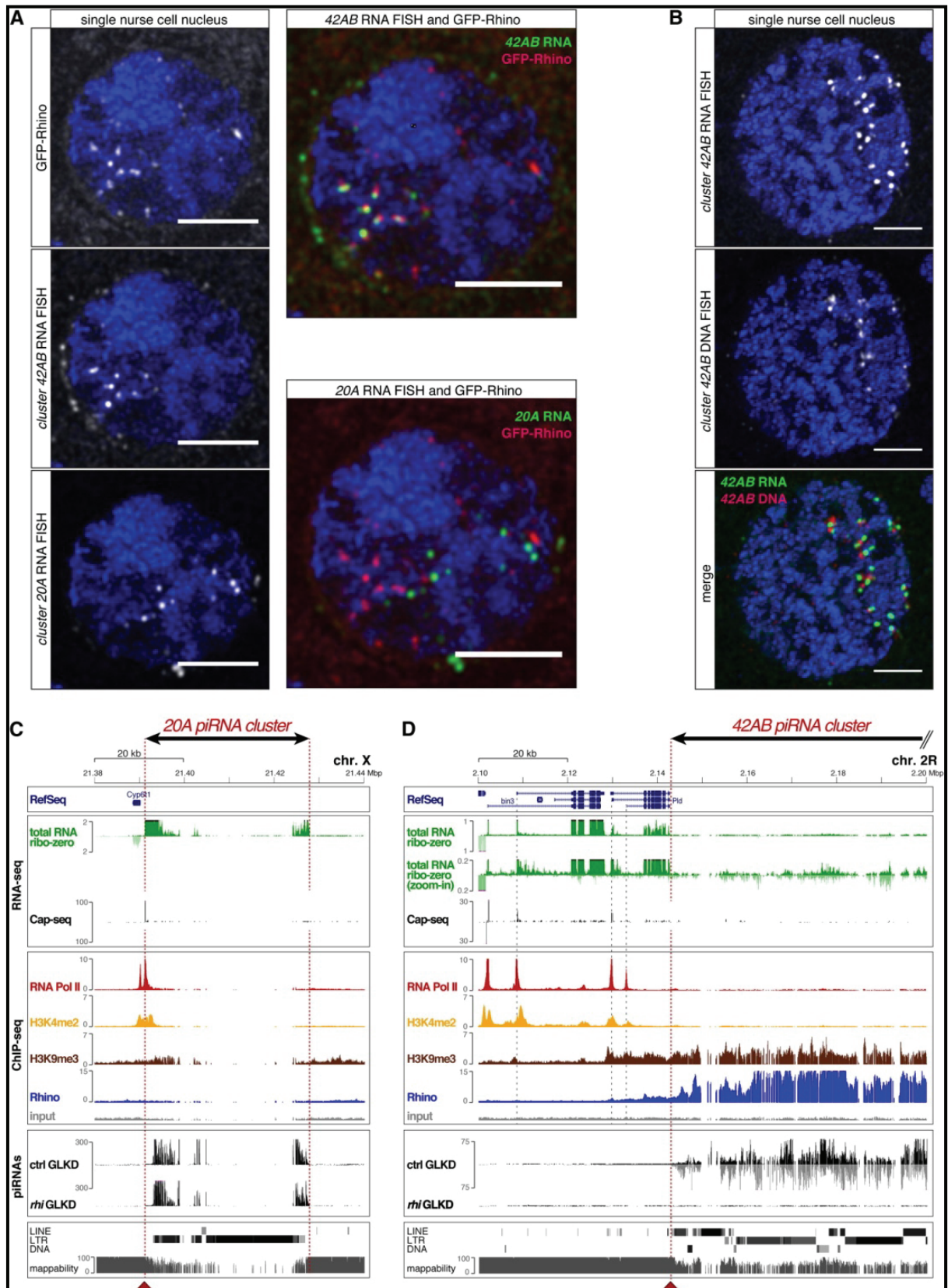


Figure 8 : Signatures transcriptionnelles des clusters bidirectionnels et unidirectionnels.

(A) Expérience de RNA Fish (hybridation *in situ* fluorescente) en vert des clusters 42AB bidirectionnel et 20A unidirectionnel (encore appelé cluster 2) en colocalisation avec la protéine de fusion Rhino-GFP en rouge. Rhino colocalise avec les ARNs du cluster 42AB mais pas avec ceux du cluster 20A. (B) Expériences de DNA (en rouge) et RNA (en vert) Fish du cluster 42AB. Les ARNs du cluster 42AB colocalisent avec le cluster 42AB. (C-D) Expériences de RNA-seq (total RNA et cap-seq), ChIP-seq de l'ARN polII, H3K4me2, H3K9me3 et Rhino ainsi que séquençage à haut débit des piRNAs en contexte sauvage et mutant pour *rhino* pour les clusters 20A (C) et 42AB (D). En (C) analyse concernant le cluster 20A qui est le cluster unidirectionnel de référence, localisé sur le chromosome X situé à proximité du gène *Cyp6t*. Le RNA-seq nous montre qu'il y a production uniquement du transcrit sens et le Cap-seq qui permet de visualiser la localisation de la coiffe des transcrits naissants, nous montre que ce cluster produit des transcrits sens coiffés. Les analyses de ChIP-seq des protéines ARN polII, H3K4me2, H3K9me3 et Rhino nous montrent un enrichissement de l'ARN pol II et H3K4me2 au début du cluster 20A. Concernant la marque H3K9me3, il y a un enrichissement de part et d'autre des bornes du cluster. Il n'y a pas d'enrichissement de Rhino. Le profil de piRNAs montre la présence de piRNAs sens uniquement en contexte sauvage et ce profil est peu affecté en contexte mutant pour *rhino*. Le dernier module nous montre que le cluster 20A est composé essentiellement d'ETs à LTR et la présence de quelques LINEs. En (D) il s'agit de la même analyse faite sur le cluster 42AB qui est le cluster bidirectionnel de référence, localisé sur le bras droit du chromosome 2. Ce cluster est localisé à proximité du gène *bin3* et *Pld*. Le RNA-seq nous montre que le cluster 42AB est effectivement transcrit dans les deux sens mais de façon faible. Nous n'observons pas d'enrichissement de la coiffe au niveau du cluster 42AB contrairement au gène adjacent. Cela signifie que les transcrits issus du cluster 42AB ne sont pas coiffés. Les ChIP-seq de l'ARN polII et H3K4me2 ne présentent pas non plus d'enrichissement au niveau du cluster 42AB contrairement aux gènes adjacents. Par contre, il y un fort enrichissement en marque H3K9me3 et en Rhino. L'analyse des piRNAs montre la présence de piRNAs sens et antisens en contexte sauvage qui sont perdus en contexte mutant pour *rhino*. Nous constatons également que le cluster 42AB est composé essentiellement de LINEs et ETs à LTR avec la présence de quelques transposons, alors que les régions adjacentes en contiennent très peu. L'ensemble de ces résultats montre que les clusters 42AB bidirectionnel et 20A unidirectionnel ont des signatures transcriptionnelles différentes. (Mohn *et al.* 2014)

La protéine Rhino est une protéine hétérochromatique composée de trois domaines : un domaine CD (pour « *chromo domain* ») en N terminal qui lui permet des interactions avec la chromatine (elle se lie aux marques H3K9me2/3), un domaine hinge au centre et un domaine CSD (pour « *chromo-shadow domain* ») en C terminal. Bien qu'il y ait transcription de ces clusters, le ChIP seq de l'ARN polII ne révèle pas un fort enrichissement de l'ARN PolII sur les clusters mais on trouve un fort enrichissement sur les promoteurs adjacents au cluster (Mohn *et al.* 2014) (**Figure 8**). Les travaux de Mohn montrent la présence d'un complexe de protéines spécifiques aux clusters bidirectionnels appelé le RDC complexe. En effet, il a été montré la présence sur la chromatine de la protéine Rhino qui reconnaît les marques H3K9me2/3 par le biais de son domaine CD mais également que la protéine Rhino est elle-même liée au niveau du domaine CSD à la protéine Deadlock au niveau de son domaine N terminal (Klattenhoff *et al.* 2009 ; Mohn *et al.* 2014). La protéine Deadlock a été

identifiée comme impliquée dans la biogenèse de piRNAs lors d'un crible par RNAi réalisé dans le but d'identifier de nouveaux facteurs impliqués dans la biogenèse de piRNAs germinaux (Czech *et al.* 2013). Il a été montré que le knockdown de *deadlock* diminue la quantité de piRNAs sens et antisens des clusters bidirectionnels mais pas des clusters unidirectionnels. Cela montre que cette protéine a une action spécifique aux clusters bidirectionnels. De plus, il a été montré que cette protéine colocalise avec la protéine Rhino (Czech *et al.* 2013 ; Mohn *et al.* 2014). La protéine Deadlock est composée d'un domaine N terminal et C terminal. La protéine Deadlock interagit avec la protéine Cutoff. Cutoff est homologue à la protéine Rai1 chez la levure, la protéine Dom3Z chez les métazoaires et DXO chez la souris. Ces protéines sont connues pour avoir un rôle dans le contrôle qualité des extrémités 5' des transcrits naissants (ARNm) en se liant aux ARNm et en les dégradant (Jiao *et al.* 2010 ; Jiao *et al.* 2013) ainsi que dans la terminaison de la transcription (Kim *et al.* 2004). Néanmoins l'activité exonucléase 5' vers 3' de Rai1/DXO n'est pas conservée dans Cutoff (Chen *et al.* 2007 ; Pane *et al.* 2011). Les expériences d'immunofluorescence montrent que ces trois protéines sont présentes spécifiquement dans les noyaux des cellules germinales, forment des foci et colocalisent toutes les unes avec les autres (Pane *et al.* 2011, Mohn *et al.* 2014).

Le complexe RDC est donc spécifique des clusters bidirectionnels et pourrait bloquer la terminaison de la transcription des gènes adjacents aux clusters de piRNAs ce qui permettrait à l'ARN polII de transcrire ces clusters (**Figure 9**) (Mohn *et al.* 2014). Une seconde hypothèse, avancée par les travaux conjoints de Mohn et Zhang, est que ce complexe pourrait également favoriser l'absence d'épissage des transcrits précurseurs de piRNAs (**Figure 9**). De plus, la protéine Cutoff pourrait interagir avec l'extrémité 5' des transcrits précurseurs de piRNAs et empêcher leur dégradation par les protéines Rat1/Xrn2 et/ou bloquer le recrutement à l'extrémité 5' du complexe de coiffe (CBC pour « *cap-binding complexe* »). Tout ceci conduirait à une transcription non-canonique des clusters de piRNAs (**Figure 9**).

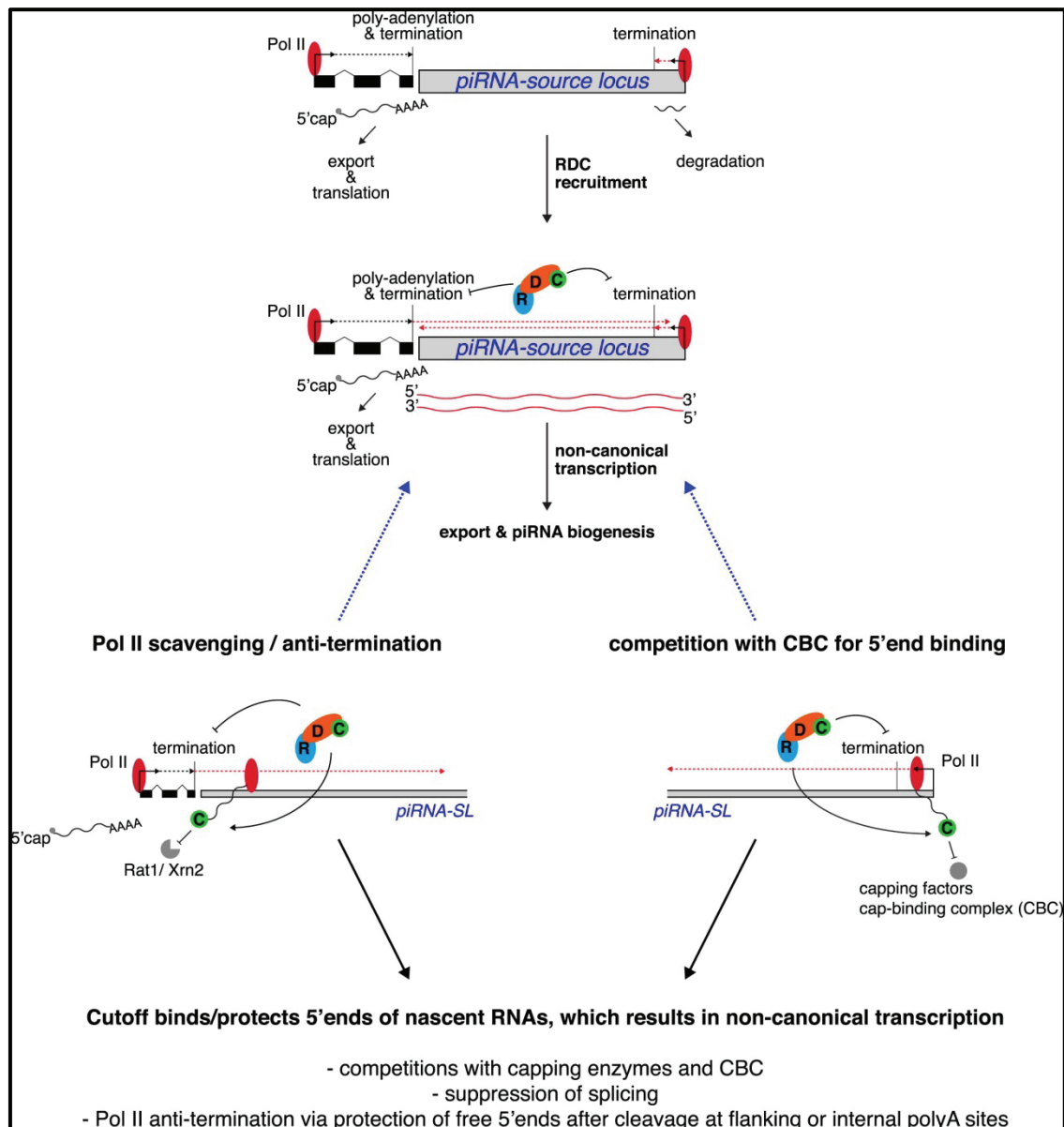


Figure 9: Modèle de transcription non canonique des clusters bidirectionnels.

Le complexe RDC permettrait de bloquer la terminaison de la transcription des gènes adjacents aux clusters de piRNAs ce qui permettrait une transcription par l'*ARN polIII*. De plus, la protéine Cutoff pourrait se lier aux transcrits naissants et bloquer la dégradation de ces transcrits par Rat1 et Xrn2 et/ou bloquer le recrutement du complexe de coiffe (CBC). (Mohn *et al.* 2014)

La recherche de nouvelles protéines impliquées dans la voie des piRNAs a permis d'identifier la protéine UAP56 (Zhang *et al.* 2012 et Zhang *et al.* 2014). UAP56 (Hel25E) est une protéine possédant un domaine DEAD box qui permet la fixation des ARNs. Elle est impliquée dans le remodelage de la chromatine (Eberl *et al.* 1997), dans l'export des ARNm hors du noyau (Gatfield *et al.* 2001) ainsi que dans la régulation l'épissage alternatif *via* la spliceosome (Park *et al.* 2004 ; Katzenberger *et al.* 2009). De plus, des femelles hétéro-alléliques pour *UAP56* sont stériles et montrent des défauts dans la localisation asymétrique

des ARNm (Eberl *et al.* 1997 ; Meignin and Davis 2008) tout comme certaines protéines de la voie des piRNAs. Des expériences d'immunofluorescence montrent qu'UAP56 est nucléaire et colocalise avec les foci de Rhino (Zhang *et al.* 2014). En perte de fonction pour *rhino*, une réduction drastique de la présence d'UAP56 est observée dans le noyau. De plus, la perte de fonction pour *UAP56* affecte la localisation des autres protéines de la voie des piRNAs dans le nuage, à l'exception de Piwi, ainsi qu'une baisse des piRNAs germinaux et plus particulièrement ceux issus des clusters bidirectionnels. Le nuage est une région cytoplasmique qui entoure le noyau d'apparence plus dense et qui contient les protéines de la voie des piRNAs (Eddy EM, 1974 et 1975 ; Lim *et al.* 2007 ; Olivieri *et al.* 2010). L'ensemble de ces résultats montre qu'UAP56 est impliquée dans la biogenèse de piRNAs germinaux et plus particulièrement des précurseurs de piRNAs issus de clusters bidirectionnels. L'effet de la perte de fonction d'*UAP56* sur la localisation des ARNm suggère que cette protéine pourrait intervenir dans le transport des précurseurs de piRNAs pour les amener dans le cytoplasme, et plus particulièrement dans le nuage (Pane *et al.* 2007). Mais il reste encore de nombreuses questions quant à la spécification de précurseur de piRNAs. Il a longtemps été proposé qu'une protéine particulière pourrait rendre ces transcrits spécifiques et reconnus comme précurseur de piRNAs.

Il a été montré qu'il y a des défauts d'épissage de précurseurs de piRNAs qui sont dépendants des protéines Rhino, Cutoff et UAP56 (Zhang *et al.* 2014 ; **Figure 10**). Des expériences de RT q-PCR spécifiques des sites d'épissage montrent que les précurseurs de piRNAs ne sont pas épissés contrairement aux transcrits des gènes codants. Ils décrivent un très bel exemple du locus *sox102F* qui est composé de séquences uniques et produit quatre transcrits distincts dans les cellules somatiques (**Figure 10**). Dans les cellules germinales, par contre, ce même locus produit des piRNAs. En contexte mutant pour *rhino* et *cutoff*, il y a non seulement perte de la production de piRNAs mais également une augmentation de transcrits au niveau des exons en adéquation avec le profil d'un gène codant. Des expériences de RT q-PCR sur les sites d'épissage montrent qu'en ovaire, il y a plus de transcrits de *sox102F* non épissés qu'épissés par rapport aux carcasses (tissus somatiques de référence). Par contre en mutant *rhino* et *cutoff*, il y a plus de transcrits de *sox102F* épissés.

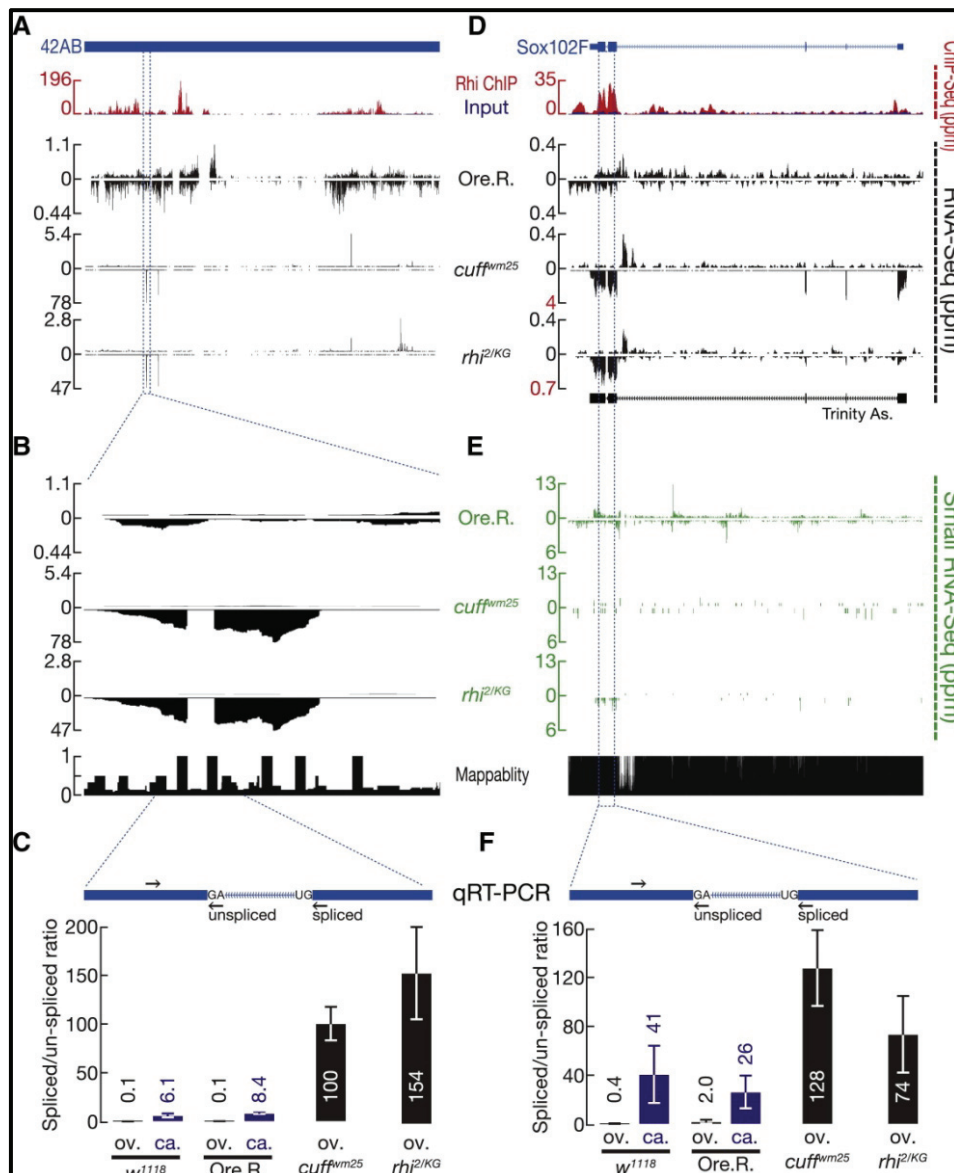


Figure 10 : Rhino et Cutoff suppriment l'épissage des clusters 42AB et *sox102F*.

(A) et (D) montrent les profils de ChIP-seq de Rhino et de RNA-seq en contexte sauvage, mutant pour *cutoff* et *rhino* pour les clusters 42AB et *sox102F*. Rhino est présent sur le cluster 42AB et sur le cluster *sox102F* au niveau des régions transcrites. Le profil RNA seq, en contexte sauvage, montre qu'il y a des transcrits sens et antisens qui sont perdus sauf pour certaines petites régions pour le cluster 42AB en contexte mutant pour *cutoff* et *rhino*. Par contre, ils sont réduits sur le cluster *sox102F* sauf au niveau des introns de ce gène inséré en sens inverse dans génome. Il y a une augmentation des transcrits antisens au niveau des introns de ce cluster. (B) Il s'agit d'un zoom du cluster 42AB sur une région contenant un intron. En contexte sauvage, il y a une transcription faible mais quasi-continue de la région. Alors qu'en contexte mutant pour *cutoff* et *rhino*, il y a une plus forte transcription de cette région mais avec un trou qui correspond à la présence de l'intron. En (C) et (F), il s'agit d'une analyse par RT q-PCR de la région d'épissage de l'intron révélé en (B) sur 42AB et d'une région de *sox102F* contenant un intron. Le ratio transcrits épissés/non épissés est calculé. En (C) concernant le cluster 42AB, il y a plus de transcrits non-épissés dans les deux fonds génétiques sauvages en ovaire par rapport à la carcasse. Par contre, il y a plus de transcrits épissés dans les mutants *cutoff* et *rhino*. En (F) concernant le cluster *sox102F*, on retrouve les mêmes résultats qu'en (C) pour le cluster 42AB. (E) montre les profils des petits ARNs

respectivement du cluster *sox102F* en contexte sauvage et mutant pour *cutoff* et *rhino*. Il y a production de petits ARNs sens et antisens tout le long du cluster en contexte sauvage. En contexte mutant pour *cutoff* et *rhino*, il y a réduction des petits ARNs sens et anti-sens mais il y a présence de quelques petits ARNs antisens spécifiquement au niveau de la jonction de deux exons. (Zhang *et al.* 2014)

Un modèle est alors proposé par cette équipe qui permettrait d'expliquer comment les précurseurs de piRNAs peuvent être distingués des autres transcrits. En effet, l'une des grandes énigmes de cette voie de piRNAs est de savoir comment les précurseurs de piRNAs sont différenciés des autres et conduits dans le nuage pour y être clivés en piRNAs. Ce modèle propose que la présence de Rhino sur les clusters de piRNAs permette le recrutement de Cutoff et donc la formation du complexe RDC. Cutoff pourrait interagir avec l'extrémité des transcrits naissants et à la place du complexe CBC. Ensuite, il y aurait recrutement de UAP56 et l'absence du complexe CBC qui aurait pour conséquence le non-épissage des transcrits. Ces transcrits chargés par Cutoff et UAP56 seraient reconnus comme précurseurs de piRNAs (**Figure 11**).

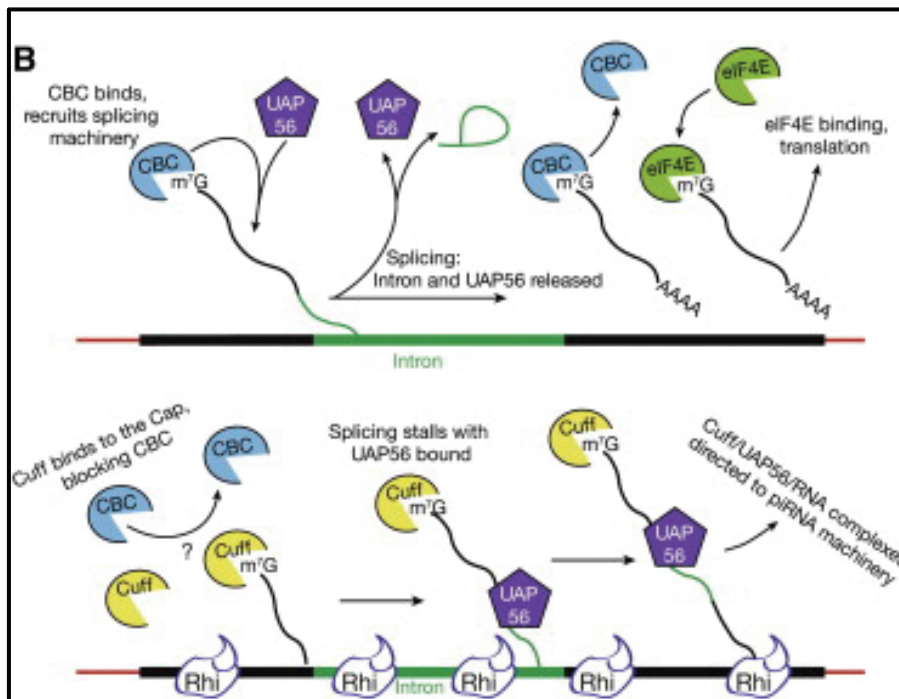


Figure 11: Modèle du rôle du Rhino, Cutoff et UAP56 dans la biogenèse des piRNAs.

(En haut) Séquence d'ADN contenant un intron. Cette région est transcrite et génère un transcrit sur lequel il va y avoir recrutement du complexe CBC au niveau de l'extrémité du transcrit naissant. Cela aura pour conséquence le recrutement d'UAP56 qui permettra l'épissage de l'intron, puis il y aura un relargage d'UAP56. Ensuite, il y aura recrutement de eIF4E qui permettra d'initier la traduction du transcrit. (En bas) Même séquence d'ADN avec présence de Rhino. Cela aura pour conséquence le recrutement de Cutoff qui bloquera le recrutement du complexe CBC. UAP56 sera recrutée mais l'absence du complexe CBC

bloquera l'épissage de l'intron et cela aura pour conséquence une stabilisation de UAP56 sur le transcrit. Le complexe transcrit, Cutoff et UAP56, pourrait être à l'origine de la spécificité des précurseurs de piRNAs. (Zhang *et al.* 2014)

Il a été montré qu'au sein de la lignée germinale, il y a deux populations de clusters : les clusters dépendants de Rhino et les clusters indépendants de Rhino (Mohn *et al.* 2014). Le cluster référence des clusters indépendants de Rhino est le cluster 2 qui est unidirectionnel. Mais les données suggèrent qu'il y aurait également des clusters bidirectionnels indépendants de Rhino (Mohn *et al.* 2014). Une comparaison moléculaire d'un cluster bidirectionnel dépendant de Rhino et d'un cluster bidirectionnel indépendant de Rhino, telle qu'elle a été menée dans la **figure 8**, peut apporter beaucoup de réponses.

- Les clusters unidirectionnels germinaux

Concernant les clusters unidirectionnels, il y en a un seul référencé et appelé cluster 2 (Brennecke *et al.* 2007) (**Figure 7**). Il est localisé dans la région péricentromérique du chromosome X au site cytologique 20A. Il est composé de fragments d'ETs dégénérés et tous insérés en antisens. Ce cluster n'est pas sensible à Rhino c'est-à-dire qu'il n'est pas observé de baisse de la production des piRNAs en condition mutant pour *rhino* bien que ce mutant affecte la localisation d'Aub et AGO3, ainsi que de Piwi, mais uniquement très précocement (Klattenhoff *et al.* 2009 ; Mohn *et al.* 2014 ; **Figure 8**). De plus, des expériences de ChIP-seq de Rhino montrent qu'il y a peu ou pas de Rhino sur le cluster 2 (Klattenhoff *et al.* 2009 ; Mohn *et al.* 2014 ; **Figure 8**). L'étude de la signature transcriptionnelle du cluster 2 (**Figure 8** : cluster 20A) montre que ce cluster est différent du cluster bidirectionnel 42AB. L'analyse de RNA-seq montre la présence de transcrits uniquement sens au début et à la fin du cluster 2 et l'analyse de Cap-seq montre la présence d'un transcrit coiffé issu de la transcription du cluster 2. De plus, les expériences de ChIP-seq de l'ARN PolIII et H3K4me2 montrent un enrichissement de cette protéine et de la marque histone au début du cluster 2 au niveau du TSS (pour « *Transcription Start Site* ») et de l'émergence du transcrit coiffé. Les expériences de ChIP-seq de H3K9me3 montrent un enrichissement de cette marque au début et à la fin de ce cluster et sur les régions adjacentes. Ces résultats font du cluster 2 une région qui ressemble à une unité de transcription génique, ce qui est le cas des clusters de piRNAs décrit chez la souris (Li *et al.* 2013). Il n'y a donc pas de recrutement du complexe RDC sur le cluster unidirectionnel germinale.

Bien que ce cluster 2 ne soit pas sensible à Rhino, il est sensible aux protéines Cutoff et AGO3 (Klattenhof *et al.* 2009 ; Pane *et al.* 2011 ; Mohn *et al.* 2014). En effet, en contexte

mutant pour *cutoff* et *AGO3*, il y a une baisse de la production des piRNAs issus du cluster 2 (Klattenhoff *et al.* 2009 ; Pane *et al.* 2011). Il semblerait que les protéines Cutoff et AGO3 aient un effet sur la biogenèse générale des piRNAs germinaux (Li *et al.* 2009 ; Pane *et al.* 2011). En effet, des expériences d'immunomarquages de chambres ovariennes en contexte mutant pour *AGO3* montrent que les protéines Aub et Vasa ne sont plus présentes dans le nuage et que Piwi n'est plus présent dans le noyau contrairement au phénotype sauvage (Li *et al.* 2009). Concernant le mutant *cutoff*, il y a perte des protéines AGO3, Aub, Vasa et Tejas (**Figure 6**) dans le nuage, alors que la localisation de Piwi ne semble pas être affectée (Pane *et al.* 2011). Ces résultats semblent contradictoires car les travaux de Klattenhoff en 2009 montrent qu'en contexte mutant pour *rhino*, les 3 protéines Piwi, Aub et AGO3 impliquées dans la biogenèse des piRNAs ont également leurs localisations cellulaires perturbées. Néanmoins, l'absence de Rhino sur ce cluster ne change pas le fait que le cluster 2 est localisé dans une région hétérochromatique enrichie en protéines telles que HP1. En effet, il a été montré par ChIP-seq que les régions riches en ETs sont enrichies en marques H3K9me3 dans les ovaires et HP1 dans les ovaires et dans les têtes (Moshkovich *et al.* 2010 ; Rangan *et al.* 2011).

L'ensemble de ces résultats suggère que les modèles proposés précédemment pour les clusters bidirectionnels, concernant le blocage de la terminaison de la transcription et du recrutement du complexe CBC (**Figure 9**), ne soient pas valables pour le cluster unidirectionnel germinale puisque ce cluster possède une transcription canonique. De plus, les travaux de Zhang en 2012 sur la caractérisation de la protéine UAP56 dans la biogenèse des piRNAs montrent que cette protéine n'est pas impliquée dans la biogenèse des transcrits issus du cluster 2. Ceci implique également que le modèle proposé en **Figure 11** n'est pas valable pour le cluster 2. En effet, s'il n'y a pas la présence du complexe RDC pour bloquer le recrutement du complexe CBC au niveau de la coiffe des transcrits naissant, UAP56 sera recrutée par le complexe CBC, ce qui impliquera que l'épissage se fera normalement, conduisant à un relargage d'UAP56. Il n'y aurait donc pas non plus de recrutement de Cutoff au niveau de la coiffe des transcrits naissants ni de stabilisation de la protéine UAP56 sur ces transcrits. Les transcrits précurseurs de piRNAs issus du cluster unidirectionnel germinale sont donc différents de ceux issus des clusters bidirectionnels.

Les modèles proposés par Mohn (**Figure 9**) et Zhang (**Figure 11**) permettent de proposer que la présence des protéines Cutoff et UAP56 au niveau des transcrits issus des clusters bidirectionnels serait suffisante pour définir ces transcrits comme transcrits

précurseurs de piRNAs par la machinerie de piRNAs. Mais ces différences de profil transcriptionnel soulèvent une question : comment les transcrits issus des clusters unidirectionnels sont-ils reconnus en tant que précurseurs de piRNAs dans les tissus germinaux ? Ces résultats mettent le doigt sur l'une des énigmes de ces voies de piRNAs : quelles sont la ou les protéines qui permettent de donner à un transcrit le statut de précurseur de piRNAs ?

b/ La biogenèse des piRNAs primaires

La biogenèse des piRNAs se fait donc à partir des précurseurs de piRNAs qui sont des ARNs simple brin contrairement aux autres voies de petits ARNs : les miRNAs et les endo-siRNAs (voir partie 3 e) qui sont générés à partir d'un précurseur double brin par les protéines Dicer (Kim *et al.* 2009). Les piRNAs sont associés aux protéines de la sous-famille PIWI qui fait partie de la famille ARGONAUTE. Les piRNAs ont une taille comprise entre 24 et 29nt chez la drosophile contrairement aux miRNAs qui ont une taille de 22-23nt et aux endo-siRNAs qui ont une taille de 21nt. Cette différence de taille est peut-être due à l'empreinte laissée sur les ARNs précurseurs par les protéines associées lors de la biogenèse des piRNAs. En effet, la protéine fixée à un transcrit va protéger de la dégradation par des nucléases, la séquence du transcrit recouverte par la protéine. Nous verrons ici comment sont générés les piRNAs dit primaires ainsi que les protéines impliquées, bien que cette étape de biogenèse soit encore très peu connue (**Figure 12**).

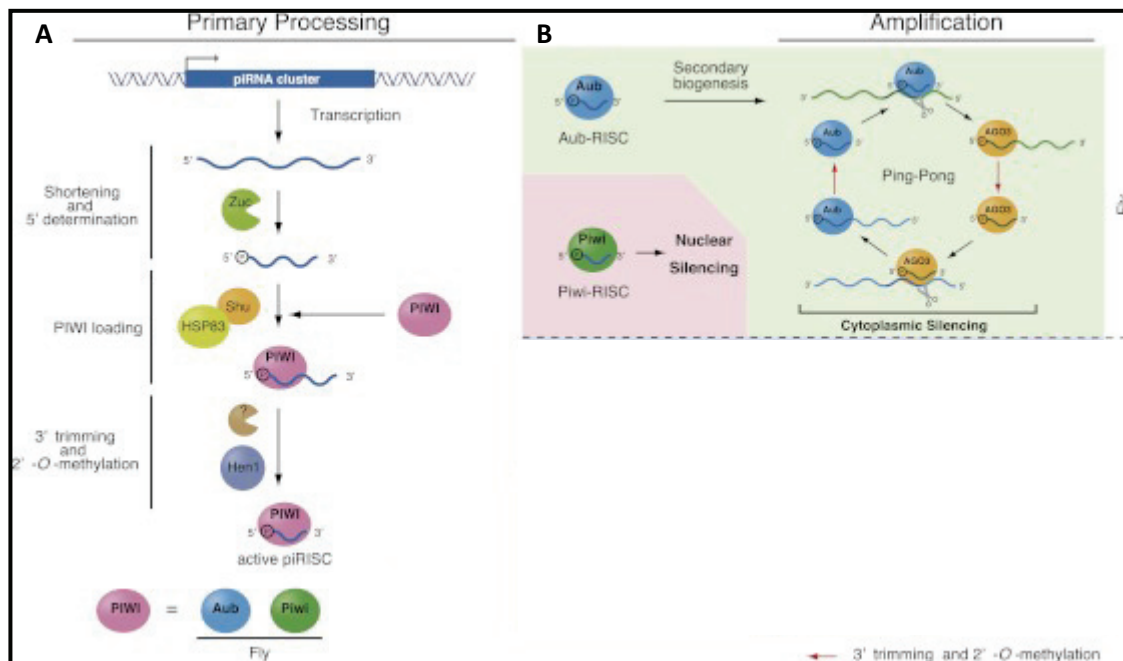


Figure 12 : Modèle de la biogenèse de piRNAs primaires et secondaires chez la drosophile en contexte germinal.

(A) Biogenèse des piRNAs primaires chez la drosophile. Le cluster de piRNAs produit un transcrit qui va être clivé par la protéine Zucchini pour être chargé ensuite par les protéines PIWI : Piwi et Aub. La protéine Hen1 va permettre la 2'O-methylation de l'extrémité 3' après la dégradation en 3' par une ou des protéines inconnues. Les piRNAs 2'O-méthylés et chargés par les PIWI protéines sont appelés piRNAs primaires. (B) Biogenèse de piRNAs secondaires chez la drosophile : les piRNAs chargés par Aub vont reconnaître un transcrit complémentaire et Aub va cliver l'extrémité 5' de ce transcrit qui va être pris en charge par la protéine AGO3. Il y a ensuite dégradation de l'extrémité 3' et 2'-O-methylation pour générer un piRNA secondaire. Ce piRNA secondaire chargé par AGO3 pourra à son tour reconnaître un transcrit complémentaire, le cliver à l'extrémité 5'. Ce transcrit sera alors pris en charge par Aub, l'extrémité 3' va être dégradée et 2'-O-méthylée et générer ainsi des piRNAs secondaires chargés par Aub. Il y aura ainsi amplification des piRNAs secondaires. Les piRNAs primaires chargés par Piwi vont eux entrer dans le noyau pour aller réprimer les séquences homologues cibles. (Ishizu *et al.* 2012)

Il a été montré que les clusters de piRNAs en contexte germinale, qu'ils soient unidirectionnels ou bidirectionnels, produisent des transcrits simple brin appelés précurseurs de piRNAs. Ces longs transcrits vont être conduits jusqu'aux pores nucléaires probablement par la protéine UAP56 mais il y a sûrement d'autres protéines impliquées et non identifiées. Une fois arrivés aux pores nucléaires, ces précurseurs de piRNAs vont passer dans le cytoplasme dans une région subcellulaire optiquement dense et diffuse appelée le nuage. Une des premières protéines décrites dans le nuage est la protéine Vasa qui est requise pour le développement de la lignée germinale et dans la production des piRNAs (Hay *et al.* 1990,

Lasko and Ashburner, 1990 ; Liang *et al.* 1994 ; Pane *et al.* 2007 ; **Figure 5**). Il s'agit d'une protéine possédant un domaine DEAD (box helicase) qui est un motif de fixation des ARNs (Linder and Jankowsky, 2011). La protéine Vasa serait donc capable de fixer les précurseurs de piRNAs afin qu'ils soient clivés tout comme il l'a été montré chez *Bombyx mori* (Xiol *et al.* 2014). Il y a d'autres hélicases présentes dans le noyau pouvant intervenir également telles que SpnE dont les mutants affectent la biogenèse des piRNAs (Vagin *et al.* 2004 ; **Figure 6**). Les facteurs responsables du clivage et de la dégradation des extrémités 5' et 3' sont encore inconnus mais il a été montré que la protéine Zucchini pouvait être un bon candidat pour le clivage des précurseurs de piRNAs (Nishimasu *et al.* 2012). En effet, basée sur des études cristallographiques, la structure de la protéine Zucchini chez la drosophile révèle une activité endonucléase contrairement aux articles précédents qui proposaient Zucchini comme étant une phospholipase (Huang *et al.* 2011). Zucchini a été montrée comme étant associée aux précurseurs de piRNAs et capable de générer des extrémités 5' mono-phosphate, qui sont l'une des caractéristiques des piRNAs matures (Ipsaro *et al.* 2012 ; Nishimasu *et al.* 2012 ; **Figure 12-A**). La particularité de Zucchini est que cette protéine est localisée sur la membrane des mitochondries. Des travaux de Pane en 2007 ont montré que Zucchini est localisée dans le nuage. La question est : quelle est l'implication de la mitochondrie et son rôle potentiel dans la biogenèse des piRNAs ? L'une des hypothèses est que les mitochondries pourraient fournir l'ATP nécessaire aux hélicases telles que Vasa pour clamber les transcrits (Xiol *et al.* 2014).

Ces résultats montrent que Zucchini serait l'endonucléase qui permettrait de faire le clivage 5' des précurseurs de piRNAs et ainsi de générer des pré-piRNAs. Ces pré-piRNAs seraient chargés par les protéines Piwi et Aub dans le nuage à l'aide des protéines Shutdown et Hsp83, qui sont des protéines co-chaperones dont la perte de fonction conduit à des défauts de localisation des protéines de la voie des piRNAs (Schüpbach and Wieschaus 1991 ; Olivieri *et al.* 2012 ; Specchia *et al.* 2010 ; **Figure 12-A**). Le mécanisme d'action de ces protéines reste encore inconnu. Les travaux de Pane en 2007 ont montré l'existence d'une interaction physique entre Zucchini et Aub. Il est possible que la formation d'un complexe avec ces protéines permette le recrutement des protéines Piwi et Aub. Car ce n'est qu'une fois les pré-piRNAs chargés par Piwi ou Aub, qu'il y aurait dégradation de l'extrémité 3' par une ou plusieurs protéines inconnues (**Figure 12-A**). Les cribles menés pour la recherche de nouveaux gènes impliqués dans la voie des piRNAs n'ont pas permis d'identifier de nucléase. Il est possible que cette fonction soit effectuée par plusieurs nucléases différentes, il y aurait donc redondance ce qui ne permettrait pas d'isoler les nucléases impliquées (Czech *et al.*

2013 ; Muerdter *et al.* 2013). Une fois la dégradation de l'extrémité 3' faite, la protéine Hen1 qui est une RNA méthyltransférase va venir modifier l'extrémité 3' en effectuant une 2'-O-methylation (Horwich *et al.* 2007 ; Saito *et al.* 2007 **Figure 12-A**). Cette modification est, selon les espèces, également présente sur les siRNAs mais pas sur les miRNAs (Horwich *et al.* 2007). Néanmoins, l'importance de cette modification reste incertaine car en contexte mutant pour le gène *Hen1*, il n'y a pas de phénotype particulier détecté (Horwich *et al.* 2007 ; Saito *et al.* 2007). Il a été observé un biais de séquences des piRNAs chargés par Piwi et Aub. En effet, la majorité des piRNAs chargés par Piwi et Aub ont un Uracile en position 1 à l'extrémité 5' (Brennecke *et al.* 2007 ; **Figure 12**). A ce jour ce biais n'est pas expliqué. Il est possible que cela provienne du clivage de Zucchini qui cliverait préférentiellement après une uridine ou des protéines PIWI qui se fixerait préférentiellement à un pré-piRNA commençant par une uridine. Cette dernière hypothèse est supportée par des travaux sur les protéines ARGONAUTE qui montrent que le site actif de fixation d'ARNs nécessite une spécificité du nucléotide en 5' (Frank *et al.* 2010 et 2012).

A l'issue de l'ensemble de ces modifications, nous obtenons donc des piRNAs primaires chargés par Piwi ou Aub. D'autres protéines identifiées lors de cribles se trouvent être impliquées dans la biogenèse des piRNAs primaires mais leur rôle précis est encore inconnu. C'est le cas de la protéine Armitage qui est une RNA helicase (Tomari *et al.* 2004 ; Saito *et al.* 2010 ; Olivieri *et al.* 2010 ; **Figure 6**). En contexte mutant pour *Armi*, la localisation de Piwi est affectée mais pas celle d'Aub et AGO3 (Saito *et al.* 2010). Ces données suggèrent que la protéine *Armi* serait impliquée dans le chargement des pré-piRNAs par Piwi. Cette fonction serait possible grâce à la présence de Vreteno, une protéine à domaine Tudor dont la perte de fonction conduit à une perturbation de la production des piRNAs primaires et la localisation de Piwi, Aub et AGO3 (Olivieri *et al.* 2010 ; Handler *et al.* 2011 ; Zamparini *et al.* 2011 ; **Figure 6**). Cette protéine interagit physiquement avec Piwi, Aub et AGO3 et Armitage dans les cellules nourricières. (**Figure 6**). L'ensemble de ces données permet de proposer que ces protéines sont impliquées dans le recrutement de Piwi et Aub sur les pré-piRNAs.

Comme nous venons de le voir, il existe de nombreuses inconnues dans cette partie de la biogenèse des piRNAs. Cependant, le fait que les piRNAs primaires soient chargés par Piwi ou Aub va orienter leur devenir. En effet, les piRNAs chargés par Piwi vont être capables de rentrer dans le noyau pour aller agir sur les copies euchromatiques des ETs ou sur les clusters producteurs de piRNAs (**Figure 12-B**). Par contre, les piRNAs chargés par la

protéine Aub vont rentrer dans une boucle d'amplification appelée cycle Ping-Pong et générer les piRNAs secondaires (**Figure 12-B**).

c/ La biogenèse des piRNAs secondaires

Les piRNAs primaires chargés par Aub vont permettre la production de piRNAs secondaires à l'aide de la troisième protéine PIWI : AGO3. Ces deux protéines Aub et AGO3 sont spécifiques de la lignée germinale et sont localisées dans le nuage (Brennecke *et al.* 2007). Elles possèdent toutes les deux une activité endonucléase tout comme Piwi. Cependant, l'activité endonucléase de Piwi n'est pas indispensable (Saito *et al.* 2006 et 2009 ; Brennecke *et al.* 2007 ; Gunawardane *et al.* 2007 ; Li *et al.* 2009 ; Malone *et al.* 2009 ; **Figure 6, 12 et 13**). Ce sont les activités endonucléases d'Aub et AGO3 qui permettent de générer les piRNAs secondaires.

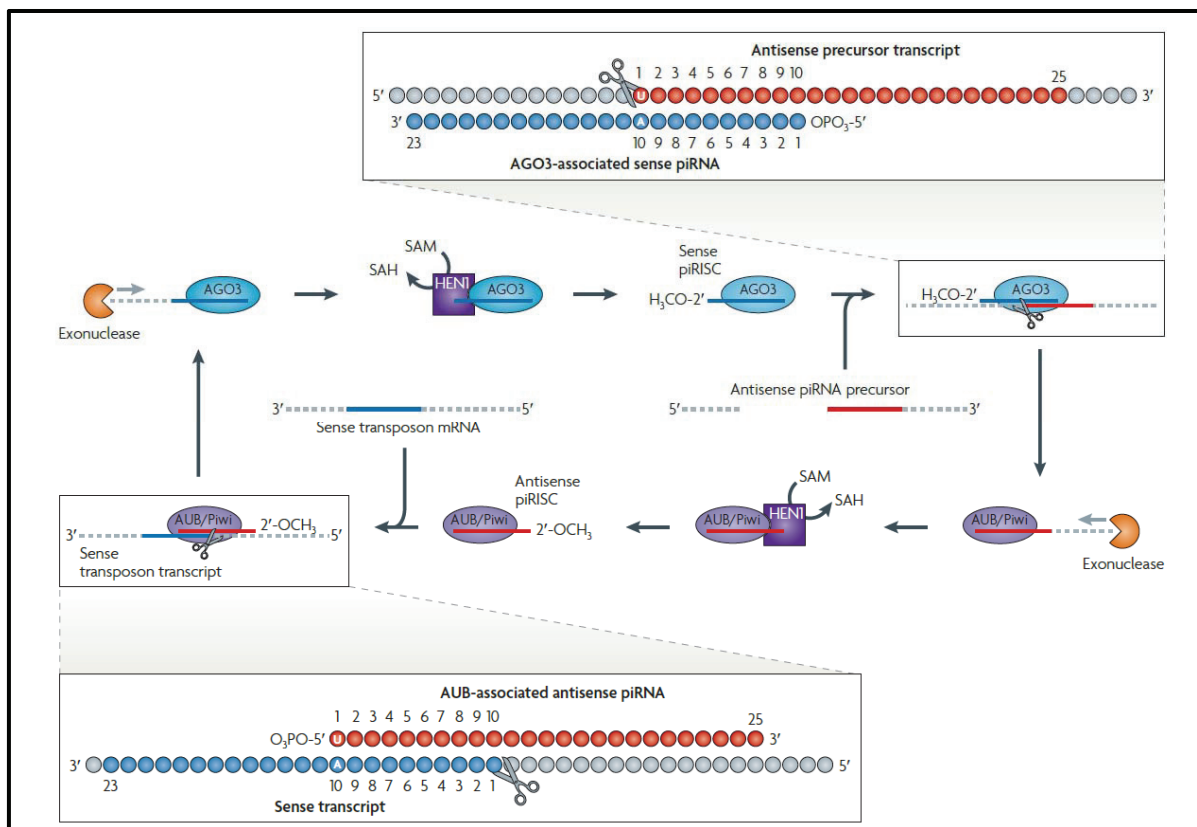


Figure 13 : Biogenèse de piRNAs secondaires (suite).

Biogenèse de piRNAs secondaires (complément de la **Figure 11**). Dans cette Figure nous pouvons nous rendre compte du biais de séquence 1U en extrémité 5' des piRNAs en rouge chargés par Piwi et Aub en violet. Lors de la biogenèse des piRNAs secondaires appelée cycle Ping-Pong, il est observé un biais de séquences des piRNAs secondaires avec un nucléotide A en position 10 en bleu (cadre du bas). En effet, le piRNA chargé par Aub avec un 1U va être complémentaire d'un transcrit bleu (précurseurs de piRNAs ou transcrits d'une copie active d'un ET), le transcrit complémentaire bleu va être clivé par Aub aura un nucléotide A en

position 10 car Aub va cliver à 10 nucléotides de son extrémité 5' (cadre du bas). Le piRNA en bleu va être pris en charge par AGO3, l'extrémité 3' va être dégradée et 2'-O-méthylée par Hen1. Le piRNA en bleu chargé par AGO3 va être complémentaire à un transcrit rouge précurseur de piRNAs et AGO3 va cliver le transcrit rouge en 5' à 10 nucléotides de son extrémité 5' avec un 1U (cadre du haut). Ceci va générer un piRNA secondaire rouge pris en charge par Aub. De la même façon, il y aura dégradation de l'extrémité 3' et 2'-O-méthylation de cette extrémité 3' par Hen1. Un piRNA secondaire rouge chargé par Aub peut à son tour produire un piRNA secondaire et ainsi de suite. (Ghildiyal and Zamore 2009)

Les piRNAs primaires chargés par Aub proviennent des précurseurs de piRNAs qui peuvent être sens et antisens. Il y aurait dans le nuage des transcrits précurseurs sens et antisens et des transcrits sens issus des ETs actifs. Les piRNAs primaires vont donc trouver dans le nuage des transcrits qui leur sont complémentaires. Aub va cliver ce transcrit complémentaire à 10 nucléotides de son extrémité 5' de la position de son premier nucléotide 1U (**Figure 13**). La protéine AGO3 va être recrutée par un mécanisme peu connu et prendre en charge le transcrit clivé en 5' et qui porte un nucléotide Adénine en position 10 (**Figure 12-B et 13**). Il va y avoir dégradation de l'extrémité 3' par des protéines non identifiées et 2'-O-méthylation de l'extrémité 3' par la protéine Hen1. Ceci génère donc un piRNA secondaire chargé par AGO3. Ce piRNA secondaire chargé par AGO3 va à son tour reconnaître un transcrit complémentaire présent dans le nuage. Grâce à son activité endonucléase, AGO3 va cliver les transcrits complémentaires à 10 nucléotide de son extrémité 5' et générer un transcrit avec une uridine en position 1. Par un mécanisme inconnu, la protéine Aub va être recrutée et le transcrit clivé va être chargé en grande majorité par Aub et une très faible proportion par Piwi (Brennecke *et al.* 2007). Il y aura de la même façon dégradation de l'extrémité 3' et 2'-O-méthylation de celle-ci. Il y a donc production de piRNAs secondaires chargés par Aub. On obtient donc un mécanisme d'amplification de la production de piRNAs secondaires encore appelé cycle Ping-Pong (Brennecke *et al.* 2007).

Des cribles par RNAi ont permis d'identifier des protéines qui pourraient participer à cette amplification. C'est le cas de protéines à domaine TUDOR appelé Tudor, Qin/Kumo (Siomi *et al.* 2010 ; Zhang, Koppetsch *et al.* 2014 ; **Figure 6**). La protéine Tudor est localisée dans le nuage et interagit physiquement avec Aub et AGO3 et servirait de plateforme pour la production de piRNAs secondaires par le biais de ses 11 domaines TUDOR qui permettent les interactions protéine-protéine (Nishida *et al.* 2009 ; **Figure 6**). La protéine Qin/Kumo possède un domaine RING, deux domaines B-box et cinq domaines TUDOR (**Figure 5**). Cette protéine est localisée dans le nuage et perturbe la localisation des protéines présentes dans le nuage comme Aub et AGO3 (Anand and Kai, 2012). Il a été constaté qu'en contexte mutant

pour *qin/kumo*, il y a perturbation des piRNAs secondaires et plus particulièrement du ratio de piRNAs secondaires chargés par Aub et AGO3 (Anand and Kai, 2012). Cette protéine permettrait le bon recrutement des protéines Aub et AGO3 lors de la biogenèse de piRNAs secondaires. Ce mécanisme d'amplification requiert donc la présence d'un certain nombre de protéines probablement encore non identifiées ou identifiées mais dont le rôle exact n'a pas encore été mise à jour. Il est très probable cependant que les protéines qui composent le nuage, puissent potentiellement former des complexes. Ceux-ci permettraient le rapprochement des transcrits (aussi bien précurseurs de piRNAs que transcrits des ETs) des protéines PIWI afin de générer un pool de piRNAs primaires et secondaires.

Ce modèle d'amplification présente de nombreuses inconnues. Néanmoins, si on analyse ce modèle, il est possible de faire plusieurs constatations. Tout d'abord, les piRNAs primaires, issus du clivage de Zucchini, ne sont chargés que par Piwi et Aub. AGO3 ne charge que des piRNAs secondaires. Mais dans le pool de piRNAs secondaires, il y en a aussi chargés par Aub et en moindre mesure par Piwi. Il est à ce stade impossible, par les techniques d'analyses bio-informatiques actuelles, d'identifier un piRNA primaire d'un piRNA secondaire chargé par Aub car il présente le même biais de séquence à savoir un 1U. Par contre, il existe un biais de séquence qui permet de déterminer ce que l'on appelle la signature Ping-Pong car il se crée des couples de piRNAs sens et antisens qui se chevauchent de 10nt et donc avec un piRNA commençant par un U en position 1 et son piRNA complémentaire avec nucléotide A en position 10. Cette signature est spécifique du cycle Ping-Pong et de la lignée germinale car Aub et AGO3 ne sont présents que dans la lignée germinale. De plus, le cluster 2 qui est un cluster unidirectionnel, ne produit que des transcrits sens ne présentant pas de signature Ping-Pong (Brennecke *et al.* 2007). Ceci permet de soulever une remarque qui est que lors de cette amplification, les transcrits sens des ETs peuvent participer à la production de piRNAs secondaires. Bien que le cluster 2 produise des transcrits précurseurs sens, tous les fragments ETs insérés dans ce cluster sont insérés en sens inverse. Cela signifie que les piRNAs primaires issus de ce cluster sont sens par rapport au sens de transcription du cluster mais la séquence de ces piRNAs est la séquence antisens des fragments ETs. Ceci signifie que les piRNAs primaires sont complémentaires aux transcrits des ETs actifs. Il serait donc possible de générer des piRNAs secondaires issus des copies d'ETs actifs comme cela a été montré avec l'élément *I* (Brennecke *et al.* 2008). Ceci devrait permettre d'observer une signature Ping-Pong ce qui n'est pas le cas. Ceci peut être dû au fait qu'il y a peu de piRNAs secondaires générés et qu'ils ne sont pas détectés à l'échelle d'observation de ce cluster. Une autre explication serait que les analyses de profils de petits

ARNs sont faites sur les séquences uniques du cluster (pour être sûr d'analyser uniquement les piRNAs issus de ce cluster). Donc par définition, les piRNAs secondaires générés à partir des copies d'ETs actifs sont éliminés des analyses. Ce type de résultats nous montre que les analyses que l'on peut faire sur les petits ARNs comportent de nombreux biais non seulement techniques lors du séquençage mais également lors de l'analyse des données. Les conclusions que l'on peut tirer de ces analyses doivent être prises avec précaution.

La biogenèse de piRNAs secondaires permet de générer un pool de piRNAs secondaires qui peuvent être impliqués dans plusieurs processus. La petite proportion de piRNAs secondaires chargés par Piwi peut potentiellement entrer dans le noyau pour agir sur ses cibles (copies actives des ETs ou clusters). Lors de la biogenèse de piRNAs secondaires, les transcrits issus de copies actives d'ETs peuvent être clivés en piRNAs secondaires et cela signifie que ces transcrits sont dégradés. Il s'agit d'un mécanisme de régulation par PTGS (pour « *Post transcriptionnal Gene Silencing* »). La biogenèse des piRNAs secondaires permet une amplification de la quantité de piRNAs secondaires. Il a été montré qu'il y a une accumulation des protéines Aub et Piwi chargées avec leurs piRNAs au pôle postérieur des embryons, qui est le lieu où se forment les cellules de la future lignée germinale (Harris and Madocnald, 2001 ; Megosh *et al.* 2006 et Brennecke *et al.* 2008). Ce dépôt de piRNAs maternels semble essentiel à la mise en place de la biogenèse de piRNAs pour la descendance et la régulation de ces ETs. Ceci permettrait le maintien *trans*-générationnel de la production de piRNAs et donc de la régulation des ETs. Des résultats allant dans ce sens ont été trouvés lors de l'étude de l'émergence du cluster de piRNAs que j'étudie et ce point sera discuté dans les parties Résultats et Discussion.

C/ Répression des ETs euchromatiques par les piRNAs germinaux

La biogenèse des piRNAs permet la régulation des ETs présents dans le génome et dont la transposition peut être délétère pour la future descendance. Nous allons voir comment ces ETs peuvent être régulés par les piRNAs.

a/ Régulation Post-transcriptionnelle : PTGS

La biogenèse de piRNAs secondaires se fait par le biais d'une boucle d'amplification à partir des protéines Aub et AGO3. Lors de ce processus, les piRNAs chargés par l'une ou

l'autre des protéines Aub ou AGO3 vont reconnaître un transcrit qui leur est complémentaire et ce transcrit peut être, soit un précurseur de piRNAs, soit un transcrit issu d'une copie active d'un ET. Dans le cas où le transcrit complémentaire est un transcrit issu d'une copie active d'un ET, ce transcrit va être clivé par les protéines PIWI pour générer un piRNA secondaire. Le fait de produire un piRNA secondaire va provoquer la dégradation de ce transcrit.

Les travaux sur l'élément *I*, un ET qui peut conduire à des phénomènes de dysgénésie hybride, a permis de mettre en évidence l'existence d'un PTGS lié au modèle Ping-Pong (Brennecke *et al.* 2008). La dysgénésie est observée lorsque des femelles d'une souche sans élément *I* sont croisées par des mâles possédant des copies euchromatiques de l'élément *I*. La descendance issue de ce croisement présente une forte stérilité liée la transposition active de ce transposon (Picard *et al.* 1976). Par contre, le croisement réciproque donne une descendance fertile. L'analyse de piRNAs ovariens des souches parentales et de leurs descendance a permis de montrer qu'il y a production de deux types de piRNAs. Les piRNAs issus des clusters de piRNAs et des piRNAs issus des copies actives des éléments *I*, identifiés grâce à la divergence de séquences des éléments dans les clusters de piRNAs. Il a été montré dans le cas des éléments *I* que les piRNAs issus de copies actives sont sens tandis que ceux des piRNAs issus des clusters sont sens et antisens. Ces piRNAs issus des copies actives ne peuvent exister que s'ils ont été générés par le clivage d'Aub (**Figure 13**). C'est d'ailleurs de cette façon qu'a été proposé le modèle de Ping-Pong par Brennecke en 2007, car il a été observé un biais de séquences de clusters. Néanmoins, l'analyse des clusters de piRNAs a montré que la majorité des clusters de piRNAs sont bidirectionnels et donc produisent des précurseurs aussi bien sens qu'antisens. Donc un cluster de piRNAs bidirectionnel peut produire des piRNAs primaires sens et antisens et des piRNAs secondaires sens et antisens indépendamment des transcrits de copies actives d'ETs. L'analyse de profil de piRNAs est complexe car il est difficile d'isoler les différents piRNAs ovariens et de leur attribuer une origine non seulement tissulaire (présence des deux types de tissus) mais également génomique.

b/ Régulation transcriptionnelle : TGS

La régulation des ETs peut donc passer par une régulation post-transcriptionnelle de leurs transcrits *via* le cycle Ping-Pong mais cette régulation ne peut avoir lieu que dans la lignée germinale, puisque Aub et AGO3 sont spécifiques de ce tissu. L'un des mécanismes de répression des ETs serait la formation de l'hétérochromatine qui permet de bloquer la

transcription en empêchant le recrutement de l'ARN polIII. De plus, les ETs sont retrouvés dans des régions hétérochromatiques riches en marques H3K9me3 et HP1, qui sont affectées par un mutant *spnE* impliqué dans la voie des piRNAs (Klenov *et al.* 2007 ; Rangan *et al.* 2011 ; **Figure 6**). Il s'agit d'un mécanisme difficile à étudier à cause précisément de la nature répétée des ETs et de la coexistence des clusters de piRNAs et des copies « actives » des ETs. L'une des hypothèses suspectées depuis longtemps est que Piwi serait à l'origine du dépôt de ces marques hétérochromatiques et qu'il y aurait un lien entre les voies de piRNAs et formation d'hétérochromatine. Ceci est suggéré par le fait que la protéine Piwi, qui peut charger des piRNAs est localisée dans le noyau (Brennecke *et al.* 2007). Des études sur un mutant de *piwi* qui affecte uniquement sa capacité à entrer dans le noyau (délétion du domaine NLS (pour « nuclear localisation signal ») : *piwi^{Nt}*) ont montré qu'il y a perte de la régulation des ETs et que le profil de marques chromatinienne de certains ETs est affecté (Klenov *et al.* 2011 ; 2014). En effet, des expériences de ChIP q-PCR ont montré qu'en contexte hétérozygote pour *piwi^{Nt}*, il y a présence de H3K9me3 et HP1 sur certains ETs mais qu'en contexte mutant pour *piwi* (*piwi^{Nt/2}*), il y a une réduction de H3K9me3 et HP1 et une augmentation de la marque H3K4me2. Néanmoins, ces résultats ne s'appliquent pas à l'ensemble des ETs car cela dépend de leur localisation génomique, de leur lieu de transcription et du cluster qui les régule. En effet, il a été observé que tous les ETs ne répondent pas de la même façon en contexte mutant pour *piwi^{Nt}* (Klenov *et al.* 2014). Pour certains ETs, la présence de H3K9me3 et HP1 réprime leur transcription mais pour d'autres le fait d'être enrichis en H3K9me3 et HP1 non seulement n'affecte pas leur transcription mais au contraire la favorise. C'est en partie pour cela qu'il est difficile d'étudier le TGS des ETs par la voie de piRNAs. Néanmoins, des études menées avec des transgènes contenant des séquences d'ETs et insérés de façon aléatoire dans le génome, ont montré la répression du gène rapporteur associé aux transgènes (Sentmanat and Elgin, 2012). La délétion de la région ciblée par les piRNAs affecte la répression du transgène. De plus, ils ont montré que ces transgènes permettaient le recrutement de marques H3K9me3 et HP1 au niveau de la région d'insertion, euchromatique à l'origine.

Un second élément permettant de penser qu'il y a une interaction entre la voie de piRNAs et la formation de l'hétérochromatine, est une étude faite sur l'interaction physique des protéines Piwi et HP1 (Brower-Toland *et al.* 2007). En effet, ils ont montré sur chromosomes polytènes de glandes salivaires de larves que Piwi colocalise avec la protéine HP1a mais pas HP1b et c. De plus, des expériences de double hybride en levure ont montré que Piwi interagit physiquement avec HP1a sous forme de dimère au niveau d'un motif

PRVKV de Piwi. Pour confirmer ces résultats, ils ont testé fonctionnellement cette interaction. Ils ont utilisé un transgène rapporteur *white* variégué : w^{var} (sensible au PEV : position effect variegation) qui montre un phénotype d'œil blanc en contexte sauvage. Les yeux sont rouges en contexte mutant pour *piwi*. S'ils insèrent un transgène portant un allèle sauvage de *piwi*, dans des mouches mutées pour *piwi*, il y a à nouveau répression du transgène w^{var} , les yeux des mouches sont clairs. Il y a donc sauvetage fonctionnel. Par contre, avec un transgène portant une forme mutée de *piwi* au niveau du motif PRVKV, il n'y a pas sauvetage et les yeux sont très rouges. Cela montre que le motif PRVKV de Piwi est bien impliqué dans le silencing par la formation de l'hétérochromatine. Ces résultats soulèvent une question : est-ce que cette répression est faite par Piwi dans les cellules de l'œil à l'âge adulte ou est-ce que cette répression est l'action de Piwi précocement au cours du développement, qui aurait déposé des marques répressives qui seraient maintenues ensuite dans les tissus somatiques jusqu'à l'âge adulte ? Des études faites sur les cellules somatiques ovariennes montrent que Piwi est capable de réprimer les ETs par TGS. L'argument le plus pertinent est une étude faite spécifiquement dans les tissus germinaux qui montre qu'une déplétion de Piwi dans les cellules germinales conduit à une dérégulation de certains ETs accompagnée d'une diminution de la quantité HP1 et H3K9me2 sur ces mêmes ETs comparé au sauvage (Wang *et al.* 2011).

Bien que ces données soient indirectes, elles permettent de montrer qu'il est possible de faire un lien entre la voie de biogenèse des piRNAs et la formation de l'hétérochromatine. Un cas de boucle RNAi et formation d'hétérochromatine a été mis en évidence chez *S. pombe* via un complexe appelé le RITS (pour « *RNA-induced Initiation of Transcriptional gene Silencing* ») (Volpe *et al.* 2002 ; Verdel *et al.* 2004). La purification du complexe RITS a montré la présence des protéines AGO1, Chp1 (protéine hétérochromatique à chromo-domaine) et la protéine Tas3 (binding protein) avec des siRNAs dépendants de Dicer. La protéine Chp1 est liée aux régions répétées centromériques et requise pour la méthylation de H3K9 ainsi que la localisation de la protéine Swi6 (protéine homologue de HP1) (Partridge *et al.* 2000 ; 2002). Verdel et ses collaborateurs ont également montré que la protéine Tas3 était importante pour la méthylation de H3K9 et le recrutement de Swi6 et son interaction avec Chp1 sont essentiels pour l'assemblage du RITS car elle permet d'associer Chp1 et AGO1 (Schalch *et al.* 2011). La protéine AGO1 est également présente dans le complexe RISC (RNA-Induced Silencing Complex) qui permet de cibler un ARNm spécifique et de l'inactiver par la complémentarité de séquences avec le siRNA contenu dans le complexe RISC. La présence d'AGO1 dans le complexe RITS suggère la présence de siRNAs. Les

petits ARNs identifiés dans le complexe RITS montrent qu'ils font la taille des siRNAs et qu'ils sont dépendants de Dicer. De plus, leur analyse par Northern-Blot a montré que ces siRNAs sont issus de régions hétérochromatiques centromériques spécifiques. En effet, l'analyse des siRNAs chez *S. pombe* a montré que les siRNAs sont issus de deux régions centromériques distinctes : *dh repeats* et une région en aval de *dg repeats* (Reinhart and Bartel 2002). L'ensemble de ces données a permis de proposer un modèle dans lequel André Verdel et ses collaborateurs proposent que des siRNAs issus de régions hétérochromatiques centromériques sont générés puis pris en charge par un complexe AGO1-Chp1-Tas3 pour former un complexe RITS actif (**Figure 14**). Ce complexe RITS va pouvoir reconnaître un transcrit naissant complémentaire aux siRNAs et permettre la formation d'hétérochromatine au niveau de cette région. De plus, il a été montré que le recrutement du siRNA sur le transcrit naissant permettait le recrutement d'un complexe RDRC (pour « *RNA-dependant RNA polymerase Complex* ») qui permet de polymériser le brin complémentaire du transcrit naissant formant ainsi un ARN double-brins reconnu par Dicer qui génère les siRNAs (Creamer *et al.* 2011, revue ; **Figure 14**). Il s'agit donc d'une boucle de rétro-contrôle où les siRNAs issus des régions centromériques ciblent et détruisent les transcrits de ces mêmes régions en produisant des siRNAs.

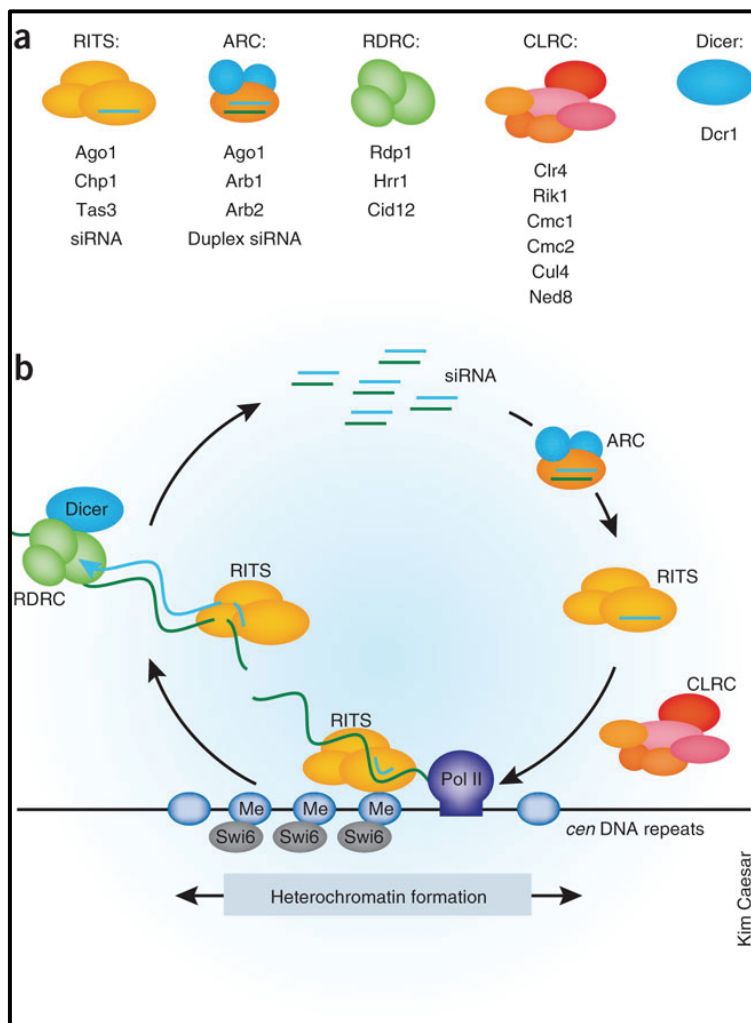


Figure 14 : Boucle RNAi-formation d'hétérochromatine via le complexe RITS chez *S. pombe*.

(a) Les différents complexes intervenant dans le complexe RITS. (b) Mécanisme moléculaire du RITS sur les régions centromériques répétées. Les transcrits de ces régions répétées sont clivés en siRNAs par Dicer et vont être pris en charge par le complexe ARC (*Ago1 chaperone complexe*) puis RITS. Le complexe RITS chargé avec un siRNA va reconnaître les transcrits naissants de ces régions répétées centromériques et permettre le recrutement de protéines hétérochromatiques. Dans un second temps, ce complexe RITS va permettre l'initiation de production d'ARN double-brins à partir du transcrit naissant grâce au complexe RDRC qui permettra avec la protéine Dicer la production de siRNAs et donc la dégradation de

ce transcrit. (Bulher *et al.* 2007)

Ce modèle est particulièrement intéressant car il permet de montrer que les siRNAs peuvent permettre la formation d'hétérochromatine au niveau de ses cibles et à la fois permettre le maintien de sa production. Si on applique ce modèle aux piRNAs, il est possible d'envisager une boucle similaire : l'existence des piRNAs favorise la production de nouveaux piRNAs. Ce modèle permet de soulever une question majeure. Si on admet que Piwi chargé par un piRNA réprime par TGS une séquence qui lui est homologue, de qui parle-t-on : des séquences d'ETs actifs ou des fragments de séquences des ETs dans les clusters ? Comment un piRNA chargé par Piwi peut-il discriminer une séquence d'ETs actifs d'un fragment de séquence d'ETs présent dans un cluster de piRNAs ? Ce problème n'est pas toujours pris en compte dans la littérature.

c/ Maintien des clusters de piRNAs

Nous avons vu qu'il existe un certain nombre d'arguments en faveur de l'existence d'un lien entre RNAi et formation d'hétérochromatine. Or, l'analyse chromatinienne des clusters a montré que les clusters sont insérés dans des régions hétérochromatiques et sont enrichis en marques hétérochromatiques H3K9me3 et Rhino dans le cas des clusters bidirectionnels et enrichis en HP1 dans le cas des clusters unidirectionnels impliqués dans le RCCs (pour « *Repressive Chromatin Centers* ») (Yoon *et al.* 2008 ; Rangan *et al.* 2011 ; Mohn *et al.* 2014 ; **Figure 8**). En effet, il a été montré entre autres par Rangan et ses collaborateurs en 2011 que la formation d'hétérochromatine est primordiale à la répression des ETs.

De plus, les travaux de Mohn et ses collaborateurs en 2014 montrent qu'en contexte mutant pour *piwi*, il y a une réduction de Rhino sur les clusters bidirectionnels et que cela affecte la production de piRNAs de ces clusters. L'ensemble de ces données permet de suggérer que Piwi permettrait le maintien de l'état hétérochromatique des clusters, primordial à la biogenèse des piRNAs (**Figure 15**). La voie de piRNAs peut être considérée comme une boucle de régulation dans laquelle des clusters de piRNAs hétérochromatiques produisent des piRNAs primaires et secondaires qui permettent la régulation des copies d'ETs actifs mais aussi le maintien des propriétés fonctionnelles des clusters de piRNAs hétérochromatiques eux-mêmes (**Figure 15**). Ce modèle ressemble à ce qui est proposé pour le mode de fonctionnement du RITS chez *S pombe*.

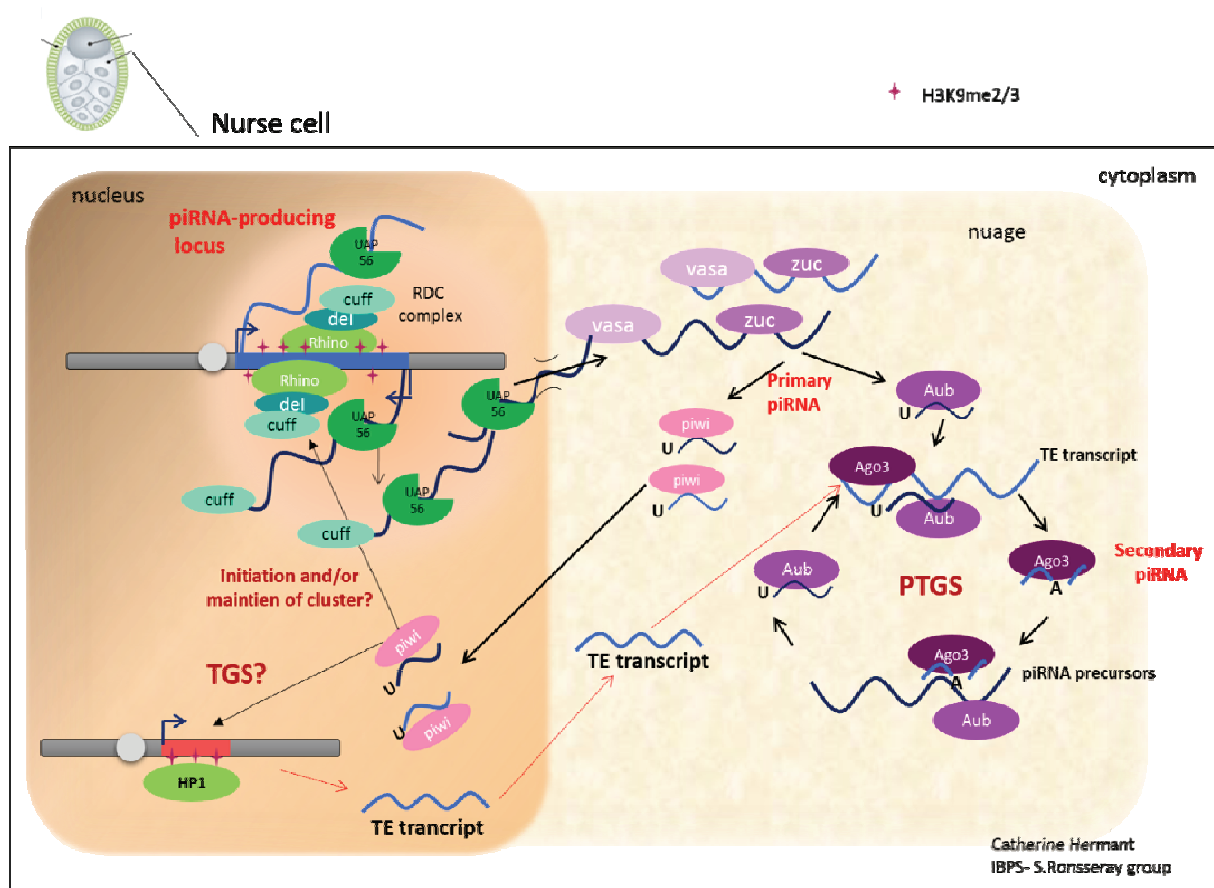


Figure 15 : Modèle de biogenèse des piRNAs et leur mode régulation dans les cellules nourricières.

Les clusters bidirectionnels de piRNAs en bleu produisent des transcrits précurseurs de piRNAs sens et antisens. Ils vont être clivés en piRNAs primaires puis secondaires dans le nuage par les protéines Zucchini et les protéines PIWI. Les piRNAs chargés par Piwi vont être capables de rentrer dans le noyau et d'interagir avec leurs cibles, grâce à la complémentarité de séquences avec les transcrits naissants. Dans le cas des copies d'ETs actifs en rouge, il y aura formation d'hétérochromatine par le recrutement de la marque H3K9me3 et de la protéine HP1. Dans le cas des clusters de piRNAs en bleu, il y aurait dépôt de la marque H3K9me3 qui permettra le recrutement de la protéine Rhino et donc la formation du complexe RDC, puis le maintien des capacités de production de piRNAs des clusters. Dans le cas où les copies d'ETs actifs seraient transcrites, les transcrits sens en bleu pourraient être clivés en piRNAs par la boucle d'amplification appelé cycle Ping-Pong et donc dégradés par PTGS.

Ce modèle expliquerait en partie la biogenèse et le mode de régulation des piRNAs issus de clusters bidirectionnels. Mais qu'en est-il du cluster 2 unidirectionnel ? Il a été montré qu'il n'y a pas la présence du complexe RDC sur ce cluster 2 (Mohn *et al.* 2014 ; **Figure 8**). Il ne fonctionne donc pas comme un cluster bidirectionnel. Néanmoins, ce cluster 2 se trouve dans une région péri-centromérique hétérochromatique. Il est possible que Piwi chargé par son piRNA sens produit par le Ping-Pong va reconnaître par homologie de séquences les transcrits antisens produit par le cluster 2 et permettre le recrutement de HP1.

Une autre possibilité est que la nature hétérochromatique constitutive des régions péricentromériques ne requiert pas l'intervention des piRNAs et favorise, par un déterminisme génétique, la production des transcrits précurseurs de piRNAs. Un autre cluster présente des contradictions : les TAS du bras chromosomique 3R, un cluster bidirectionnel de piRNAs. Ce cluster produit des piRNAs avec un fort biais de piRNAs sens. Les TAS sont des régions hétérochromatiques constitutives pour lesquels une mutation de *piwi* conduit à une augmentation de la marque H3K9 me3 et HP1 associée à une dérégulation des ETs (Yin et Lin 2007). Ceci signifie que selon les marques chromatinienne et leur quantité, un cluster de piRNAs peut produire ou non des précurseurs de piRNAs. Donc, la notion d'hétérochromatine, que ce soit dans le cas de la répression des ETs actifs ou dans le maintien des clusters de piRNAs, reste encore imprécise et parfois même contradictoire.

Ce modèle soulève une autre question. Pourquoi y aurait-il un recrutement de Rhino sur certains clusters bidirectionnels seulement et au contraire un recrutement de HP1 sur certains clusters bidirectionnels et unidirectionnels ainsi que sur les copies actives d'ETs ? L'idée est que Piwi permet le recrutement de la marque H3K9me3 et le recrutement de HP1 de façon générale, mais il existerait probablement des facteurs spécifiques (marques histones ou protéines chromatinienne) sur certains clusters bidirectionnels qui permettraient en plus le recrutement de la protéine Rhino. Il y aurait donc une balance entre les protéines HP1 et Rhino au niveau de ces clusters qui favoriserait ou pas la production de piRNAs.

Le bilan de cette partie est que la biogenèse des piRNAs ainsi que leur mode de régulation sont encore très partiellement compris et certaines données se contredisent. L'hétérogénéité des ETs, de leur localisation et de leur état chromatinien fait qu'il est difficile de tirer des conclusions générales sur cette voie germinale de « *Silencing* ».

Nous venons de voir la biogenèse de piRNAs germinaux qui permettent une régulation efficace des ETs. Mais cette régulation est aussi nécessaire dans les cellules folliculaires somatiques des ovaires (**Figure 4 et 5**) mais aussi dans tous les autres tissus somatiques.

2/ Régulations somatiques par les petits ARNs non-codants chez la drosophile

Certains ETs ont la faculté de transposer dans n'importe quelle cellule. Bien que la transposition au niveau de la lignée germinale puisse avoir des répercussions sur la future descendance, il est également nécessaire que leur transposition soit réprimée également dans les tissus somatiques. En effet, ces transpositions peuvent générer des mutations qui peuvent participer à des processus de cancérogenèse. De plus, il a été montré que certains ETs avaient la capacité de transposer et de passer d'une cellule d'un tissu à une cellule d'un autre tissu. C'est le cas par exemple de *Gypsy*, rétrotransposon à LTR transcrit dans les cellules folliculaires, qui possède la capacité d'envahir les cellules germinales et dont la régulation est faite par le locus *flamenco* en cellules folliculaires (Lécher *et al.* 1997 ; Mevel-Ninio *et al.* 2007). Il y a vis-à-vis de la régulation des ETs, deux types de tissus somatiques, les tissus somatiques de la carcasse et les tissus somatiques folliculaires qui entourent les cellules germinales (**Figure 4 et 5**). Nous verrons dans cette partie qu'il existe deux types de régulations différents : la voie somatique des piRNAs qui prédomine dans les tissus somatiques folliculaires et la voie des siRNAs qui prédomine dans les tissus somatiques de la carcasse.

A/ Biogenèse des piRNAs somatiques folliculaires

Les cellules folliculaires sont des cellules somatiques qui entourent les cellules nourricières et l'ovocyte. Il existe des ETs qui ont la capacité de transposer à partir des cellules folliculaires vers les cellules germinales. Un système de régulation a été mis en place en cellules folliculaires, qui ressemble à celui observé dans la lignée germinale. En effet, ces cellules folliculaires contiennent des clusters de piRNAs qui permettent la production de piRNAs uniquement primaires car nous verrons que ces cellules n'ont pas les protéines nécessaires à la biogenèse de piRNAs secondaires. Ces piRNAs primaires sont chargés par Piwi et peuvent entrer dans le noyau pour réprimer leurs cibles. Toutes ces analyses ont pu être faites sur des cultures de cellules issues de cellules folliculaires appelées OSS (pour « *Ovarian Somatic Sheet* ») et OSC (pour « *Ovarian Somatic Cells* ») et confirmées dans de nombreux cas sur ovaire.

a/ Les clusters de piRNAs et la production des transcrits précurseurs de piRNAs

Tout comme dans les cellules germinales, les piRNAs peuvent être issus de clusters. La particularité des clusters de piRNAs somatiques, est qu'ils sont principalement unidirectionnels. Le cluster de piRNAs somatiques de référence est *flamenco/COM* est un cluster spécifique des cellules somatiques folliculaires. Ce cluster est localisé sur le chromosome *X* dans une région péricentromériques hétérochromatique qui fait environ 180Kb. Il a été montré capable de réguler différentes familles d'ETs : *ZAM*, *Idéfix* et *Gypsy* (Tcheressi *et al.* 2002 ; Desset *et al.* 2003). Il est composé à 85% d'ETs qui sont majoritairement tous insérés dans la même orientation, en antisens du sens de transcription (Robert *et al.* 2001 ; Celniker *et al.* 2002 ; Bergman *et al.* 2006 ; Brennecke *et al.* 2007 ; Malone *et al.* 2009). Il a été montré que, comme les clusters 42AB et cluster 2 de la lignée germinale, le cluster *flamenco/COM* est sensible au mutant *dSETDB1* (ou *eggless*) qui est une méthyle transférase de l'histone H3K9 et donc enrichie en marques H3K9me3 (Rangan *et al.* 2011). En contexte mutant pour *dSETDB1*, il y a une réduction de la quantité de piRNAs issus de ces clusters (Rangan *et al.* 2011). En effet, l'analyse des petits ARNs de ce cluster a montré qu'il y a production que de piRNAs issus du brin plus et que ce cluster ne produit des piRNAs que dans les cellules folliculaires (Brennecke *et al.* 2007 ; Malone *et al.* 2009, **Figure 7**). Il a été pendant longtemps difficile de détecter les transcrits précurseurs de *flamenco/COM*. Une étude de Saito et ses collaborateurs en 2010 a permis d'isoler des transcrits issus de *flamenco/COM* de 25 à 70 nt dans les Yb body, structure cytoplasmique particulière contenant certaines protéines impliquées dans la biogenèse de piRNAs somatiques. Ceci montre que les piRNAs somatiques sont générés à partir de plus grands transcrits. Mais ce n'est que récemment qu'une étude transcriptionnelle de *flamenco* a pu être menée (Goriaux *et al.* 2014). Cette étude a permis d'identifier le promoteur de *flamenco* nommé *Inr DPE promoter* et localisé à 1743pb du gène *DIP1* situé en amont de *flamenco/COM*. Cette étude a permis de montrer que le cluster *flamenco/COM* produisait des transcrits épissés différents dont le seul exon commun est l'exon 1 et que l'expression de ce cluster dépend d'un facteur de transcription CI (pour « *cubitus interruptus* »). De plus, des expériences de RNA Fish des transcrits de *flamenco* ont permis de montrer que les transcrits précurseurs de piRNAs de *flamenco/COM* étaient localisés sous forme de foyer ponctiforme appelé DOT COM dans le noyau ou à sa périphérie dans le cytoplasme et que le DOT COM

est situé à proximité des Yb body cytoplasmiques (Dennis *et al.* 2013 ; Murota *et al.* 2014). Il y aurait donc une accumulation de transcrits précurseurs de piRNAs somatiques à la périphérie du noyau en vis à vis du Yb body cytoplasmique contenant les protéines nécessaires à la biogenèse de piRNAs primaires.

b/ Biogenèse des piRNAs primaires somatiques

Une fois les transcrits précurseurs de piRNAs produits, ils vont être emmenés dans le cytoplasme par des protéines encore non identifiées et ce jusqu'aux Yb body. Les Yb body sont des structures cytoplasmiques contenant les protéines nécessaires à la biogenèse des piRNAs primaires (**Figure 6 et 16**). Les protéines retrouvées dans les Yb body sont Yb, protéines à domaines hélicases et DEAD, spécifiques de la lignée somatique folliculaire (Saito *et al.* 2010), So Yb, protéines à domaine DEAD (Handler *et al.* 2011), Shutdown (Olivieri *et al.* 2012), Vreneto (Handler *et al.* 2011) et Armitage (Saito *et al.* 2010) impliquées également dans la biogenèse de piRNAs primaires germinaux. Dans les cellules folliculaires, il n'y a donc que la production de piRNAs primaires. En effet, les précurseurs de piRNAs folliculaires vont être clivés en 5', pour générer des piRNAs prématures, par Zucchini présent sur la membrane des mitochondries localisées à proximité des Yb body, aidé par des protéines contenues dans les Yb body (Malone *et al.* 2009 ; **Figure 16**). Ces piRNAs prématures vont être chargés par la protéine Piwi et tout comme pour les piRNAs germinaux, il y aura dégradation de l'extrémité 3' et 2'O-méthylation par la protéine *Hen1* ou une autre protéine non identifiée. Il est ainsi généré des piRNAs primaires matures chargés par Piwi uniquement car c'est la seule protéine PIWI exprimée dans les cellules folliculaires (Malone *et al.* 2009).

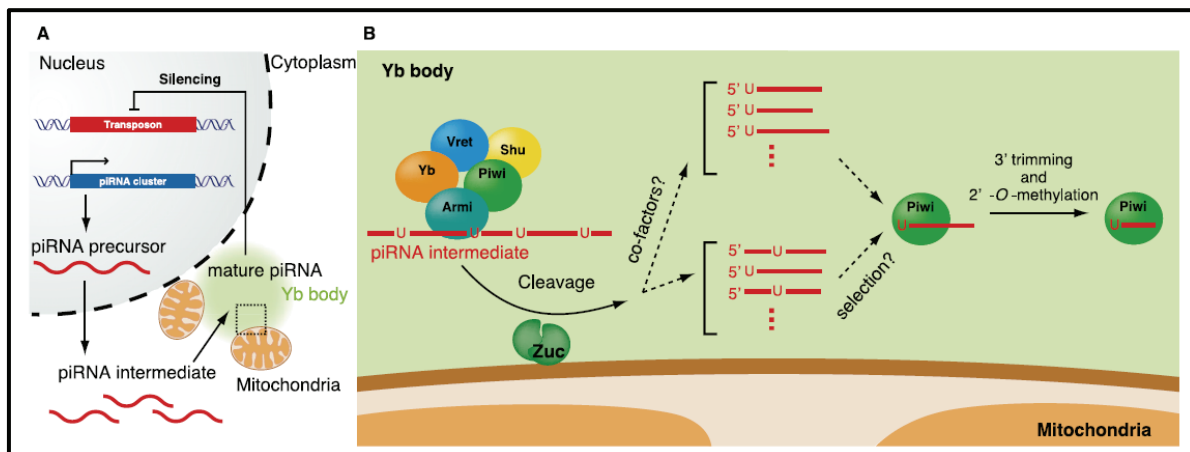


Figure 16 : Biogenèse des piRNAs primaires folliculaires.

(A) Modèle de production des piRNAs folliculaires et de leur mode de répression. Production de piRNAs à partir des clusters de piRNAs en bleu qui produiront des piRNAs matures par l'intervention de protéines contenues dans les Yb body et à la surface des mitochondries et chargés par la protéine Piwi. Les piRNAs matures chargés par Piwi rentrent ensuite dans le noyau pour réprimer les copies actives d'ETs en rouge. (B) Schéma expliquant le rôle et les interactions entre les Yb body et la mitochondrie. La mitochondrie se trouve accolée aux Yb body. Les protéines à domaine Hélicase Armitage et Yb présentes dans les Yb body vont former un complexe avec les précurseurs de piRNAs. Zucchini qui se trouve à la surface de la membrane des mitochondries va cliver les piRNAs précurseurs apportés par le complexe Armi-Yb à l'extrémité 5'. *In vitro*, Zucchini ne clive pas avec un biais de séquence 5' mais tout comme pour les piRNAs germinaux, un biais de séquence 1U a été montré pour les piRNAs folliculaires. Il est donc possible que des cofacteurs soient responsables de ce biais ou que seuls les piRNAs prématures commençant par un uracile soient sélectionnés et chargés par Piwi. Il y aurait donc une sélection des piRNAs 1U par Piwi. Piwi va donc prendre en charge les piRNAs prématures qui seront dégradés en 3' par une ou des protéines inconnues puis 2'-O-méthylés par des protéines non identifiées. La présence de protéines Vreteno et Shutdown a été montrée dans les Yb body mais leurs rôles sont encore indéfinis. (Ishizu *et al.* 2012)

Le mécanisme de biogenèse des piRNAs folliculaires reste encore peu compris. Les mécanismes de biogenèse des piRNAs primaires germinaux et folliculaires montrent des facteurs communs mais également des facteurs tissus présentant une spécificité tissulaire (Czech *et al.* 2013 ; Muerdter *et al.* 2013 ; Handler *et al.* 2013). L'identification de facteurs dans ces deux tissus a fait l'objet d'études exhaustives par le biais de cribles et leur rôle dans la biogenèse et/ou la répression des ETs en cours (Czech *et al.* 2013 ; Muerdter *et al.* 2013 ; Handler *et al.* 2013). De plus, un certain nombre de protéines ont été identifiées dans la voie de silencing de piRNAs mais leur rôle exact n'est pas encore clarifié.

B/ Mécanismes de répression folliculaires des éléments transposables

Les piRNAs matures, qui sont produits dans les cellules folliculaires et chargés par Piwi vont permettre le silencing des ETs potentiellement actifs. Etant donné qu'il n'y a pas la présence des protéines Aub et AGO3 et donc de cycle Ping-Pong, il n'y aurait à priori donc pas de mécanisme de PTGS *via* le cycle Ping-Pong dans les cellules folliculaires. Il a été montré qu'il y a du PTGS dans ces cellules (Dufourt *et al.* 2011). A l'aide d'un transgène *UAS-GFP-Idefix*, il a été montré une répression de la GFP sans qu'il y ait de recrutement des marques H3K9me3, H3K27me2/3 et des protéines HP1 et PRC1 PC et PRC1 PH au niveau de ce transgène. En l'absence de la séquence d'*Idefix* sur ce même transgène, il y a expression de la GFP. Ceci montrerait que la répression de la GFP se ferait par PTGS et non par TGS. Néanmoins, il est possible que la répression de la GFP passe par des facteurs chromatinien non identifiés.

Il est probable que les ETs soient aussi réprimés par TGS. C'est ce que montrent les travaux des Sienski et ses collaborateurs en 2012. Il a été montré que la protéine Piwi est nucléaire et n'a pas d'activité de clivage (Saito *et al.* 2009 et 2010 ; **Figure 6**). Malgré tout, il a été montré que c'est par Piwi chargé par les piRNAs matures que se fait la répression (Sienski *et al.* 2012). En effet, ces auteurs ont montré en cellules OSS/OSC, culture cellules ovariennes folliculaires, qu'en contexte mutant pour *piwi*, il y a une augmentation d'ARN pol II sur les ETs actifs qui s'accompagne d'une augmentation de leur transcrits naissant par rapport au sauvage. Il a été également observé une diminution des marques H3K9me3 sur ces séquences ETs comparé au sauvage. De plus, ces travaux ont permis de montrer l'implication de Maelstrom (Mael) dans la voie de répression par les piRNAs (Findley *et al.* 2003 ; Lim *et al.* 2007). Il a été également montré qu'en contexte mutant pour *mael*, il y a une augmentation de l'ARN pol II et des transcrits naissant des ETs actifs mais pas de changement du niveau de H3K9me3, ni de changement de la quantité de piRNAs matures (Sienski *et al.* 2012). Ces résultats indiquent que Mael agirait en aval de l'action de Piwi. La protéine HP1 serait également impliquée (*cf* partie régulation transcriptionnelle par TGS en germinale) puisque HP1 colocalise avec Piwi sur chromosomes polytènes et sa déplétion conduit à une dérégulation des ETs en tissu germinale (Brower-Toland *et al.* 2007 ; Wang *et al.* 2011). Il a été alors proposé un modèle pour expliquer la régulation transcriptionnelle des ETs actifs dans les cellules folliculaires (Sienski *et al.* 2012 ; **Figure 17**).

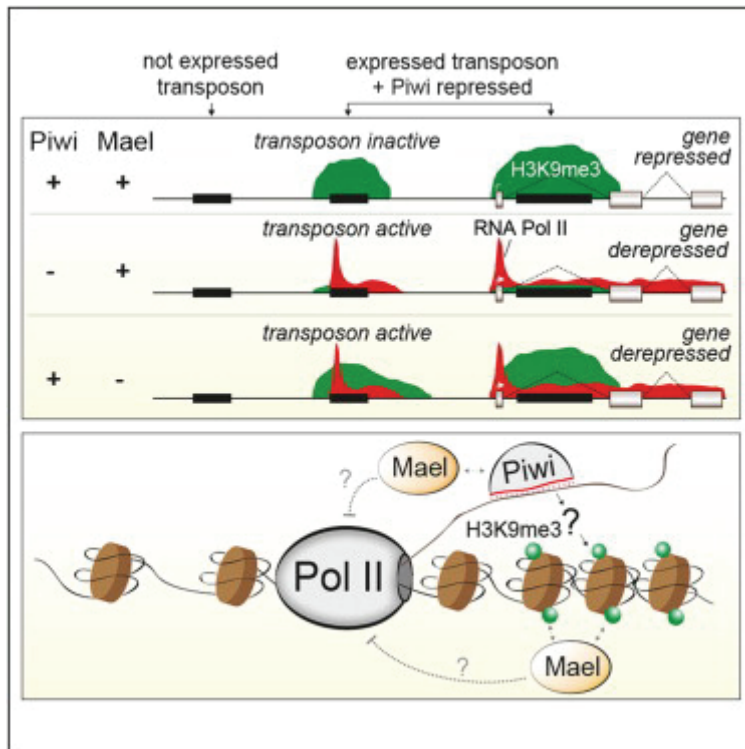


Figure 17 : Modèle de régulation transcriptionnelle des ETs dans les cellules folliculaires.

(En haut) En contexte sauvage, les auteurs présentent trois types de séquences en noir : 1- les ETs non-exprimés et non sensibles à Piwi qui ne montrent ni la présence de H3K9me3, ni celle de l'ARN pol II ; 2- les ETs sensibles à Piwi, qui montrent une accumulation de la marque H3K9me3 au niveau des séquences d'ETs en noir qui sont alors inactifs ; 3- des gènes contenant des séquences ETs en noir qui sont alors déréprimés et montrant également la présence

de la marque H3K9me3. En contexte mutant pour Piwi uniquement, il n'y a pas de changement pour les ETs insensibles à Piwi. Par contre, il y a une diminution de la marque H3K9me3 au niveau des deux autres catégories ainsi qu'une augmentation de l'ARN pol II conduisant à une activation des ETs sensible à Piwi et une expression des gènes contenant des fragments d'ETs. En contexte mutant pour *mael* uniquement, il n'y a pas de changement pour la première catégorie d'ETs insensibles à Piwi. Par contre, il y a présence d'H3K9me3 et d'ARN Pol II au niveau des séquences d'ETs sensibles à Piwi et des gènes contenant un fragment d'ETs conduisant à leurs expressions. (En bas) Modèle pour le silencing des ETs. Les transcrits naissants des ETs pourraient être reconnus *via* l'homologie de séquences par les piRNAs chargés par Piwi. Ceci permettrait le recrutement des marques H3K9me3 favorisant la répression des ETs mais de façon incomplète. Il y aurait alors intervention de Mael, soit par le biais de Piwi, soit par le biais de marques H3K9me3 qui bloquerait le recrutement de l'ARN pol II et donc permettrait le silencing des ETs et des gènes contenant des fragments d'ETs. (Sienski *et al.* 2012)

Ce modèle propose qu'il y ait recrutement des piRNAs chargés par Piwi au niveau des transcrits naissant qui permettrait alors le dépôt de marques H3K9me3. La protéine Mael serait alors à son tour recrutée et bloquerait le recrutement de l'ARN Pol II. Il y aurait inhibition de la transcription des ETs. Cette étude a été menée dans les cellules OSC/OSS mais il est possible qu'une telle régulation transcriptionnelle se réalise également dans les cellules germinales puisque Piwi y est également nucléaire ainsi que Mael. Toute la difficulté de l'analyse de cette voie de silencing par les piRNAs est que de nombreuses études sont menées sur ovaires entiers, ce qui pose le problème du mélange des deux types cellulaires. Il est donc difficile de conclure sur les résultats. Il existe des lignées de cellules folliculaires mais pas de cellules germinales provenant de la drosophile. Plusieurs équipes ont essayé de

séparer les cellules par FACS car ces cellules folliculaires sont très petites comparées aux cellules germinales. Néanmoins, le fait que les 16 cellules germinales soient reliées par des ponts cytoplasmiques ne rend pas aisée cette expérience.

La régulation des ETs dans les tissus folliculaires passerait donc par une régulation transcriptionnelle ou TGS principalement. Un papier récent de Mohn et ses collaborateurs en 2014 a permis de montrer l'implication des piRNAs chargés par Piwi dans le maintien des clusters de piRNAs dans le cas de clusters bidirectionnels. Comme il l'a été discuté pour les clusters unidirectionnels germinaux, les clusters unidirectionnels somatiques pourraient faire l'objet d'une boucle de rétrocontrôle positif par les piRNAs chargés par Piwi par le biais du maintien des marques H3K9me3 et HP1 au niveau de ces clusters. Néanmoins, il est possible que la nature hétérochromatique constitutive des régions où sont localisés les clusters de piRNAs ne requiert pas l'intervention des piRNAs ni de Piwi.

Nous venons de voir que les ETs dans les tissus ovariens somatiques et germinaux sont régulés par une catégorie des petits ARNs, les piRNAs. Mais qu'en est-il de la régulation des ETs dans les tissus somatiques non ovariens ?

C/ Biogenèse des siRNAs : petits ARNs non-codants dans les tissus somatiques

En 1998, Fire et Mello ont montré l'existence d'une régulation *via* dsRNA (pour « *double strand RNAs* ») ou ARN interférent (RNAi) ainsi que la description du mécanisme chez *C.elegans* (Fire *et al.* 1998). Ils ont montré que l'injection d'ARN double brins pouvait induire une répression faible d'un ARNm complémentaire et qu'elle est également maintenue dans la descendance ce qui suggère un mécanisme d'amplification des ARNs interférents. Ces travaux leur ont valu le Prix Nobel de Physiologie et de Médecine en 2006. De nombreuses études menées dans différents organismes ont ensuite montré l'existence d'une régulation *via* les dsRNA appelés siRNA provenant d'agents exogènes comme des infections virales, des transgènes répétés inversés ou de longs dsRNAs injectés (Guilidyal and Zamore 2009). Ce mécanisme permettrait la protection du génome contre l'entrée d'ARN exogène double brins puisque ces dsRNAs initient la production de petits ARNs (siRNA) qui permettent par la suite la dégradation de tout transcrit qui leur est complémentaire (Hamilton *et al.* 1999 ; Zamore *et al.* 2000 ; Hammond *et al.* 2000).

Ensuite, il a été montré chez les plantes et *C.elegans*, qu'il pouvait y avoir une production endogène de siRNAs à partir de régions contenant des ETs ou des séquences très répétées (Hamilton *et al.* 2002 ; Ambros *et al.* 2003 ; Zilberman *et al.* 2003). Ces siRNAs endogènes sont plus généralement appelés endo-siRNA et permettent la régulation de gènes et des ETs.

C'est sur la biogenèse des siRNAs endogènes chez la drosophile et leur mode répression que je me focaliserai par la suite. Tout comme les piRNAs, les endo-siRNAs proviennent de locis de production qui peuvent être, entre autres, des locis bidirectionnels pouvant être composés d'ETs, des transcrits chevauchant complémentaires ou encore des ARNs structurés en tige-boucle (Ghildiyal *et al.* 2008 ; Okamura *et al.* 2008 ; Okumara, Balla *et al.* 2008 ; Chung *et al.* 2008 ; Czech *et al.* 2008 ; **Figure 18 et 19**). La production des siRNAs requiert l'intervention des protéines DICER, une *dsRNA-spécifique RNase III* de la famille des Ribonucléases qui a une activité endonucléase très conservée dans le règne animal et végétal. Il y a deux protéines Dicer chez la drosophile : Dicer-1 impliquée dans la biogenèse des miRNAs et Dicer-2 impliquée dans la biogenèse des siRNAs (Lee *et al.* 2004). Dicer-2 va permettre de générer des petits ARNs double-brins d'environ 21nt (Bernstein *et al.* 2001, **Figure 18**). Ensuite, les protéines de la famille ARGONAUTE et plus particulièrement AGO2 chez la drosophile, vont former le complexe appelé RISC (pour « *RNA-induced silencing complex* ») qui va prendre en charge les siRNAs double brins clivés par Dicer-2. L'un des brins appelé le brin passager sera dégradé. Le brin conservé appelé brin guide avec un groupement 5' monophosphate, servira à la reconnaissance du transcrit cible et sa dégradation par la protéine AGO2 qui possède une activité endonucléase (Hammond *et al.* 2000, Miyoshi *et al.* 2005). Le brin guide va être modifié à son extrémité 3' par la protéine Hen1, tout comme les piRNAs, et donc 2'-O-méthylé (Horwich *et al.* 2007 ; Saito *et al.* 2007). Les analyses d'immunoprécipitation de la protéine AGO2 ont montré que les siRNAs matures simples brins ont une taille exacte de 21nt, et ont une orientation sens et antisens (Czech *et al.* 2008 **Figure 20**).

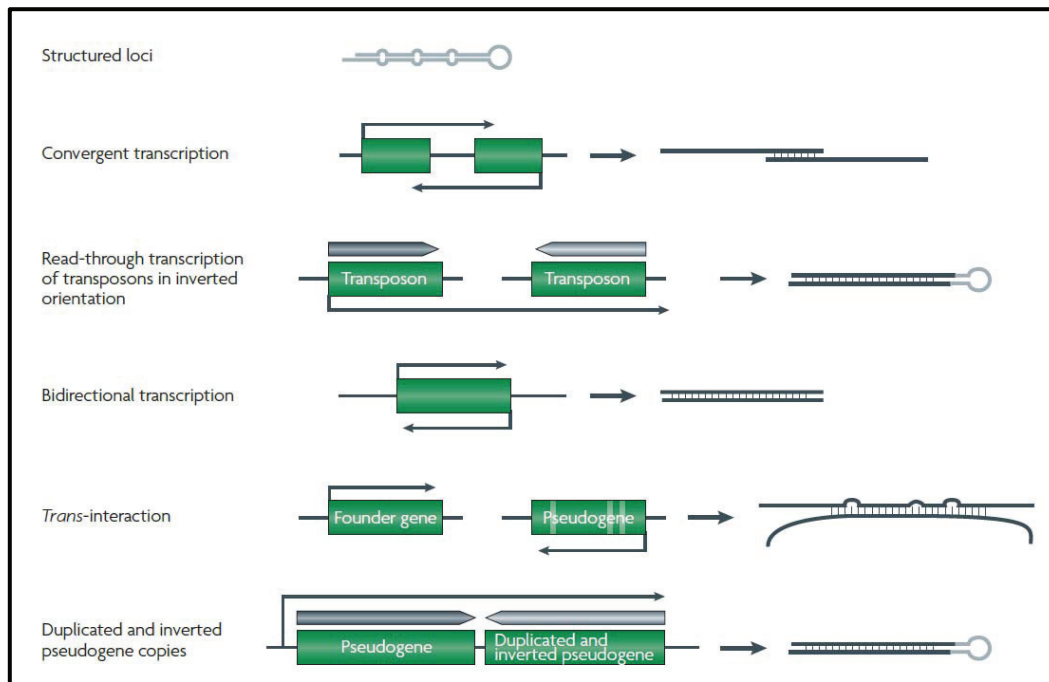


Figure 18 : Sources génomiques des précurseurs des endo-siRNAs chez la drosophile et les mammifères.

Les endo-siRNAs peuvent être issus de loci structurés issus d'une structure 2D d'un long brin d'ARN, de transcrits chevauchants complémentaires ou de loci bidirectionnels. (Ghildiyal and Zamore 2009)

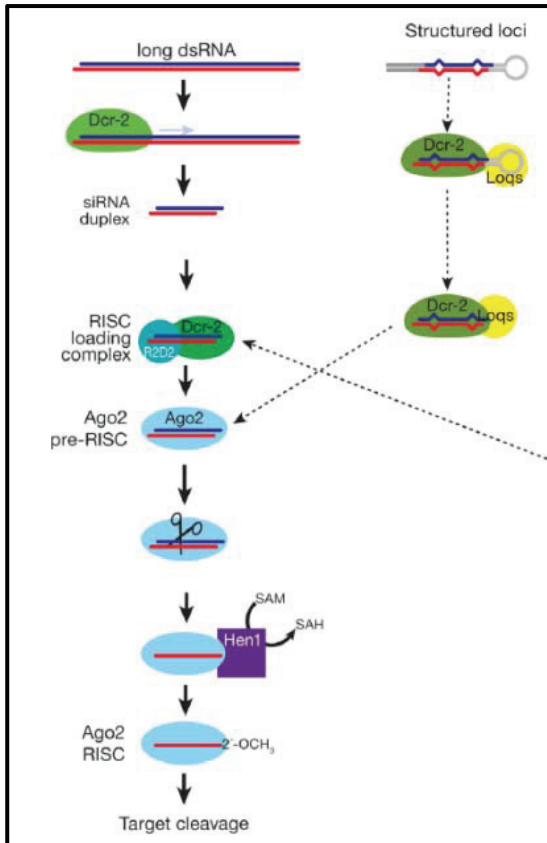


Figure 19 : Biogenèse des siRNAs chez la drosophile.

Les précurseurs de siRNAs peuvent avoir deux origines : des longs ARNs double brins ou des ARNs structurés en forme de tige-boucle (**Figure 18**). Dans le cas des précurseurs issus de longs ARNs double-brins, la protéine Dicer-2 va cliver les précurseurs de siRNAs en siRNAs double-brins à bout sortant appelés duplex siRNAs. Ces duplex siRNAs vont être pris en charge par le complexe RISC contenant entre autres la protéine AGO2. Il y aura dégradation du brin bleu appelé brin passager. La protéine Hen1 va permettre la 2'-O-méthylation de l'extrémité 3'. Une fois méthylé le siRNA chargé par le complexe RISC est mature pour le clivage de sa cible. Les précurseurs issus d'ARNs structurés en tige-boucle sont clivés par Dicer-2 mais nécessitent l'intervention de la protéine Loqs pour être pris en charge par le complexe RISC. (Ghildiyal and Zamore 2009)

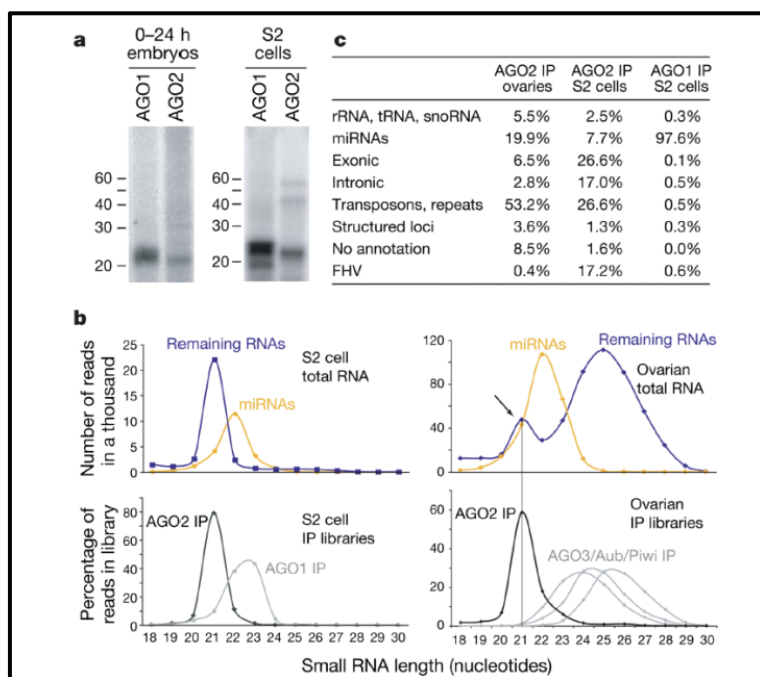


Figure 20 : AGO2 lie les endo-siRNAs.

(a) ARNs isolé à partir d'IP d'AGO1 impliqué dans la biogenèse des miRNAs et AGO2 issus d'embryon et de cellules S2 de drosophile. Dans les deux types de cellules, on constate qu'AGO2 et AGO1 sont liés à des petits ARNs d'une taille inférieure à 30nt et les petits ARNs liés à AGO2 sont légèrement plus petits que ceux liés à AGO1 (b) Profil de taille des petits ARNs isolés par des IP précédentes. En cellules S2, l'analyse des ARNs totaux montre la présence de deux populations : les miRNAs (22-

23nt) et une population de petits ARNs de 21nt. L'IP d'AGO2 et AGO1 sur les cellules S2 montre qu'AGO2 est liée aux petits ARNs de 21nt et AGO1 aux miRNAs. Dans les tissus germinaux, le profil des ARNs totaux montre la présence de trois populations de petits ARNs : ARNs de 21nt, les miRNAs et ARNs de 24-29nt : les piRNAs. L'IP d'AGO2 montre qu'AGO2 est lié aux ARNs de 21nt. Les IP d'AGO3, Aub et Piwi montrent que ces protéines sont liées aux piRNAs. (c) Annotation des petits ARNs issus des IP précédentes. L'IP d'AGO2 en ovaire et en cellules S2 montrent qu'AGO2 est liée à des petits ARNs issus majoritairement de transposons et en plus des régions exoniques dans le cas des cellules S2 contrairement à l'IP de AGO1 qui est liée majoritairement aux miRNAs. (Czech *et al.* 2008)

D/ Mécanisme de répression somatiques des éléments transposables

L'étude des endo-siRNAs somatiques chez la drosophile montre qu'ils sont entre autres impliqués dans le silencing des ETs (Ghildiyal *et al.* 2008 ; Czech *et al.* 2008 ; Chung *et al.* 2008 ; **Figure 20**). En effet, le complexe RISC chargé avec son siRNA va permettre de cibler un ARNm complémentaire. AGO2 l'un des composants du complexe RISC possède une activité endonuclease V (Rand *et al.* 2004). Cette activité endonucléase permettrait de dégrader un ARNm complémentaire ciblé par le siRNA et donc induit un mécanisme de PTGS.

Il a été montré que les ETs peuvent être réprimés par TGS par le recrutement des complexes Polycomb: PRC1 et 2 (Dufourt *et al.* 2011). En effet, il a été montré en larve par l'utilisation de transgène *UAS-GFP-Idefix* activé par un Gal4 spécifique, un recrutement des protéines de PRC1 PC et PH ainsi que de la marque H3K27 me2/3 reconnue par le complexe PRC2 au

niveau de ce transgène. Ceci montre que dans les tissus somatiques il y aurait également la formation de structures répressives au niveau des ETs impliquant les protéines du groupe polycomb.

E/ siRNAs produit par les clusters de piRNAs

Nous venons de voir qu'il existe une régulation *via* les petits ARNs dans les tissus somatiques et germinaux. Les ETs sont réprimés par la voie des siRNAs dans les tissus somatiques non ovariens et par la voie des piRNAs dans les tissus somatiques ovariens. Néanmoins, la dichotomie n'est pas si nette. En effet, il a été observé à plusieurs reprises que des siRNAs sont produits à partir des clusters de piRNAs dans les tissus somatiques (Ghildiyale *et al.* 2008) mais aussi dans les tissus ovariens (Czech *et al.* 2008, Chung *et al.* 2008 ; Malone *et al.* 2009 ; **Figure 20**). Même si la présence des siRNAs dans les tissus germinaux à plusieurs fois été observé, leur fonctionnalité n'a pas été clairement analysée.

F/ piRNAs like dans les tissus somatiques

De la même façon que des siRNAs sont produits dans les tissus ovariens, des piRNAs ont été détectés dans des tissus somatiques : dans la tête de drosophile (Yan *et al.* 2011). En effet, des petits ARNs de la taille de piRNAs et montrant une signature Ping-Pong ont été identifiés dans les têtes de drosophiles et sont appelés piRNA-like (Yan *et al.* 2011). De plus, il a été détecté la présence des protéines Aub et AGO3 dans les mushroom body (neurones $\alpha \beta$ du cerveau), ainsi qu'une dérégulation de certains ETs en contexte mutant pour Aub et AGO3 (Perrat *et al.* 2013). La biogenèse et le rôle de ces piRNAs restent encore à être élucidés. L'une des parties de ma thèse a consisté à étudier l'implication possible des piRNAs dans la tête. Je présenterai donc mes résultats et discuterai de ce phénomène dans la suite de ce manuscrit.

Au cours de ma thèse, j'ai étudié l'émergence d'un nouveau cluster de piRNAs composé de transgène *P-lacZ*. Ces transgènes ont été initialement générés pour étudier la répression de l'élément *P* dans la lignée germinale. Donc avant de vous présenter mes travaux, je vais vous présenter ce mécanisme de répression *via* les transgènes *P-lacZ* appelé TSE (pour « *Trans Silencing effect* »).

III/ Clusters de piRNAs transgéniques : outils pour l'étude de l'émergence de loci producteurs de piRNAs

L'origine de l'utilisation des transgènes *P-lacZ* est l'étude de la répression de l'élément *P*, de la famille des transposons, dans la lignée germinale femelle de drosophile (**Figure 21**). En effet, il a été constaté un phénomène de dysgénésie des hybrides lié au croisement entre individus contenant ou pas des éléments *P* (Kidwell, 1977). Lorsque des femelles sans élément *P* (provenant de souches M, collectées avant 1970) sont croisées avec des mâles contenant des éléments *P* (provenant de souches P), la descendance femelle de ce croisement présente à 29°C un phénotype de gonades atrophiées, appelé GD (pour « Gonade dysgénétique ») (Kidwell, 1977 ; Bingham *et al.* 1982 ; O'Hare *et al.* 1983). Au contraire, lors du croisement réciproque, la descendance femelle est normale (Engels 1989 ; Rio 1991 ; Ronsseray 1993). L'étude des populations naturelles a montré qu'une copie de l'élément *P* insérée dans l'hétérochromatine subtélomérique (TAS) du chromosome *X* est capable de réprimer les autres copies actives de l'élément *P* présentes dans le génome (Ronsseray *et al.* 1991 ; Ronsseray *et al.* 1996 ; Marin *et al.* 2000 ; Simmons *et al.* 2004). Cette capacité de répression de l'élément *P* transmis maternellement est appelée cytotype P et est spécifique de la lignée germinale (Engels 1979 -1989, Ronsseray *et al.* 1991-1996 ; Marin *et al.* 2000). Le phénotype de stérilité GD observé chez les femelles dysgénétiques est dû à une transposition des éléments *P*. Dans le cas des femelles non dysgénétiques, l'élément *P* télomérique apporté maternellement suffit à réprimer les copies actives des éléments *P* apportées paternellement (**Figure 22A**). Il n'y a donc pas de transposition de l'élément *P*.

Bien avant la mise en évidence de la voie de silencing par les piRNAs, il était déjà décrit des cas d'effets transgénérationnels sur le long terme pour la transmission ou la mise en place des capacités répressives des ETs.

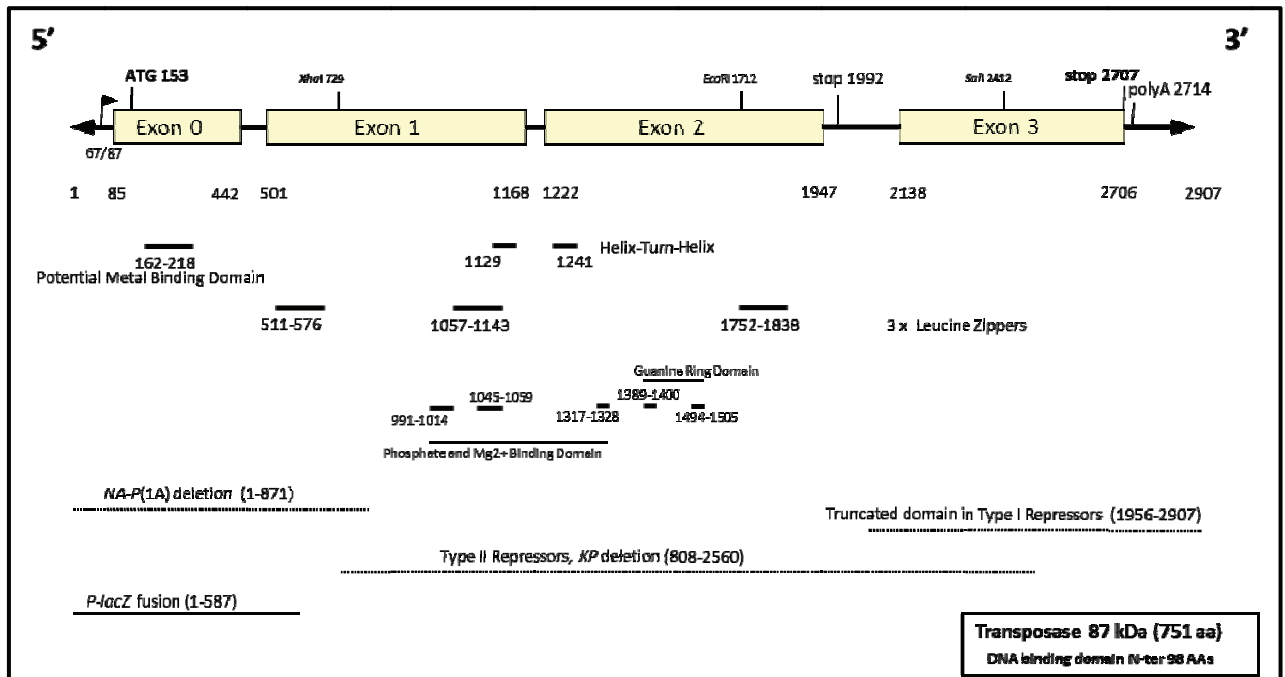


Figure 21 : Structure de l'élément *P* chez *Drosophila melanogaster*.

L'élément *P* est composé de 4 exons. Les traits noirs en gras représentent les localisations des domaines et structures identifiées. Les traits pointillés représentent les délétions importantes identifiées dans la formation d'une protéine potentiellement répresseur de *P*. En effet, l'élément *P* peut produire une protéine répresseur de 66KDa qui bloque l'action de la transposase dans les tissus somatiques. Le rôle de cette protéine en contexte germinale n'est pas clair. Il s'agit d'un mécanisme de répression qui s'ajoute à celui par le small RNA silencing. Le trait noir fin représente la séquence du promoteur de *P* utilisé dans les transgènes *P-lacZ*.

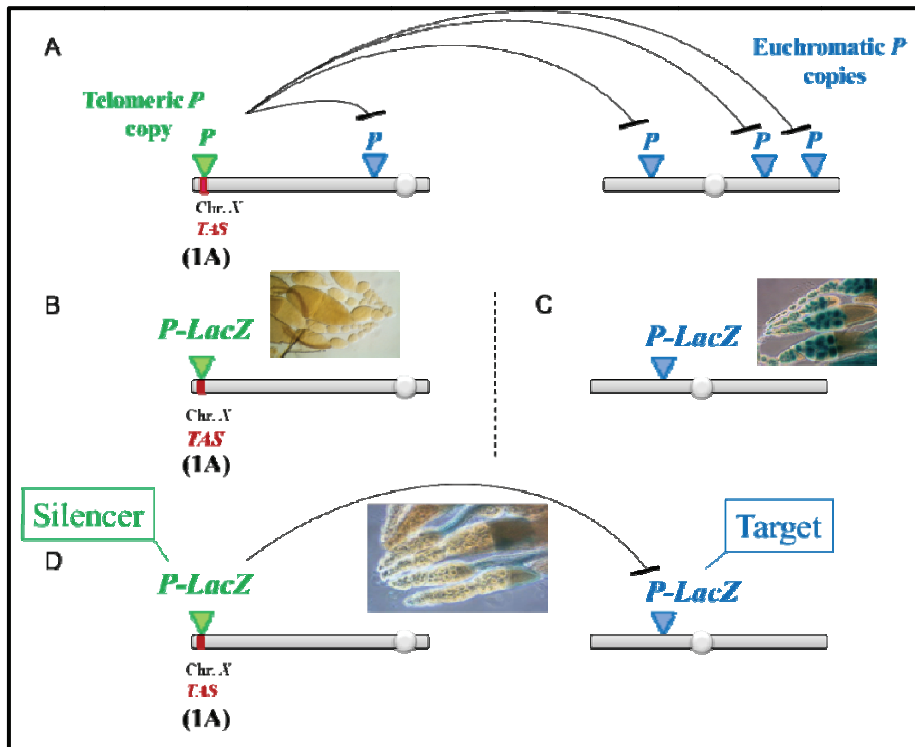


Figure 22 : Répression de l'élément *P* et d'un transgène *P-lacZ*.

Coloration *LacZ* d'ovaires de drosophile. (A) Répression de copies actives d'élément *P* euchromatiques en bleu par un élément *P* télomérique en vert inséré dans les TAS du chromosome *X* au site cytotologique 1A. (B) Expression dans les ovaires d'un transgène *P-lacZ* en vert inséré dans les TAS du chromosome *X* au site cytotologique 1A. Les ovaires sont blancs car les TAS ne contiennent pas d'enhancers actifs en ovaires. (C) Expression dans les ovaires d'un transgène *P-lacZ* en bleu inséré dans l'euchromatine à proximité d'un enhancer de la lignée germinale. Les ovaires sont bleus car le transgène s'exprime et produit la protéine β gal. (D) Répression d'un transgène *P-lacZ* euchromatique en bleu appelé cible par un transgène *P-lacZ* télomériques en vert, appelé silencer, inséré dans les TAS du chromosome *X*. Les ovaires d'une femelle contenant ces deux transgènes sont blancs. Le transgène cible ne produit pas de protéine β gal (C).

1/ Mécanisme de répression de transgènes par les piRNAs : Le *Trans-Silencing Effect* (TSE)

Les transgènes *P-lacZ* ont été utilisés dans le but de visualiser cette répression dans la lignée germinale (**Figure 21**). En effet, ces transgènes contiennent une fusion du promoteur de *P* et du gène rapporteur *LacZ* (Bellen *et al.* 1989 ; **Figure 21**). Ces transgènes sont des transgènes « *enhancers-trap* » puisqu'ils doivent être insérés à proximité de régions enhancer pour s'exprimer. Il a été ainsi obtenu une large collection de souches portant des transgènes *P-lacZ* insérés à travers le génome. Certaines de ces lignées contenaient un transgène s'exprimant dans la lignée germinale (**Figure 22C**) et d'autres contenaient un transgène dans les TAS et notamment inséré dans les TAS du chromosome *X* tout comme l'élément *P*

régulateur télomérique (**Figure 22B**). Le transgène *P-lacZ* inséré dans les TAS ne présente pas de coloration LacZ, montrant que les TAS sont dépourvus d'enhancer fonctionnant en lignée germinale femelle. Il a ensuite été testé si une femelle contenant un transgène *P-lacZ* dans les TAS pouvait réprimer un transgène *P-lacZ* inséré dans l'euchromatine apporté par le mâle. La coloration des ovaires des femelles issues de ce croisement montre l'existence d'une répression du transgène *P-lacZ* cible par le transgène télomérique puisque les ovaires sont blancs alors que lorsque le transgène *P-lacZ* euchromatique est seul, les ovaires sont bleus (Roche et Rio 1998 ; Ronsseray *et al.* 2003 ; **Figure 22D et C**). Ils ont donc montré qu'un transgène *P-lacZ* inséré dans les TAS pouvait réprimer un transgène *P-lacZ* homologue dans l'euchromatine. Ce mécanisme de répression avec les transgènes a été appelé TSE ou *Trans Silencing Effect* (Roche et Rio 1998). Le TSE est donc un mécanisme de répression dans lequel un transgène hétérochromatique, appelé silencer, va réprimer en *trans* un transgène euchromatique homologue appelé cible. Les propriétés fonctionnelles et moléculaires ont été étudiées par mon équipe.

2/ Propriétés génétiques et moléculaires du TSE

- Spécificité tissulaire

Tout d'abord, l'analyse du TSE a montré que cette répression apparaît spécifique de la lignée germinale femelle (Josse *et al.* 2008). L'utilisation de transgènes cibles s'exprimant d'une part chez l'adulte (dans le cerveau, les corps gras, les glandes salivaires) et d'autre part dans les larves (dans les disques d'ailes, d'œil et de pattes), montre qu'il n'y a pas de répression de ces transgènes dans ces différents tissus somatiques. De plus, un transgène cible s'exprimant dans les testicules (lignée germinale mâle) ne subit pas non plus de répression. Il a également été montré qu'un transgène silencer n'était pas capable de réprimer un transgène cible ne s'exprimant que dans les cellules folliculaires (Josse *et al.* 2008). Ceci signifie que cette répression est spécifique de la lignée germinale et plus particulièrement de la lignée germinale femelle.

- Effet maternel

Il a également été montré par l'équipe que cette répression présente un effet maternel tout comme la répression de l'élément *P* (Ronsseray *et al.* 2001). En effet, lorsque le transgène silencer est apporté maternellement, il y a répression du transgène cible (**Figure**

22D et Figure 23D) alors que lors du croisement réciproque, le transgène cible s'exprime normalement. Cela indique qu'il y a une composante cytoplasmique nécessaire à la répression (Josse et al. 2007). En effet, par génétique, l'équipe a montré qu'il fallait deux composantes pour faire le TSE : une composante cytoplasmique apportée maternellement et une composante chromosomique apportée maternellement ou paternellement (Figure 23). De plus, cette expérience a montré que ces deux composantes pouvaient être apportées séparément. Cette notion de composante cytoplasmique apportée maternellement soulève l'idée que la femelle doit déposer un constituant cytoplasmique dans l'embryon.

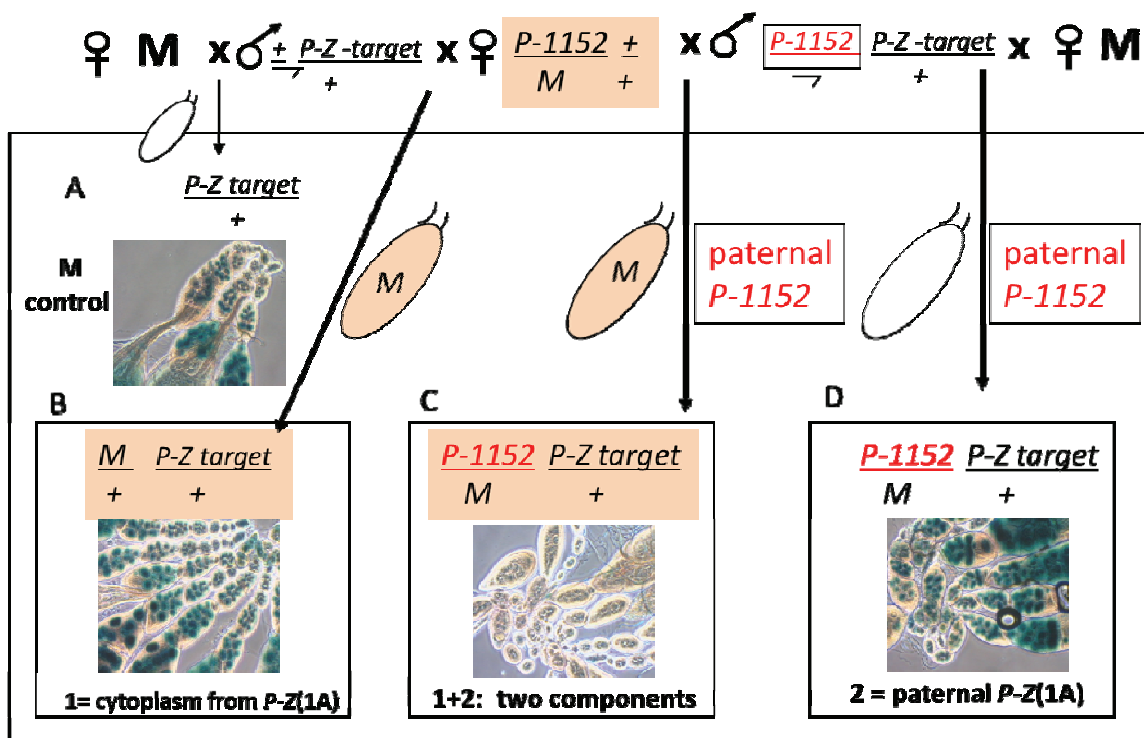


Figure 23 : Le TSE requiert deux composantes : une composante chromosomique (le transgène lui-même) et une composante cytoplasmique.

Coloration *LacZ* d'ovaires de drosophile. (A) Contrôle d'expression de cible. Des femelles M (sans élément *P* et sans transgène *P-lacZ*) sont croisées avec des mâles portant un transgène *P-lacZ* cible. Le transgène cible s'exprime dans les ovaires qui montrent une coloration bleue. (B) Des femelles contenant un transgène silencer appelé *P-1152* sont croisées par des mâles contenant un transgène cible. La descendance ayant reçu uniquement le transgène cible et la composante cytoplasmique des femelles en rose montre une coloration bleue des ovaires. Donc la composante cytoplasmique seule n'est pas suffisante pour réprimer le transgène cible. (C) Les mêmes femelles contenant un transgène silencer sont croisées avec des mâles contenant à la fois le transgène cible et un transgène silencer. La descendance ayant reçu paternellement le transgène silencer et la composante cytoplasmique maternellement en rose montre une coloration blanche des ovaires. Cela signifie que même si le transgène silencer a été reçu paternellement, la composante cytoplasmique a permis la mise en place de la répression. (D) Le croisement de femelles M avec les mâles contenant les transgènes cible et silencer montre que la descendance ayant reçu paternellement les deux transgènes cible et

silencer et maternellement une composante cytoplasmique de souche M montre une coloration bleue des ovaires. Cela signifie que le transgène silencer apporté paternellement n'est pas capable de réprimer un transgène cible.

- phénomène de variégation

L'analyse du TSE a montré également un phénomène appelé variégation (Josse *et al.* 2007 ; **Figure 24**). Il s'agit d'une répression tout ou rien à distribution aléatoire au sein d'un tissu. En effet, il a été montré que dans certains cas, le croisement de femelles possédant un transgène silencer avec des mâles possédant un transgène cible pouvait générer une descendance dont la coloration des ovaires montrait une alternance « aléatoire » de chambres bleues et blanches. Actuellement Pauline Marie, étudiante en thèse dans l'équipe sous la direction d'Antoine Boivin, analyse ce phénomène afin d'identifier le mécanisme moléculaire sous-jacent. La principale question que soulève la variégation est : la variégation résulte-t-elle d'une décision prise précocement au cours du développement et mémorisée ensuite jusqu'à l'ovaire adulte ?



Figure 24 : Premier cas de phénomène de variégation dans la lignée germinale chez la drosophile.

Ovaires de la descendance d'un croisement entre des femelles portant un transgène silencer télomérique et des mâles portant un transgène cible spécifique de la lignée germinale.

Le fait que le TSE présente un effet maternel et nécessite une composante cytoplasmique oriente le TSE vers un mécanisme de répression passant par les petits ARNs.

- Régulation par les piRNAs

Le TSE suggère que cette répression passe par des petits ARNs car le transgène *P-lacZ* télomérique ne peut pas coder un répresseur de *P*. Les nombreux travaux faits sur les piRNAs, notamment sur le mécanisme de biogenèse des piRNAs ont renforcé cette idée. L'équipe a testé l'effet de mutants de gènes impliqués dans cette voie : *aub*, *armi*, *piwi*, *zucchini* et *squash* (Josse *et al.* 2007 ; Todeschini *et al.* 2010). Ces expériences ont montré qu'en contexte

mutant pour tous ces gènes, le TSE est très fortement affecté. C'est-à-dire que le transgène silencer télomérique n'est plus capable de réprimer un transgène cible. De plus, une analyse de séquençage à haut débit des petits ARNs ovariens d'une lignée contenant un transgène silencer (*P-1152*) a été faite et montre la présence de deux types de petits ARNs associés au transgène *P-1152* des siRNAs et des piRNAs (de Vanssay *et al.* 2012).

Comme je l'ai décrit précédemment l'un des mécanismes de répression des piRNAs serait le TGS. En 2007, l'équipe a proposé l'implication de la formation d'hétérochromatine dans la répression des transgènes cibles *via* les petits ARNs par TGS (Josse *et al.* 2007). En effet, il a été testé l'état du TSE en contexte mutant pour Su(var)205 qui code pour HP1 (composante de l'hétérochromatine) et Su(var)3-7 qui code pour une autre protéine composante de l'hétérochromatine centromérique et télomérique. L'ensemble de ces mutants montre que le TSE est partiellement affecté. Tous ces résultats sont autant d'arguments allant dans le sens que la répression des transgènes *P-lacZ*, donc de l'élément *P*, passe par la voie de silencing des piRNAs. Les transcrits des transgènes cibles peuvent être dégradés par PTGS par le Ping-Pong mais il pourrait également y avoir une répression transcriptionnelle (TGS). En effet, il a été montré par ChIP Q-PCR qu'en présence d'un transgène silencer télomérique, il y a réduction de la marque H3K4me2/3 et de l'ARN polII alors qu'il y a une augmentation des marques H3K9me2 et H3K9me3 ainsi que de HP1 sur un transgène cible (Le Thomas *et al.* 2013).

Remarques : Nous venons de voir que le TSE passait par la voie des piRNAs. Or, les piRNAs sont présents dans les cellules germinales mais également les tissus somatiques ovariens. Dans le cas du TSE, il a été observé qu'il n'y avait pas de répression dans les cellules folliculaires. Il est possible que le fait que le TSE ne soit pas détecté dans les cellules folliculaires, résulte du fait que les clusters dans lesquels sont insérés les transgènes télomériques ne sont pas actifs dans les cellules folliculaires. Mais cela ne signifie pas que le mécanisme de répression utilisé par le TSE n'existe pas dans les tissus folliculaires. En effet, les transgènes silenciers testés sont principalement insérés dans les TAS. Or, les TAS sont considérés comme des clusters bidirectionnels. Il a été montré que le cluster 42AB, par exemple, ne s'exprime pas dans les cellules folliculaires (**Figure 7**). Il est possible d'envisager qu'un transgène inséré en antisens dans un cluster unidirectionnel comme *flamenco* (tout comme la majorité des fragments d'ETs qui le composent) s'exprimant dans les cellules folliculaires puisse réprimer un transgène cible sens. Cette expérience refléterait les conditions réelles de la voie de silencing par les piRNAs somatiques tels qu'elle a été

décrite. Donc, le fait de n'avoir pas détecté de TSE dans les cellules folliculaires dans les conditions testées ne signifie pas qu'un TSE dans les cellules folliculaires soit en soi impossible.

Le TSE permet donc de visualiser la répression de séquences contenant des fragments d'ETs par les piRNAs. Il est possible de considérer le TSE comme un « *read-out* » de la répression des ETs par les piRNAs en contexte germinale. L'étude du TSE a permis ensuite de mettre en évidence un phénomène de paramutation. Ce sujet fait l'objet principal de ma thèse.

IV/ Phénomène de paramutation lié à l'effet maternel du TSE

1/ Définition de la paramutation

La paramutation constitue une conversion épigénétique stable. Pour rappel, l'épigénétique désigne les mécanismes qui permettent la transmission de changements dans l'expression des gènes à travers la mitose et/ou la méiose alors que l'agent ayant induit le changement a disparu, et qui ne résultent pas de modifications de la séquence d'ADN.

Il est alors possible de définir des épi-allèles comme ayant la même séquence d'ADN mais ne présentant pas les mêmes propriétés épigénétiques et conduisant donc à des profils d'expression différents. Lors de la paramutation, il y a un transfert des propriétés épigénétiques d'un épi-allèle à l'autre (Brink *et al.* 1956 ; **Figure 25**). L'épi-allèle qui donne ses propriétés épigénétiques est dit « paramutagénique ». Avant la paramutation, l'épi-allèle qui reçoit les propriétés épigénétiques est dit « paramutable ». Après la paramutation, l'épi-allèle qui a reçu les propriétés épigénétiques est dit « paramuté ». Pour parler d'une vraie paramutation, il faut que l'allèle paramuté le reste de façon stable au cours des générations suivantes et qu'il soit lui-même paramutagénique. C'est-à-dire que s'il est mis en présence d'un allèle paramutable, il le transforme à son tour en allèle paramuté. Le phénomène est donc potentiellement récurrent.

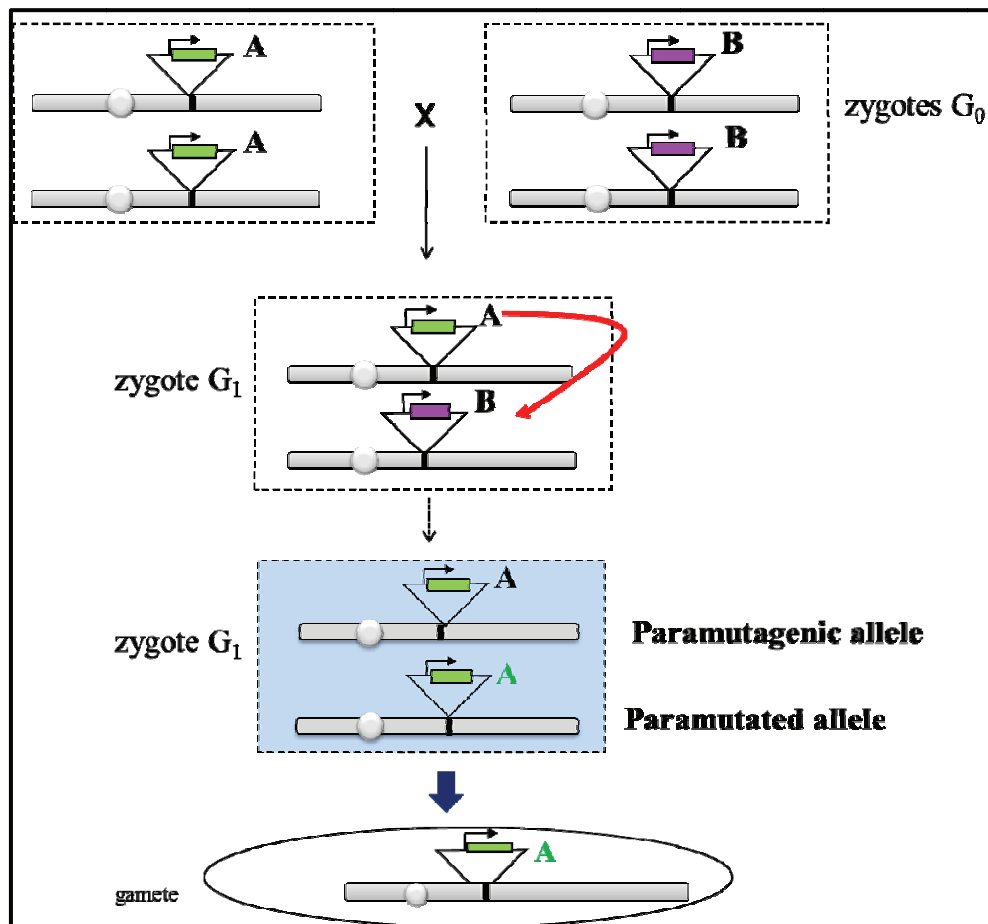


Figure 25 : Principe de la paramutation.

A en vert et B en violet sont deux épi-allèles avec une même séquence d'ADN mais possédant des propriétés épigénétiques différentes. Lorsqu'ils sont présents dans le même génome, A donne ses propriétés épigénétiques à B (B devient A vert). L'épi-allèle A qui donne ses propriétés épigénétiques est appelé paramutagénique. Le nouvel épi-allèle A vert qui a reçu les propriétés épigénétiques est dit paramuté. L'épi-allèle A vert paramuté garde ses nouvelles propriétés dans les gamètes et au cours des générations suivantes.

2/ La paramutation au locus *Booster B1* chez le maïs

Le premier cas de paramutation a été mis en évidence chez le maïs et a été décrit de façon stable uniquement dans le règne végétal (Brink *et al.* 1956). Un exemple typique de paramutation chez le maïs est celui du locus *b1*. Le locus *b1* code pour un pigment : l'anthocyanine violette (**Figure 25**). Il a été identifié deux épi-allèles différents : *B-I* et *B'*. Des plantes homozygotes pour l'épi-allèle *B-I* montrent une expression élevée du locus *b1* qui est associée à une pigmentation des tiges en violet foncé. Par contre, des plantes homozygotes pour l'épi-allèle *B'* montrent une faible expression du locus *b1* qui est associée à une pigmentation des tiges en vert clair. Le croisement de ces deux homozygotes donne une descendance hybride qui présente des tiges vertes claires. Il a été montré que ce cas de

dominance ne reflète pas une génétique classique puisque l'épi-allèle *B-I* est converti en *B'* par l'épi-allèle *B'*. Le nouvel épi-allèle paramuté est appelé *B'** (Coe 1966). Cette paramutation est stable et 100% de la descendance en G₂ est de type *B'/B'** et donc paramuté. Il a été identifié une séquence non-codante répétée en tandem à environ 100kb en amont du *start* de transcription du locus *b1* et dont le profil de méthylation des histones est différent entre les deux épi-allèles *B-I* et *B'* (Stam *et al.* 2002). En effet, les séquences répétées en tandem du locus de *B-I* montrent une chromatine plus ouverte que celles de *B'*. De plus, il a été montré l'implication de petits ARNs dans le mécanisme de paramutation. En effet, il a été montré que les séquences répétées en amont de *b1* sont transcrites dans les deux sens et pourraient être à l'origine de la production d'ARNs double brins (Alleman *et al.* 2006). Dans ce même article, il a été identifié une RdRP (pour « *RNA-dependant RNA polymerase* ») appelé MOP1 (pour « *mediator of paramutation 1* ») nécessaire à la paramutation de *B-I* en *B'** (Alleman *et al.* 2006 ; Dorweiller *et al.* 2000). La transcription bidirectionnelle des séquences répétées est nécessaire mais pas suffisante pour induire le phénomène. La paramutation dépend du nombre de répétitions en tandem. En effet, il faut 7 copies de séquences répétées en tandem pour que la paramutation soit stable au cours des générations (Alleman *et al.* 2006). Il a finalement été montré la production de siRNAs de 25nt issus des séquences répétées dans les 3 génotypes : les deux homozygotes *B-I* et *B'* ainsi que l'hybride *B-I/B'* (Siderenko *et al.* 2009 ; Arteago-Vasquez *et al.* 2010). Ces petits ARNs sont sensibles à un mutant MOP1 (Arteago-Vasquez *et al.* 2010). Cela signifie que les siRNAs sont requis dans l'établissement de la paramutation mais pas suffisants. Par contre, une fois la paramutation établie, ils ne sont plus requis pour son maintien au cours des générations. De nombreuses interrogations persistent. Pourquoi les répétitions en tandem sont-elles requises ? Comment l'épi-allèle *B-I* peut-il être fortement actif transcriptionnellement alors qu'il produit des siRNAs ? Quelle est la nature des molécules ou marques qui sont héritées lors de la paramutation et de son maintien ?

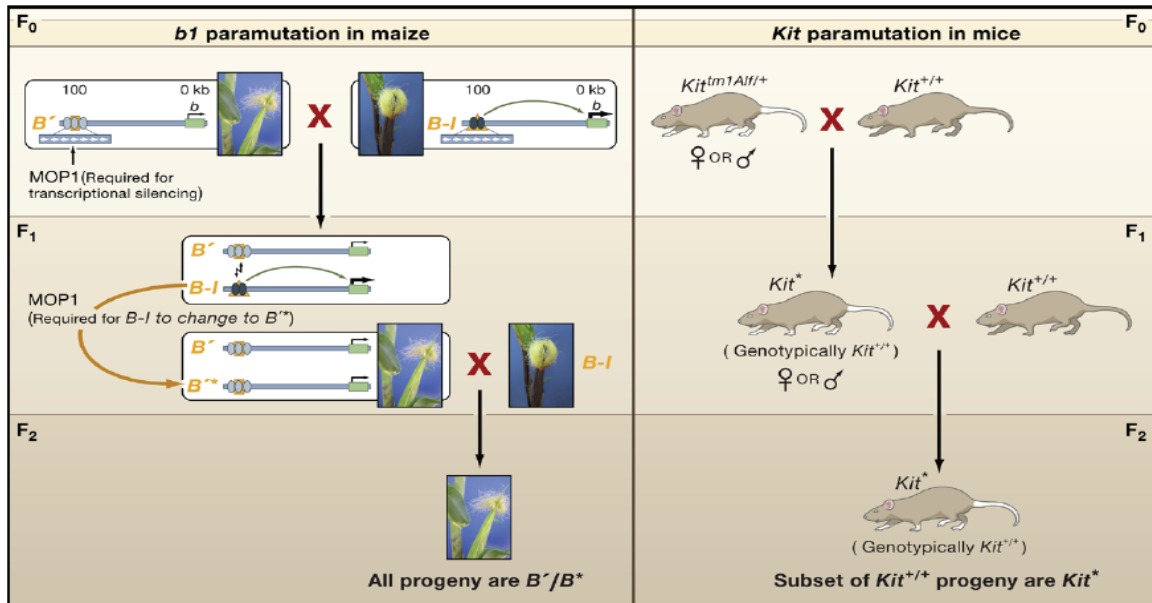


Figure 25 : Paramutation chez le maïs et la souris.

(A gauche) Paramutation chez le maïs au locus *b1* qui code pour un pigment, l'anthocyanine violette. Il existe deux épi-allèles de ce locus B' et $B-I$. Ces deux épi-allèles ont la même séquence d'ADN mais n'ont pas le même profil d'expression. Des individus homozygotes pour B' montrent une répression de *b1* et ont des tiges vertes claires alors que des individus homozygotes pour $B-I$ montrent une expression élevée de *b1* et ont des tiges violet foncé. Il a été identifié une séquence de 7 répétitions de 853pb en tandem, symbolisée en orange, en amont du start de transcription de *b1*. L'état chromatinien de cette séquence répétée est différent entre $B-I$ et B' . Dans le cas de B' , il y a une méthylation des histones de la séquence répétée responsable de la répression de *b1* symbolisée par des ronds clairs. Il a été montré que la protéine MOP1 est requise dans l'état méthylé de B' . Le croisement de deux homozygotes B' et $B-I$ donne une descendance hétérozygote avec toujours des tiges vertes claires. Cela signifie que l'épi-allèle B' a donné ses propriétés épigénétiques à $B-I$ et $B-I$ est donc devenue B'' . Si des hétérozygotes B'/B'' sont croisés avec des homozygotes $B-I$, toute la descendance F_2 aura des tiges claires et sera B'/B'' .

(À droite) Paramutation chez la souris au locus *Kit* qui code pour un récepteur de thyrosine kinase. Ces souris hétérozygote pour l'allèle Kit^{tm1Alf} , qui ne produit pas de protéine, présentent un phénotype de queue et pattes blanches. Lorsque des souris hétérozygotes sont croisées avec le sexe opposé sauvage, certains descendants sont génétiquement sauvages ($kit^{+/+}$) mais montrent le phénotype de queue et pattes blanches. Ils montrent une réduction de la quantité d'ARNm de *Kit* tout comme l'hétérozygote. Ces descendants avec un phénotype paramuté sont désignés Kit^* . La fréquence de Kit^* n'est pas de 100% et varie selon le croisement. De même que l'étendue de la zone blanche est variable d'un individu à l'autre. Bien le phénotype de queue et pattes blanches soit encore observé dans la seconde génération après out-cross de Kit^* avec $Kit^{+/+}$, la fréquence est faible et le phénotype a disparu après quatre générations. (Chandler 2007)

3/ La paramutation au locus *Kit* chez la souris

Il a été décrit un cas de paramutation instable chez la souris (Rassoulzadegan *et al.* 2006, **Figure 25**). Cette paramutation serait liée au locus *Kit* qui code pour un récepteur de thyrosine kinase impliqué dans la mélanogenèse, la différenciation des cellules germinales et l'hématopoïèse. Le mutant perte de fonction *Kit*^{tm1Alf} provoque la mort précoce des souriceaux après la naissance. Mais des souris hétérozygotes portant l'allèle *Kit*^{tm1Alf} et un allèle sauvage sont viables et présentent un phénotype particulier : les souris femelle et mâle ont la queue blanche et les pattes blanches. Le croisement d'individus hétérozygote et sauvage produit entre autres des descendants génétiquement sauvages (*Kit*^{+/+}) mais qui présentent un phénotype de queue blanche et pattes blanches. Il a été observé une réduction de la quantité d'ARNm de *Kit* pour cette descendance *Kit*^{+/+} comme pour les hétérozygotes. Ces descendants sont appelés *Kit**. L'équipe de Rassoulzadegan a proposé l'intervention de molécules d'ARNs qui seraient transmis à la descendance *via* les gamètes et qui seraient responsables de la dégradation des ARNm de *Kit*. Il a été montré, toujours par cette équipe, une accumulation d'ARNs non polyadénylés et de tailles anormales dans les hétérozygotes. Pour tester l'implication des ARNs dans le phénotype de paramutation, ils ont injecté des ARNs issus d'hétérozygotes dans des embryons sauvages et ont observé une haute fréquence d'individus paramutés. Ils ont répétés l'expérience avec des miRNAs de *Kit* et ont observé également l'apparition d'une proportion d'individus aux pattes blanches. Ils en concluent que l'héritage épigénétique serait associé à un transfert d'ARNs. Par contre, il n'a pas été identifié de profils chromatinien différents au niveau du promoteur de *Kit* entre les deux allèles muté et sauvage. Néanmoins, les régions à proximité n'ont pas été explorées. Cette paramutation est instable au cours des générations et toute la descendance n'est pas affectée. De plus, l'entendue du phénotype varie d'un individu à l'autre.

Jusqu'à nos travaux, aucune paramutation stable n'avait été décrite dans le monde animal.

V/ Présentation du sujet

L'équipe de « Répression épigénétique et ADN mobile », dans laquelle j'ai réalisé mes travaux de thèse, étudie des phénomènes de régulation à transmission épigénétique à travers la méiose. En effet, le travail initié à partir de l'analyse des mécanismes de régulation des ETs a évolué vers l'analyse des phénomènes d'héritabilité trans-générationnelle associés aux régulations médiées par les piRNAs. Nous essayons, en particulier, de comprendre comment fonctionnent ces clusters de piRNAs par le biais de loci transgéniques uniques dans le génome dont nous connaissons la localisation et la séquence. Nous pouvons donc isoler les piRNAs produits par ces transgènes et connaître leur état fonctionnel en apportant au système un transgène cible qui sera réprimé ou pas par les piRNAs produits par le transgène silencier. Nous avons, enfin et surtout, des témoins négatifs c'est-à-dire des souches dépourvues de ces loci ce qui permet d'analyser la génétique de ce système.

Mon travail de thèse s'est déroulé en 3 phases.

A mon arrivée au laboratoire un phénomène de paramutation venait d'être découvert avec un cluster hétérochromatique de transgènes. Ce cluster peut exister en deux états selon les lignées : dans certaines lignées le cluster est dans un état « actif », c'est-à-dire qu'il peut effectuer un *trans*-silencing en lignée germinale femelle. Dans d'autres lignées il n'est pas actif. Ces deux types de lignées sont totalement stables au cours des générations. Nous avons converti une lignée possédant un cluster de transgène non-silencer en cluster silencier par transmission maternelle de piRNAs homologues aux transgènes. Celle-ci représentait le premier cas stable de paramutation animale. J'ai donc, dans un premier temps, participé à la première phase de caractérisation de ce phénomène de paramutation menée en vue d'une première publication. J'ai mené l'analyse de l'effet d'un mutant (*aubergine*) sur les propriétés d'un locus paramuté.

Nous avons donc activé *de novo* un cluster de piRNAs. Or très peu de choses sont connues sur les mécanismes d'émergence des clusters de piRNAs. Dans un second temps, j'ai développé l'analyse des propriétés fonctionnelles et moléculaires de ce nouveau cluster de piRNAs par une approche gène candidat et par l'analyse des petits ARNs produits par ce cluster. Ce cluster produit deux sortes de petits ARNs : des piRNAs et des siRNAs. J'ai analysé le rôle fonctionnel de ces deux types de petits ARNs dans le silencing et dans la mise en place et/ou le maintien de la paramutation.

Nous avons ici paramuté un cluster non-silencer en cluster producteur de piRNAs par la transmission maternelle de piRNAs. Mais qui de l'œuf et de la poule est arrivé en premier ? Comment a émergé le premier cluster de piRNAs ? Nous nous sommes penchés sur la question en testant un certain nombre de conditions environnementales ou gains et pertes de fonction de gènes qui pourraient déstabiliser un cluster composé de séquences répétées qui produira un transcrit qui sera clivé en piRNAs. Dans une troisième partie, je vous présenterai ma participation à ce projet.

Résultats

Résultats

VI/ Emergence d'un locus producteur de piRNAs par un phénomène de paramutation

1/ Présentation des lignées *T-1* et *BX2*

Le TSE a été mis en évidence et caractérisé avec des transgènes télomériques comme transgènes silencer. Ces régions télomériques sont constituées de clusters de séquences hétérochromatiques non-codantes qui produisent de façon abondante des piRNAs. Afin de généraliser le phénomène de répression par le TSE, il a été activement recherché par le laboratoire des transgènes non-télomériques susceptibles d'être silencer (Ronsseray *et al.* 2001). Des lignées portant des clusters de transgènes *P{lacW}* avaient été générées pour l'étude des phénomènes de répétitions en tandem sur l'expression génique (Dorer and Henikoff 1994, 1997). Il a été montré que ces clusters de transgènes conduisent à une répression de l'expression du transgène dans l'œil (appelée RIGS pour « *Repeat Induced Gene Silencing* ») et forment des îlots hétérochromatiques en contexte euchromatique (Dorer and Henikoff 1994, 1997; Fanti *et al.* 1998). Ils pouvaient donc être de bons candidats pour générer une répression en *trans*. Deux des lignées, appelées *T-1* et *BX2*, ont particulièrement retenu l'attention car il s'agissait de deux lignées possédant le même cluster de 7 transgènes *P{lacW}* inséré au même locus (site cytologique 50C) sur le bras droit du chromosome 2. Il s'avère que la lignée *T-1* est issue de la lignée *BX2* suite à un traitement aux rayons X. Des réarrangements chromosomiques ont été identifiés par Dorer et Henikoff dans cette souche *T-1* dont une translocation des chromosomes 2 et 3 mais une analyse par Southern Blot a montré que la lignée *T-1* possède toujours 7 copies du transgène au même site cytologique. Les tests de répression de ces deux lignées ont montré que la lignée *T-1* est un fort silencer alors que la lignée *BX2* n'est pas silencer. Il s'agit donc de deux épi-allèles qui ont le même cluster de transgènes mais qui n'ont pas les mêmes propriétés épigénétiques car l'une a des capacités répressives contrairement à l'autre.

Il était probable que locus *T-1* produisait des piRNAs alors que le locus *BX2* n'en produisait presque pas. Serait-il possible de démarrer *BX2* par héritage maternel de piRNAs homologues ? Si tel est le cas, il s'agit alors d'une conversion épigénétique et plus particulièrement d'une paramutation. Ces travaux ont fait l'objet d'un article auquel j'ai participé (de Vanssay *et al.* 2012). En effet, Augustin a généré une lignée nommée *BX2**

paramutée qui est devenue Silencer et ce de façon stable au cours des générations (>100 générations). De plus, la lignée *BX2** s'est montrée paramutagénique, c'est-à-dire qu'elle est elle-même capable de transférer ses propriétés épigénétiques à un autre épi-allèle naïf et de le paramuter de façon stable. L'analyse des petits ARNs ovariens a confirmé que *T-1* produisait des petits ARNs homologues sur toute la longueur du transgène alors que le cluster *BX2^{naïf}* n'en produisait qu'une très faible quantité. Par contre, *BX2** produisait des petits ARNs tout comme *T-1*. L'analyse de ces petits ARNs a montré la présence de deux populations d'ARN : piRNAs de 24 à 28nt et ARNs de 21nt qui pourraient être des siRNAs.

Lors de la « Review » de cet article, il a été demandé de tester l'effet de mutants de gènes impliqués dans ces deux voies de silencing par les petits ARNs. L'un de mes projets de thèse a été de tester l'effet de mutations d'*aubergine*, gène impliqué dans la biogenèse des piRNAs secondaires, sur les propriétés répressives d'un locus paramuté. J'ai généré par recombinaison deux lignées : *BX2**, *aub^{N11}* (2 lignées indépendantes) et *BX2**, *aub^{QC42}* (2 lignées indépendantes). J'ai confirmé la présence de l'allèle *aub⁻* en effectuant des tests de stérilité puisque que cette mutation est viable et stérile à l'état homozygote. J'ai, ensuite, croisé mes lignées : *BX2**, *aub^{N11}* et *BX2**, *aub^{QC42}* avec respectivement des lignées *aub^{QC42}* ; *BQ16* (transgène cible) et *aub^{N11}* ; *BQ16* afin de tester les propriétés répressives de *BX2** en contexte mutant pour *aub*. J'ai donc obtenu des individus *BX2** hétéro-alléliques pour *aub* dans les deux sens de croisements dont j'ai testé leur stérilité. J'ai réalisé 4 réplicats biologiques et ai testé à chaque fois deux lignées indépendantes pour chaque génotype. J'ai, également, extrait les ARNs totaux ovariens des femelles *BX2** hétéro-alléliques pour *aub* (*BX2**, *aub^{QC42}/aub^{N11}*) et hétérozygote (*BX2**, *aub^{QC42}/Cy*), confirmé par des tests de stérilité, afin de faire le séquençage à haut débit des petits ARNs. Leur analyse est faite dans l'article Hermant *et al.* en préparation.

2/ Article: « Paramutation in *Drosophila* linked to emergence of piRNA-producing locus » (de Vanssay *et al.* 2012)

Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus

Augustin de Vanssay^{1†}, Anne-Laure Bougé^{2†}, Antoine Boivin¹, Catherine Hermant¹, Laure Teyssset¹, Valérie Delmarre¹, Christophe Antoniewski^{2†} & Stéphane Ronsseray¹

A paramutation is an epigenetic interaction between two alleles of a locus, through which one allele induces a heritable modification in the other allele without modifying the DNA sequence^{1,2}. The paramutated allele itself becomes paramutagenic, that is, capable of epigenetically converting a new paramutable allele. Here we describe a case of paramutation in animals showing long-term transmission over generations. We previously characterized a homology-dependent silencing mechanism referred to as the *trans*-silencing effect (TSE), involved in *P*-transposable-element repression in the germ line^{3–5}. We now show that clusters of *P*-element-derived transgenes that induce strong TSE^{6,7} can convert other homologous transgene clusters incapable of TSE into strong silencers, which transmit the acquired silencing capacity through 50 generations. The paramutation occurs without any need for chromosome pairing between the paramutagenic and the paramutated loci, and is mediated by maternal inheritance of cytoplasm carrying Piwi-interacting RNAs (piRNAs) homologous to the transgenes. The repression capacity of the paramutated locus is abolished by a loss-of-function mutation of the *aubergine* gene involved in piRNA biogenesis, but not by a loss-of-function mutation of the *Dicer-2* gene involved in siRNA production. The paramutated cluster, previously producing barely detectable levels of piRNAs, is converted into a stable, strong piRNA-producing locus by the paramutation and becomes fully paramutagenic itself. Our work provides a genetic model for the emergence of piRNA loci, as well as for RNA-mediated *trans*-generational repression of transposable elements.

Paramutations have been well described in plants^{1,2,8–12}. The best characterized is the *b1* paramutation in maize, which involves a small RNA silencing pathway^{13–15}, changes in DNA methylation levels and chromatin modifications¹⁶, and shows full penetrance and stability across generations. Paramutation-like phenomena involving microRNAs have been described in mice^{17,18}. However, long-term inheritance of a paramutation through generations has not been reported so far in animals.

In *Drosophila melanogaster*, transposition of *P* elements causes hybrid dysgenesis, a syndrome of genetic abnormalities including a high mutation rate, chromosome rearrangements and sterility^{19,20}. In natural populations, telomeric *P* elements inserted in heterochromatic telomere-associated sequences (TAS) are master sites for establishing *P*-element repression in the germ line^{21–23}. In laboratory lines (for example, *P-1152*), *P-lacZ* transgenes inserted in TAS mimics telomeric *P* elements by repressing germline expression of reporter transgenes inserted at distant euchromatic sites, through a homology-dependent silencing mechanism, TSE^{3–5,24}. TSE is strongly sensitive to mutations affecting the piRNA pathway^{5,25}. Its establishment involves both genetic and epigenetic components: a chromosomal copy of the telomeric silencer transgene must be either paternally or maternally inherited, and a cytoplasmic component containing small RNAs homologous to the transgene must be maternally inherited^{4,5}. In

addition to telomeric loci, we found that *T-1*, a tandem repeat cluster of *P-lacZ* transgenes inserted in the middle of chromosome arm 2R (50C), can also trigger a strong TSE⁷. *T-1* and other *P-lacZ* clusters inserted at the same locus (Supplementary Fig. 1) induce ectopic heterochromatin and show variegation of the *white* gene marker in the eye, a phenomenon termed repeat-induced gene silencing^{6,26}. However, *T-1* triggers strong silencing of various TSE reporter transgenes in the germ line⁷, whereas the other transgene clusters at this locus, including *BX2*, which contains the same number of transgene repeats as *T-1*, did not induce detectable TSE (Supplementary Table 1).

The epigenetic properties of *T-1* were analysed together with those of the *P-1152* telomeric silencer and the *BX2* cluster as controls. *T-1* and *P-1152* showed typical maternal transmission of TSE: strong repression occurred in the germ line of progeny when the silencer was maternally inherited (Fig. 1a), whereas weak or null repression was detected when the silencer was paternally inherited (Fig. 1b). *BX2* showed no repression capacity in these crosses. To analyse the relationship between TSE and piRNAs, we sequenced 19–29-nucleotide RNAs from ovaries of *T-1*, *P-1152* or *BX2* females (Supplementary Table 2). Abundant small RNAs matched the *T-1* sequences in the library from hemizygous females having inherited the *T-1* locus maternally (Fig. 1c), but not paternally (Fig. 1d). Among these species, the 23–28-nucleotide RNAs showed the typical ‘ping-pong’ signature of piRNA biogenesis²⁷, including a bias for a 5′ U (1U) and a strong tendency to form sense–antisense pairs with complementarity over their first ten nucleotides (Supplementary Fig. 2). In addition to piRNAs, short interfering RNAs (siRNAs) have been shown to be produced by previously characterized piRNA loci²⁸. Similarly, *T-1* produced a significant fraction of 21-nucleotide RNAs (Fig. 1c) that do not show the ping-pong signature of piRNAs and probably correspond to siRNAs (Supplementary Fig. 3a). In agreement with a previous report²⁹, small RNAs with similar features were produced by *P-1152* in hemizygous females having inherited the *P-1152* locus maternally (Fig. 1f). Homozygous *P-1152* females produced about twice as many piRNAs as these hemizygous females (Supplementary Fig. 4). Finally, only a very low level of small RNAs was produced that matched *BX2* in hemizygous females from the *BX2* line (Fig. 1e). Hence, maternal inheritance of *T-1*, as well as *P-1152*, is associated with both the production of piRNAs derived from these loci and the capacity of these loci to mediate TSE, thereby linking silencing and piRNAs in this system. We next tested epigenetic interactions between the *P-1152* telomeric silencer and *T-1*, and found that chromosomal and maternally transmitted components of *T-1* and *P-1152* can complement each other to induce TSE (Supplementary Fig. 5), consistent with the presence of piRNAs matching *P-lacZ* sequences in ovaries of both *T-1* and *P-1152* females.

To investigate possible transfer of epigenetic information between *T-1* and the inactive *BX2* locus, we crossed hemizygous *T-1* females

¹Laboratoire Biologie du Développement, UMR7622, CNRS-Université Pierre et Marie Curie, 9 quai Saint Bernard, 75005 Paris, France. ²Drosophila Genetics and Epigenetics, CNRS URA2578 - Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France. [†]Present addresses: Institut Jacques Monod, CNRS, UMR 7592, Université Diderot, Sorbonne Paris Cité, F-75205 Paris, France (A.d.V.); Drosophila Normal and Pathological Neurobiology, INSERM U661 - Institut de Génétique Fonctionnelle, 141 rue de la Cardonille, 34094 Montpellier, France (A.-L.B.); Drosophila Genetics and Epigenetics, Laboratoire Biologie du Développement, UMR7622, CNRS-Université Pierre et Marie Curie, 9 quai Saint Bernard, 75005 Paris, France (C.A.).

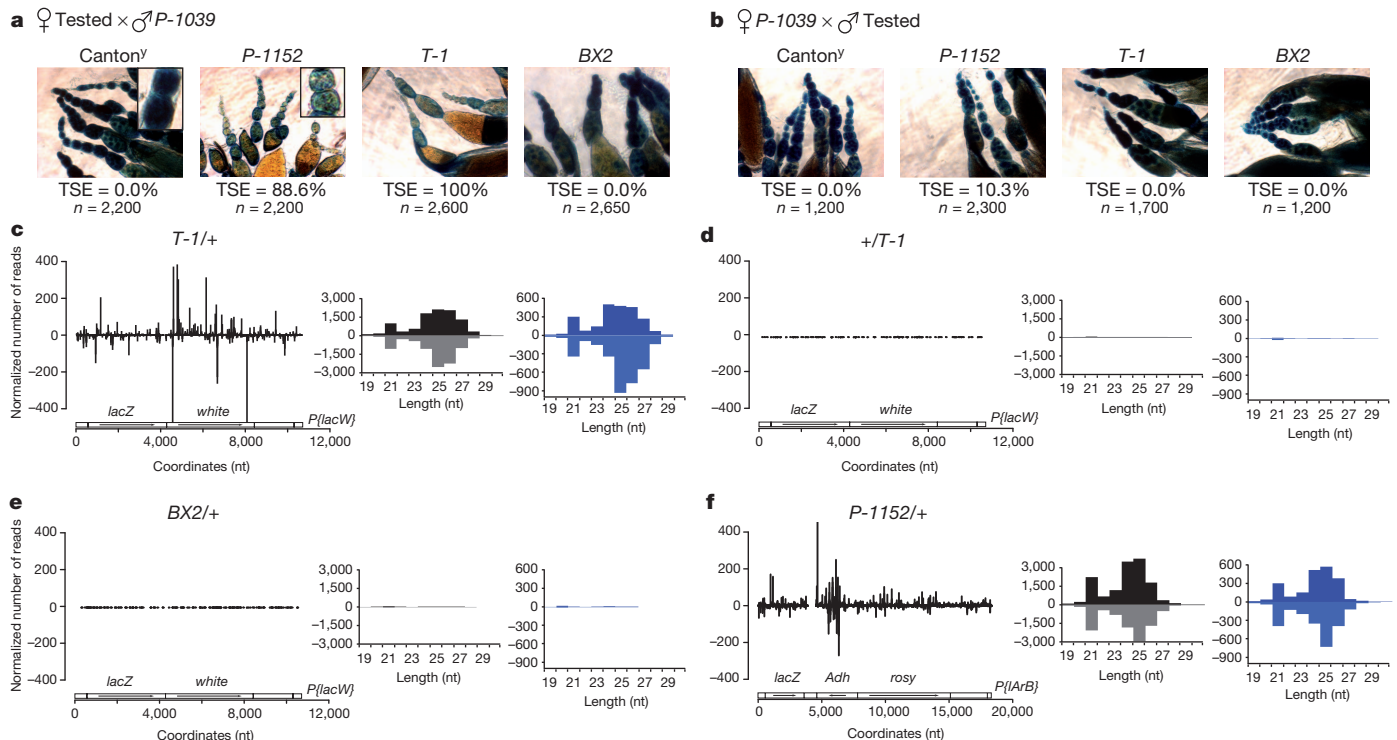


Figure 1 | Maternal inheritance of *P-1152* and *T-1* repression capacities correlates with the presence of *T-1* or *P-1152*-derived piRNAs in ovaries of female progeny. **a, b**, Maternal (**a**) and paternal (**b**) inheritance of TSE mediated by *P-1152*, *T-1* or *BX2* was tested using the *P-1039* TSE reporter transgene. *lacZ* staining of ovaries of G_1 females from the indicated crosses was performed, and TSE was expressed as the percentage of repressed egg chambers among the total number (n) of egg chambers analysed. Female and male Canton^y flies are devoid of any transgene and were used as controls. Note that *lacZ* staining of follicle cells surrounding egg chambers (shown at higher

magnification in insets) is observed in all ovaries because TSE only occurs in the germ line⁴. Original magnification, $\times 20$. **c–f**, Deep sequencing of small RNAs from ovaries of the indicated genotypes in which the maternally inherited allele is always indicated first. Plots show the abundance of 19–30-nucleotide (nt) small RNAs matching *P[lacW]* (**c–e**) or *P[lArB]* (**f**). Histograms show the length distributions of small RNAs matching *P[lacW]* or *P[lArB]* (dark bars), or only the *lacZ* sequence in these elements (blue bars). Positive and negative values correspond to sense and antisense reads, respectively.

with hemizygous *BX2* males, and recovered female progeny that had not inherited the *T-1* locus and carrying a paternally inherited *BX2* locus (Fig. 2). These females showed marked silencing of the TSE reporter transgene, indicating that the cytoplasm of *T-1* oocytes can

confer new silencing capacities to the inactive allele of the *BX2* locus. This *de novo* silencing allele will be hereafter referred to as *BX2** to differentiate it from the initial *BX2* allele never having been exposed to a *T-1* cytoplasm.

A *BX2** line was established and analysed in successive generations (Fig. 3a). Notably, second generation (G_2) *BX2** females from test crosses with males carrying a TSE reporter transgene still showed a complete TSE (Fig. 3b). This capacity to mediate TSE was fully maintained over 25 generations of the *BX2** line (TSE = 100%, $n = 4,600$). TSE remained very strong between G_{32} and G_{55} (99.4%, $n = 22,700$) showing a reversion rate less than 0.5% per generation at 25 °C (Supplementary Discussion). We conclude that maternally inherited factors from the *T-1* strain stably paramutated the *BX2* locus.

In contrast to *BX2* females, ovaries of G_2 *BX2** females contained abundant small RNAs matching the *BX2* sequence (Fig. 3c and Supplementary Table 2) with a profile similar to the one observed in *T-1* females (see Fig. 1c). The size distribution of these small RNAs showed a large peak corresponding to 23–28-nucleotide small RNAs with the piRNA ping-pong signature (Supplementary Fig. 2), as well as a discrete peak corresponding to a 21-nucleotide siRNA-like species of RNAs. Therefore, the acquired capacity of the *BX2** allele to mediate TSE correlates with the *de novo* production of *lacZ*-derived small RNAs from this locus. Finally, *BX2**-derived small RNAs were continuously produced in ovaries over at least 42 generations of a *BX2** line (Fig. 3d and Supplementary Figs 2 and 3). Together, these data indicate that the *BX2** paramutation is associated with stable production of high levels of small RNAs from the *BX2* locus in ovaries.

We next tested whether the paramutated *BX2** allele is paramutagenic. We crossed hemizygous *BX2** females with hemizygous naive *BX2* males and recovered female progeny having inherited the

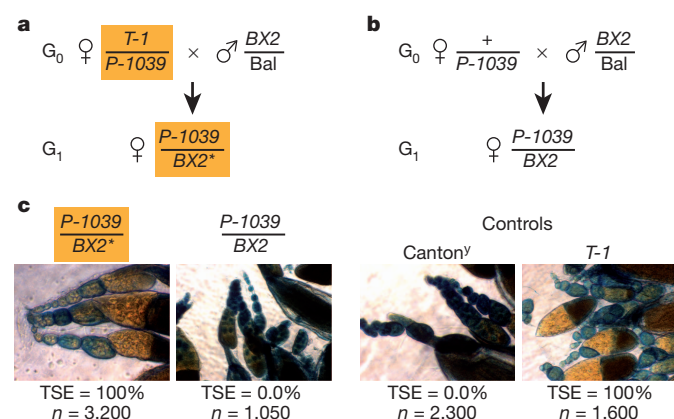


Figure 2 | Epigenetic induction of *BX2* by *T-1*. **a**, *T-1* females carrying the TSE reporter transgene *P-1039* were crossed to *BX2* males carrying a balancer chromosome (Bal). *BX2** female progeny having inherited cytoplasm from *T-1* mothers (orange background) and a *BX2* chromosome from fathers were stained for *lacZ*. **b**, Females carrying only the TSE reporter *P-1039* were crossed to *BX2* males. Female progeny from this cross were stained for *lacZ*. **c**, *P-1039*/*BX2** female progeny from the cross in **a** showed complete TSE, which was scored as indicated in Fig. 1. *P-1039*/*BX2* female progeny from the cross in **b** did not show TSE. Controls correspond to crosses between Canton^y (devoid of any transgene) or *T-1* females with *P-1039* males, which resulted in progeny showing null and complete TSE, respectively. Original magnification, $\times 20$.

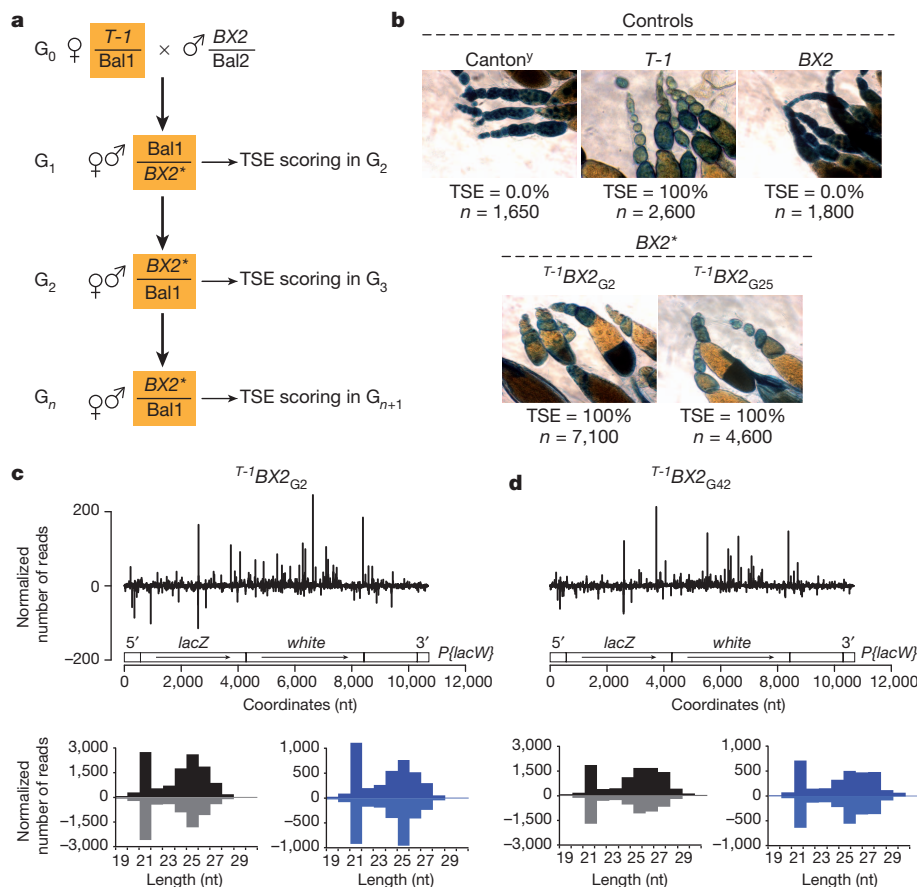


Figure 3 | *BX2 paramutation occurs and is associated to the production of small RNAs by the *BX2* cluster.** **a**, *BX2** lines were established as indicated. Bal1 and Bal2 are balancer chromosomes carrying distinct phenotypic markers. *BX2** siblings were crossed at each generation to perpetuate the *BX2** line. In addition, *BX2** females were crossed at various generations (G_n) to males carrying the *P-1039* reporter, to score the TSE of *BX2** in the G_{n+1} female progeny. **b**, TSE in *BX2** females from generations G_2 and G_{25} , and in progeny of crosses from Canton^S, *T-1* and *BX2* females with *P-1039* males as controls. *T-1 BX2*_{G25} indicates that *BX2* females inherited cytoplasm from *T-1* females 25 generations before the present cross. TSE was scored as indicated in Fig. 1. Original magnification, $\times 20$. **c**, **d**, Abundance (top) and length distribution (dark histograms) of 19–30-nucleotide small RNAs matching the *P[lacW]* transgene in ovaries from hemizygous *BX2** females from generation G_2 (**c**) and G_{42} (**d**). Length distributions of the subsets of small RNAs only matching *lacZ* are shown as blue histograms. Positive and negative values correspond to sense and antisense reads, respectively.

cytoplasm of *BX2** mothers and the *BX2* locus from fathers (Fig. 4a). This *BX2* allele was then assessed in generation G_2 for its capacity to silence a TSE reporter transgene in the germline. Notably, we observed a complete TSE (Fig. 4a), indicating that the paternally inherited *BX2* allele was paramutated through maternal inheritance of *BX2** cytoplasm. This newly paramutated *BX2* allele, which corresponds to a second-order paramutation, will be hereafter referred to as *BX2**². A *BX2**² line was established and showed stable TSE over 36 generations (Fig. 4a). Moreover, this line retained the capacity to produce large amounts of *BX2**²-derived small RNAs after 36 generations (Fig. 4b). Following an identical mating scheme, *BX2**² females were able to paramutate a paternally inherited *BX2* locus, generating a third-order *BX2**³ paramutated allele that showed full TSE capacity over 10 generations. Applying this procedure recurrently, we generated a

fifth-order paramutated *BX2**⁵ allele that showed full TSE capacity (Supplementary Fig. 6). In conclusion, the conversion of *BX2* to *BX2** by *T-1* maternal cytoplasm has all the properties of a paramutation, because it is stable over generations and the paramutated allele shows secondary paramutagenicity.

Interestingly, *T-1* also fully paramutated *C2*, another seven-copy transgene inserted at the same location (Supplementary Fig. 1), whereas lower-copy-number transgenes at this location were paramutated only transiently (Supplementary Table 3). A similar unstable paramutation interaction was also observed between the non-allelic *P-1152* and *BX2* loci (Supplementary Fig. 7).

As paramutation in this system is correlated with the production of *BX2**-derived piRNAs and siRNAs, we investigated the effect of *aubergine* and *Dicer-2* loss of function on a paramutated *BX2* cluster.

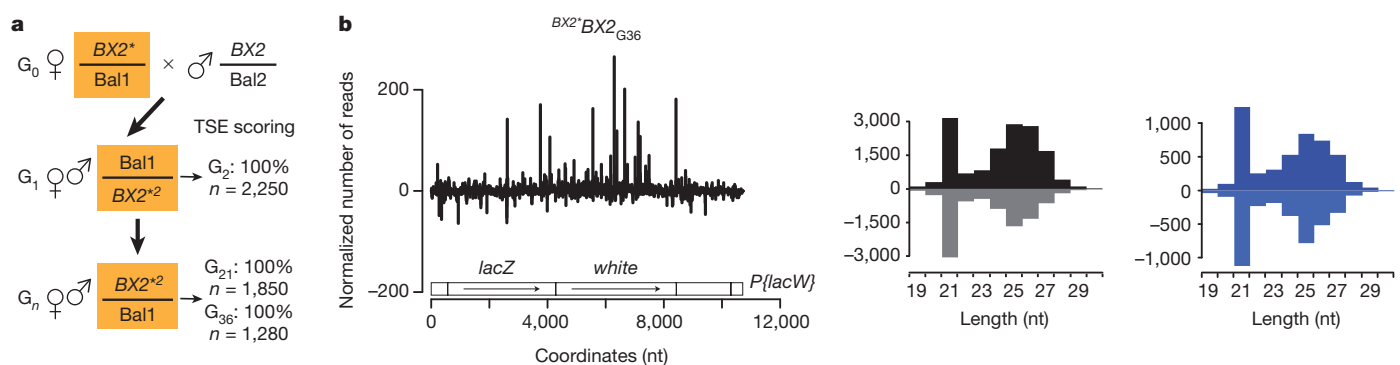


Figure 4 | Paramutated *BX2 is paramutagenic.** **a**, *BX2** females were crossed with *BX2* males and a *BX2**² line (second-order paramutation) was established as indicated. Bal1 and Bal2 are balancer chromosomes. *BX2**² siblings were crossed at various generations to perpetuate the *BX2**² line. In addition, *BX2**² females were crossed at each generation (G_n) with males carrying the *P-1039* reporter transgene to score the TSE of *BX2**² in the G_{n+1}

female progeny. **b**, Abundance (graph on the left) and length distribution (black histogram in the middle) of 19–30-nucleotide small RNAs matching the *P[lacW]* transgene in ovaries from hemizygous *BX2**² females from generation G_{36} . Length distribution of the subset of small RNAs only matching *lacZ* is shown in the blue histogram on the right.

The silencing capacity of the *BX2*² cluster was completely abolished in homozygous *aubergine* mutants, whereas strong silencing still took place in *Dicer-2* homozygous mutants (Supplementary Fig. 8). Moreover, the *BX2*² locus still showed full repression capacity after four generations in a *Dicer-2* homozygous mutant context. Hence, the *BX2*² silencing activity requires piRNAs, whereas neither *BX2*² activity nor inheritance rely on siRNAs. In maize, paramutation can be induced by a non-allelic transgene producing *b1*-repeat double-stranded RNA (dsRNA) and siRNAs¹⁵ and epigenetic inheritance of the *Kit^{tm1Alf}* mutant allele in mice seems to result from paternal as well as maternal transmission of small RNAs¹⁷. These data indicate that paramutations may in some instances involve small RNAs without interactions between alleles at the DNA or chromatin levels. Our findings that, in *Drosophila*, the *BX2* paramutation is triggered by cytoplasmic inheritance strongly support this view.

Finally, we investigated the effect of the paramutation on transcription of the *BX2* locus by quantitative polymerase chain reaction with reverse transcription (RT-qPCR). *BX2* and *BX2*² showed similar steady-state levels of both sense and antisense transcripts (Supplementary Fig. 9). This observation suggests that paramutation, rather than increasing the pool of piRNA precursor transcripts, activates their downstream processing into piRNAs. Thus, the maternally transmitted piRNAs could trigger production of primary piRNAs and/or ping-pong amplification of secondary piRNAs in the nuage. As paramutation is accompanied by *de novo* production of high levels of piRNA, it provides an invaluable model to determine the molecular events involved in the genesis of piRNA loci.

METHODS SUMMARY

All crosses were performed at 25 °C. *lacZ* expression assays were carried out using X-gal overnight staining²⁰. The *P-lacZ-white* construct (named *P{lacW}*) contains the *P-lacZ* translational fusion and is marked by the mini-*white* gene (Supplementary Fig. 1 and Supplementary Table 4). Small RNA libraries from hand-dissected ovaries were prepared using the Illumina kit and sequenced using an Illumina Genome Analyzer II or an Illumina HiSeq-2000, following the manufacturer's instructions. For library comparisons, read counts were normalized to the total number of small RNAs that matched the *D. melanogaster* genome and did not correspond to abundant cellular RNAs (ribosomal RNA, transfer RNA and small nuclear RNAs). Overlap signatures were computed for each sequence data set by collecting the appropriate RNA reads matching *P* transgenes and calculating overlap frequencies with RNA reads on the opposite strand.

Full Methods and any associated references are available in the online version of the paper.

Received 6 June 2011; accepted 16 July 2012.

Published online 26 August 2012.

- Brink, R. A. A genetic change associated with the R locus in maize which is directed and potentially reversible. *Genetics* **41**, 872–889 (1956).
- Coe, E. H. Jr. A regular and continuing conversion-type phenomenon at the B locus in maize. *Proc. Natl Acad. Sci. USA* **45**, 828–832 (1959).
- 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* **149**, 1839–1855 (1998).
- Josse, T. et al. Telomeric trans-silencing in *Drosophila melanogaster*: tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS ONE* **3**, e3249 (2008).
- Josse, T. et al. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet.* **3**, 1633–1643 (2007).
- Dorer, D. R. & Henikoff, S. Transgene repeat arrays interact with distant heterochromatin and cause silencing in *cis* and *trans*. *Genetics* **147**, 1181–1190 (1997).
- Ronssey, S., Boivin, A. & Anxolabehere, D. *P*-element repression in *Drosophila melanogaster* by variegating clusters of *P-lacZ-white* transgenes. *Genetics* **159**, 1631–1642 (2001).
- Chandler, V. L. Paramutation: from maize to mice. *Cell* **128**, 641–645 (2007).
- Hollick, J. B., Patterson, G. I., Coe, E. H., Cone, K. C. & Chandler, V. L. Allelic interactions heritably alter the activity of a metastable maize *pl* allele. *Genetics* **141**, 709–719 (1995).

- Pilu, R. et al. A paramutation phenomenon is involved in the genetics of maize *low phytic acid1-241* (*lpa1-241*) trait. *Heredity* **102**, 236–245 (2009).
- Sidorenko, L. V. & Peterson, T. Transgene-induced silencing identifies sequences involved in the establishment of paramutation of the maize *p1* gene. *Plant Cell* **13**, 319–335 (2001).
- Stam, M. Paramutation: a heritable change in gene expression by allelic interactions *in trans*. *Molecular Plant* **2**, 578–588 (2009).
- Alleman, M. et al. An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**, 295–298 (2006).
- Dorweiler, J. E. et al. mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* **12**, 2101–2118 (2000).
- Arteaga-Vazquez, M. et al. RNA-mediated trans-communication can establish paramutation at the *b1* locus in maize. *Proc. Natl Acad. Sci. USA* **107**, 12986–12991 (2010).
- Stam, M., Belele, C., Dorweiler, J. E. & Chandler, V. L. Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes Dev.* **16**, 1906–1918 (2002).
- Rassoulzadegan, M. et al. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441**, 469–474 (2006).
- Grandjean, V. et al. The *miR-124-Sox9* paramutation: RNA-mediated epigenetic control of embryonic and adult growth. *Development* **136**, 3647–3655 (2009).
- Kidwell, M. G., Kidwell, J. F. & Sved, J. A. Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility, and male recombination. *Genetics* **86**, 813–833 (1977).
- Engels, W. R. in *P Elements in Drosophila* (eds Berg, D. E. & Howe, M. M.) (American Society for Microbiology, 1989).
- Ronssey, S., Lehmann, M., Nouaud, D. & Anxolabehere, D. The regulatory properties of autonomous subtelomeric *P* elements are sensitive to a suppressor of variegation in *Drosophila melanogaster*. *Genetics* **143**, 1663–1674 (1996).
- Marin, L. et al. *P*-element repression in *Drosophila melanogaster* by a naturally occurring defective telomeric *P* copy. *Genetics* **155**, 1841–1854 (2000).
- Stuart, J. R. et al. Telomeric *P* elements associated with cytotypic regulation of the *P* transposon family in *Drosophila melanogaster*. *Genetics* **162**, 1641–1654 (2002).
- Poyhonen, M. et al. Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3 (Bethesda)* **2**, 331–338 (2012).
- Todeschini, A. L., Teyssset, L., Delmarre, V. & Ronssey, S. 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 (2010).
- Dorer, D. R. & Henikoff, S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77**, 993–1002 (1994).
- Brennecke, J. et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
- Czech, B. et al. An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453**, 798–802 (2008).
- Muerdter, F. et al. Production of artificial piRNAs in flies and mice. *RNA* **18**, 42–52 (2011).
- Lemaitre, B., Ronssey, S. & Coen, D. Maternal repression of the *P* element promoter in the germline of *Drosophila melanogaster*: a model for the *P* cytotypic. *Genetics* **135**, 149–160 (1993).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank O. Sismeiro, J.-Y. Copée, E. Mouchel-Vielh, V. Ribeiro, C. Pappalico and P. Graça for technical assistance, D. Dorer, S. Henikoff and the Bloomington Stock Center for providing stocks, and flybase.org for providing databases. We thank T. Josse for preliminary experiments. We thank J.-R. Huynh, V. Colot, N. Randsholt, A.-M. Pret, C. Carré and F. Peronnet for critical reading of the manuscript. S.R. thanks D. Anxolabehere and M. Lehmann for previous help. This work was supported by fellowships from the Ministère de l'Enseignement Supérieur et de la Recherche to A.d.V. and C.H., from the Fondation pour la Recherche Médicale to A.d.V., from the Association Nationale de la Recherche (ANR) to A.-L.B., and by grants from the Association pour la Recherche contre le Cancer to S.R. and from the ANR (project "Nuclear endosRNAs") to C.A.

Author Contributions Genetic experiments were conceived by A.d.V., A.B. and S.R., and performed by A.d.V., A.B., C.H., V.D., L.T. and S.R. L.T. conceived and performed molecular mapping of the clusters and Southern blot analysis. Deep-sequencing analysis was conceived by A.d.V., A.-L.B., S.R. and C.A., and performed by A.d.V. and A.-L.B. Bioinformatic analysis was conceived and performed by C.A. RT-qPCR was conceived and performed by A.B. S.R., A.d.V., A.B. and C.A. wrote the paper and all authors discussed the results.

Author Information Small RNA sequences have been deposited at the National Center for Biotechnology Information under accession SRP012172. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.R. (stephane.ronssey@upmc.fr) or C.A. (christophe.antoniowski@upmc.fr).

METHODS

Experimental conditions. All crosses were performed at 25 °C and involved 3–5 couples in most cases. *lacZ* expression assays were carried out using X-gal overnight staining as described previously³⁰, except that ovaries were fixed for 6 min. **Transgenes and strains.** *P-lacZ* fusion enhancer trap transgenes *P-1152*, *BQ16*, *BC69* and *P-1039* all contain an in-frame translational fusion of the *Escherichia coli lacZ* gene to the second exon of the *P* transposase gene and a *rosy* transformation marker³¹. The *P-1152* insertion (Supplementary Table 4) was mapped to the telomere of the X chromosome (cytological site 1A) and consists of two *P-lacZ* insertions in the same TAS unit and in the same orientation⁵. *P-1152* is homozygous, viable and fertile. *BQ16* is located at 64C in euchromatin of the third chromosome⁴ (Supplementary Table 4) and is homozygous, viable and fertile. *BC69* is inserted in chromosome 2 (Supplementary Table 4) in the first exon of the *vasa* gene and results in a *vasa* loss-of-function allele; consequently, it is homozygous, female and sterile. *P-1039* is located at 60B on the second chromosome (Supplementary Table 4) and is homozygous lethal. *P-1152* shows no *lacZ* expression in the ovary, *BQ16* and *BC69* are strongly expressed in the nurse cells and in the oocyte and *P-1039* shows strong *lacZ* staining in numerous tissues including the follicle cells, the nurse cells and the oocyte.

***P-lacZ* clusters.** Lines with different numbers of *P-lacZ-white* transgenes³² located at cytological site 50C on the second chromosome^{6,26} were used (Supplementary Table 4). The transgene(s) insertion site is located near the *mRpl53* gene, in an *Ago1* intron. This site is not a piRNA-producing locus, as observed for instance in the deep-sequencing data set from *P-1152* ovaries (data not shown). The *P-lacZ-white* construct contains the *P-lacZ* translational fusion and is marked by the mini-white gene (*P[lacW]*, FBtp0000204). *BX2* carries seven *P-lacZ* copies including at least one defective copy inserted in direct orientations. *T-1* derives from *BX2* following X-ray treatments (Supplementary Fig. 1). *T-1* has chromosomal rearrangements including translocations between the second and the third chromosomes. After overnight staining, weak *lacZ* expression is detected in the follicle cells of *BX2* and *T-1* female ovaries, presumably because of a position effect at 50C, but no staining is observed in the germ line (data not shown).

Lines carrying transgenes have M genetic backgrounds (devoid of *P* transposable elements), as do the multi-marked balancer stocks used in genetic experiments. The Canton^y and *w*¹¹¹⁸ lines were used as controls completely devoid of any *P* element or transgene. Crosses involving *P-1152* were performed with females carrying the telomeric transgenes in the homozygous state (except where indicated), whereas crosses performed with *BX2* or *T-1* were performed with females carrying the cluster in the heterozygous state (referred to as hemizygous in case of insertions) because of the sterility (*BX2*) and lethality (*T-1*) induced by transgene clusters.

Two strong hypomorphic mutant alleles of *aubergine* induced by EMS were used. Both of them are homozygous, female and sterile, and TSE was previously shown to be abolished by a heteroallelic combination of these alleles⁵. *aub*^{QC42} comes from the Bloomington Stock Center (stock no. 4968) and has not been characterized at the molecular level³³. *aub*^{N11} has a 154-bp deletion, resulting in a frameshift which is predicted to add 16 novel amino acids after residue 740 (refs 34, 35). *Dicer-2*^{L8116X} is a loss-of-function allele induced by EMS that has a sequence variant at residue 811 resulting in a stop codon³⁶. It is homozygous, viable and fertile.

Quantification of TSE. When TSE is incomplete, variegation is observed because 'on/off' *lacZ* expression is seen between egg chambers: that is, egg chambers can show strong expression (dark blue) or no expression, but intermediate expression levels are rarely found. TSE was quantified as previously described⁵ by determining the percentage of egg chambers with no expression in the germ line.

Deep sequencing analyses. Small RNAs from hand-dissected ovaries were cloned using the DGE-Small RNA Sample Prep Kit and the Small RNA Sample Prep v.1.5 Conversion Kit from Illumina (libraries 1 to 5), following the manufacturer's instructions, or using the TruSeq (TM) SBS v.5 Kit at Fasteris (<http://www.fasteris.com/>) (libraries 6 to 8). Libraries 1 to 5 were sequenced using an Illumina Genome Analyzer II and libraries 6 to 8 were sequenced using an Illumina Hi-Seq 2000. Sequence reads in fastq format were trimmed from the adaptor sequence 5'-TCGTATGCCGTCTTCTGCTTG-3' (libraries 1 to 5) or 5'-CTGTAGG CACCATCAAT-3' (libraries 6 to 8) and matched to the *D. melanogaster* genome release 5.43 using Bowtie³⁷, as well as to the sequences of the *P*-element constructs *P[lArB]* (FlyBase accession FBtp0000160) and *P[lacW]* (FlyBase accession FBtp0000204). Only 19–30-nucleotide reads matching the reference sequences with 0 or 1 mismatch were retained for subsequent analysis. For global annotation of the libraries (Supplementary Table 2), we used release 5.43 of fasta reference files available in FlyBase, including transposon sequences (dmel-all-transposon_r5.43.fasta) and release 18 of miRNA sequences from miRBase (<http://www.mirbase.org>).

Sequence length distributions, small RNA mapping and frequency maps were generated using in-house Python scripts and R (<http://www.r-project.org/>) to

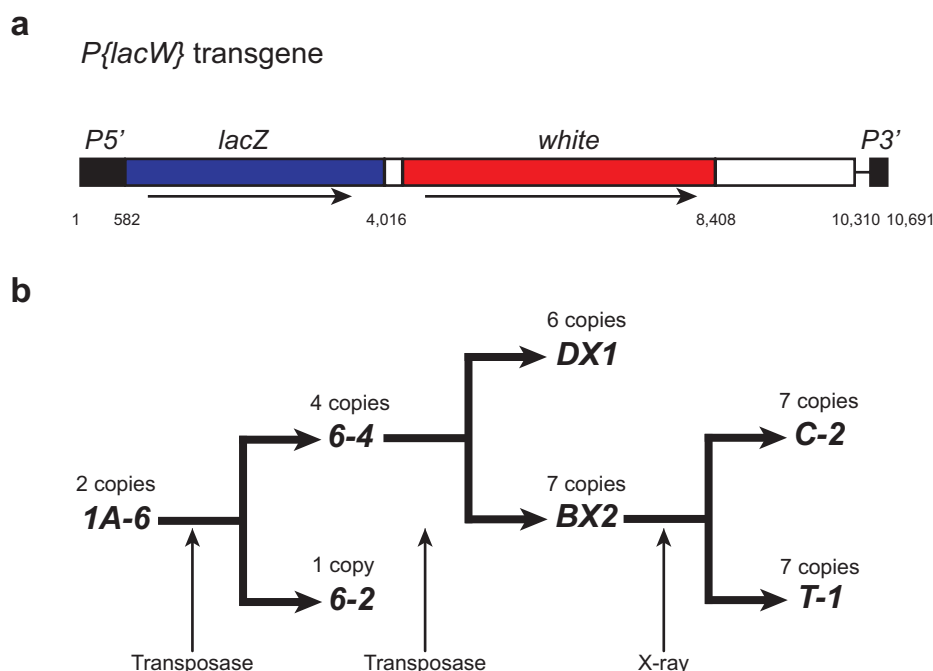
analyse Bowtie outputs. Scripts were integrated and run in a Galaxy instance hosted by the laboratory. The corresponding Mississippi suite of analysis workflows and codes is accessible from <http://www.drosophila.org> upon request. For library comparisons, read counts were normalized (Supplementary Table 2) to the total number of small RNAs that matched the *D. melanogaster* genome and did not correspond to abundant cellular RNAs (rRNA, tRNA and snoRNAs). For small RNA mapping, we matched each individual RNA sequence to *P[lArB]* or *P[lacW]* and gave to each matched position a weight corresponding to the normalized occurrence of the sequence in the small RNA library. When RNA sequences matched *P[lArB]* or *P[lacW]* repeatedly, the weight was divided by the number of hits to these *P*-element constructs.

Distributions of piRNA overlaps (ping-pong signatures) were computed by collecting, for each sequencing data set, all the 23–28-nucleotide RNA reads matching *P[lArB]* or *P[lacW]* whose 5' ends overlapped with another 23–28-nucleotide RNA read on the opposite strand. Then, for each possible overlap of 1–28 nucleotides, the number of read pairs was counted. Distributions of siRNA overlaps were computed using a similar procedure, except that 20–22-nucleotide RNA reads were collected instead of the 23–28-nucleotide RNA reads. The distributions of piRNA/siRNA overlaps were computed by collecting separately the 20–22-nucleotide and 23–28-nucleotide RNA reads matching *P[lArB]* or *P[lacW]*, and counting for each possible overlap of 1–22 nucleotides the number of read pairs across these two distinct read data sets. 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 it using the formula $z(i) = (O(i) - \text{mean}(O))/\text{standard deviation}(O)$.

RT-qPCR experiments. Total RNA was extracted (Qiagen kit) from ovaries dissected from 1A-6, *BX2* and *BX2** females and quantified (NanoDrop). Four to six biological replicates were made for each genotype. For each sample, 10 µg of RNA was treated with DNase (Fermentas). 1 µg of DNase-treated RNA was used for reverse transcription (Fermentas) using either no primer (control RT) or two primers simultaneously (specific RT): one specific to the *nanos* transcript used as the sample RNA quantification reference (5'-GGATTCGCCCTCTCTAAACC-3') and the second specific to a region of the *P[lacW]* transgene. *P[lacW]* RT primers were designed to be specific to the sense (s) or to the antisense (a) transcripts of five regions of the *P[lacW]* transgene: 5'*P*, 5'*lacZ*, 3'*lacZ*, 5'*white* and 3'*P*. Sequences are: a1 (5'-ATTCAAACCCACGGACAT-3'), a2 (5'-AGTACGAAATGCGTCGTTTAGAGC-3'), a3 (5'-GGGGAAAACCTTATTATCAGCCG-3'), a4 (5'-GCTGTTTGCTCTCTCTG-3'), s1 (5'-GTTTTCCAGTCAAGCAGT-3'), s2 (5'-AATGCGCTCAGGTCAAATTC-3'), s3 (5'-TATGGAACCGTCGATATTCAGCC-3'), s4 (5'-ATTTTGTGGTTCGAGTTC-3'), s5 (5'-TTAAGTGTACTTCGGTAAGCTTCG-3'), s6 (5'-TTTGGGAGTTTTCACCAAGG-3'). One primer was both antisense and sense (as) because it is located in the inverted repeat of the *P* element. It is (5'-TGATGAATAACATAAGGTGGTCCGTCG-3'). RT primers are shown on the transgene map (Supplementary Fig. 9). qPCR was then performed on triplicates of each RT with a primer pair specific for the *nanos* gene in order to quantify the *nanos* transcripts. Simultaneously, qPCR was performed on triplicates of the same RT using different primer pairs corresponding to the former five regions of interest of *P[lacW]*. qPCR primer sequences are: 5'*P* (5'-CTGCAAAGCTGTGACTGGAG-3' and 5'-TTTGGGAGTTTTCACCAAGG-3'), 5'*lacZ* (5'-GAGAATCCGACGGTTGTGA-3' and 5'-AAATTCAGACGGCAAACGAC-3'), 3'*lacZ* (5'-ACTATCCGACCGCCTTACT-3' and 5'-GTGGGCCATAATCAATTCG-3'), 5'*white* (5'-GTCAATGTCCGCTTCAGTT-3' and 5'-GGAGTTTTGGCACAGTTC-3') and 3'*P* (5'-CCACGGACGCTAAGGCTAA-3' and 5'-GTCGCAAGAGACATCCACT-3'). The same series of dilutions composed of a mix of different RT preparations was used to normalize the quantity of *nanos* transcripts in all RT preparations leading to standard quantity (Sq) values for *nanos* transcripts in specific RT (using *nanos* primer = Sq(*nanos*)) or in control RT (without primer = Sq(control *nanos*)) preparations. A series of dilutions of a plasmid containing the *P[lacW]* transgene was used to normalize the quantity of transcripts of the clusters leading to Sq values for cluster transcript (Sq(specific) and Sq(control specific)). Variations between technical triplicates seem to be very low when compared to variations between biological replicates. The mean of the three technical replicates was then systematically used (Sq). The measure of the quantity of transcripts from a given region for one biological sample was then calculated using the formula: $(\text{Sq(specific)} - \text{Sq(control specific)}) / (\text{Sq(nanos)} - \text{Sq(control nanos)})$. This allowed us to eliminate the background noise due to both sense and antisense transcripts (Sq(control transcript)) and to take into account variations in the quantity of RNA between biological samples (Sq(*nanos*)).

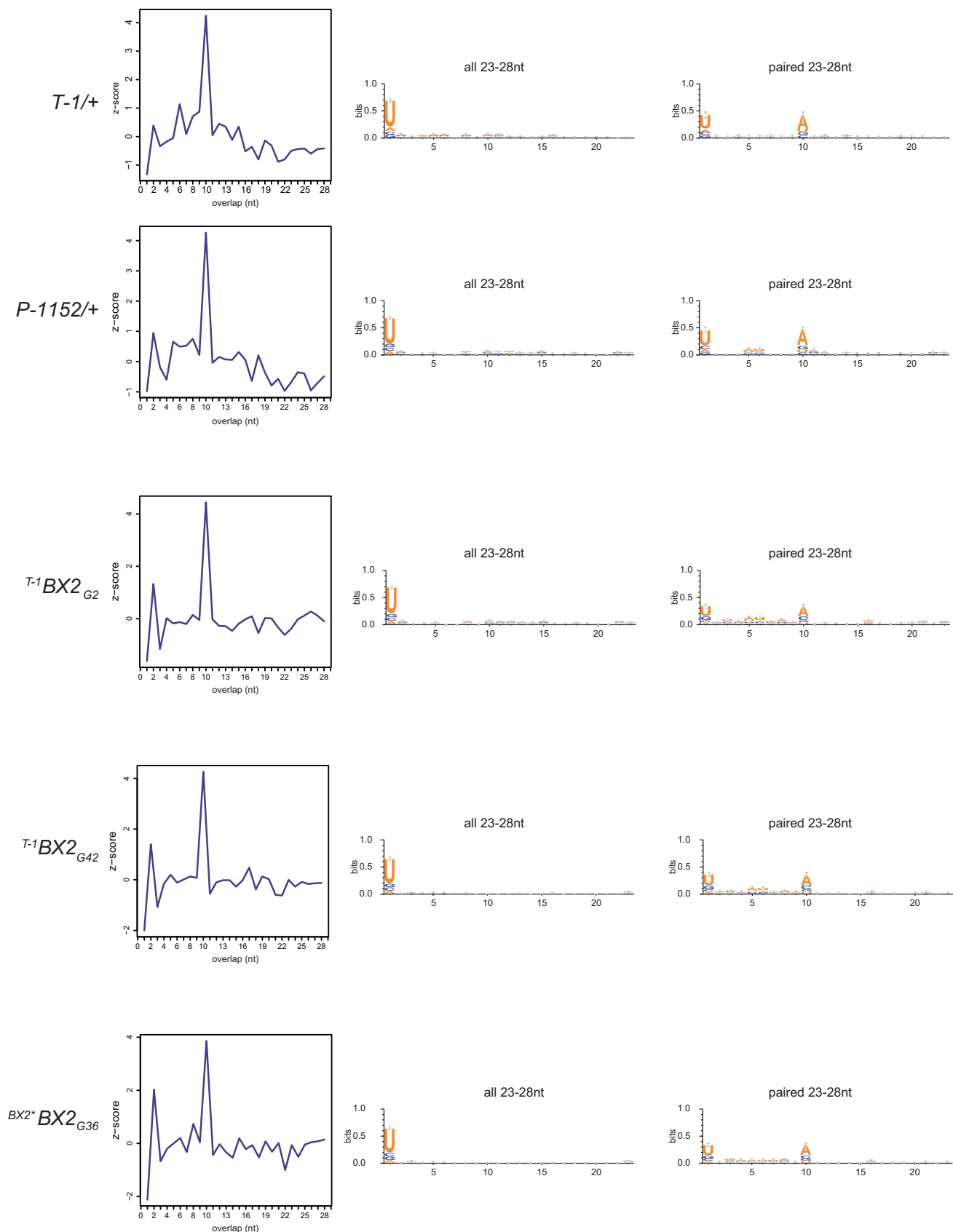
31. O'Kane, C. J. & Gehring, W. J. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl Acad. Sci. USA* **84**, 9123–9127 (1987).

32. Bier, E. *et al.* Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev.* **3**, 1273–1287 (1989).
33. Schupbach, T. & Wieschaus, E. Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**, 1119–1136 (1991).
34. Wilson, J. E., Connell, J. E., Schlenker, J. D. & Macdonald, P. M. Novel genetic screen for genes involved in posterior body patterning in *Drosophila*. *Dev. Genet.* **19**, 199–209 (1996).
35. Harris, A. N. & Macdonald, P. M. *aubergine* encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* **128**, 2823–2832 (2001).
36. Lee, Y. S. *et al.* Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69–81 (2004).
37. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).



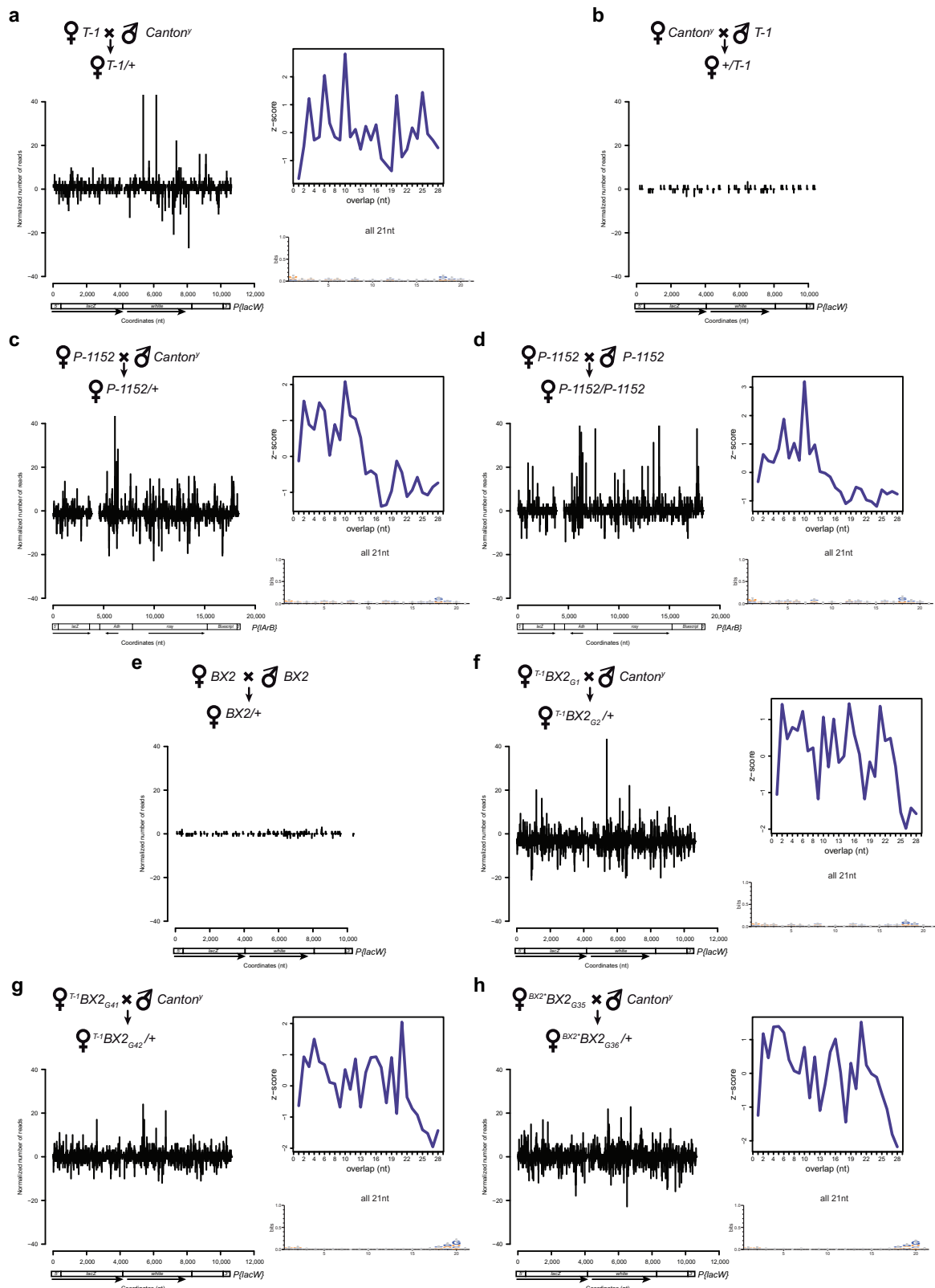
Supplementary Figure 1. *P-lacZ* transgene clusters at the cytological site 50C of chromosome 2.

(a) The *P{lacW}* transgene (Flybase FBtp0000204) contains an in-frame translational fusion of the *E. coli lacZ* gene to the second exon of the *P*-element transposase gene and a mini-*white* marker gene. The P5' and P3' black boxes indicate *P* element sequences. **(b)** Chromosomes carrying varying numbers of tandemly repeated *P{lacW}* transgenes were generated previously^{1,2}. The relationships between *P{lacW}* clusters are indicated. Transposase-mediated mobilization of the initial 1A-6 *P{lacW}* tandem repeat generated both the single-copy 6-2 and the four-copy cluster 6-4. Additional mobilization generated the six-copy DX1 and the seven-copy BX2 clusters. X-ray mutagenesis of BX2 males² generated two additional strains carrying the 7-copy clusters C-2 and T-1, respectively. No alterations of the 7-copy cluster structure in these strains were detected by Southern analysis³, but they show various rearrangements at distant chromosomal sites. In addition, clusters appear stable over time in Southern analysis (data not shown). The single insertion 6-2 is fertile and viable at the homozygous state. The other clusters are lethal or sterile at the homozygous state and are therefore maintained in fly stocks over a *CyO* balancer chromosome.

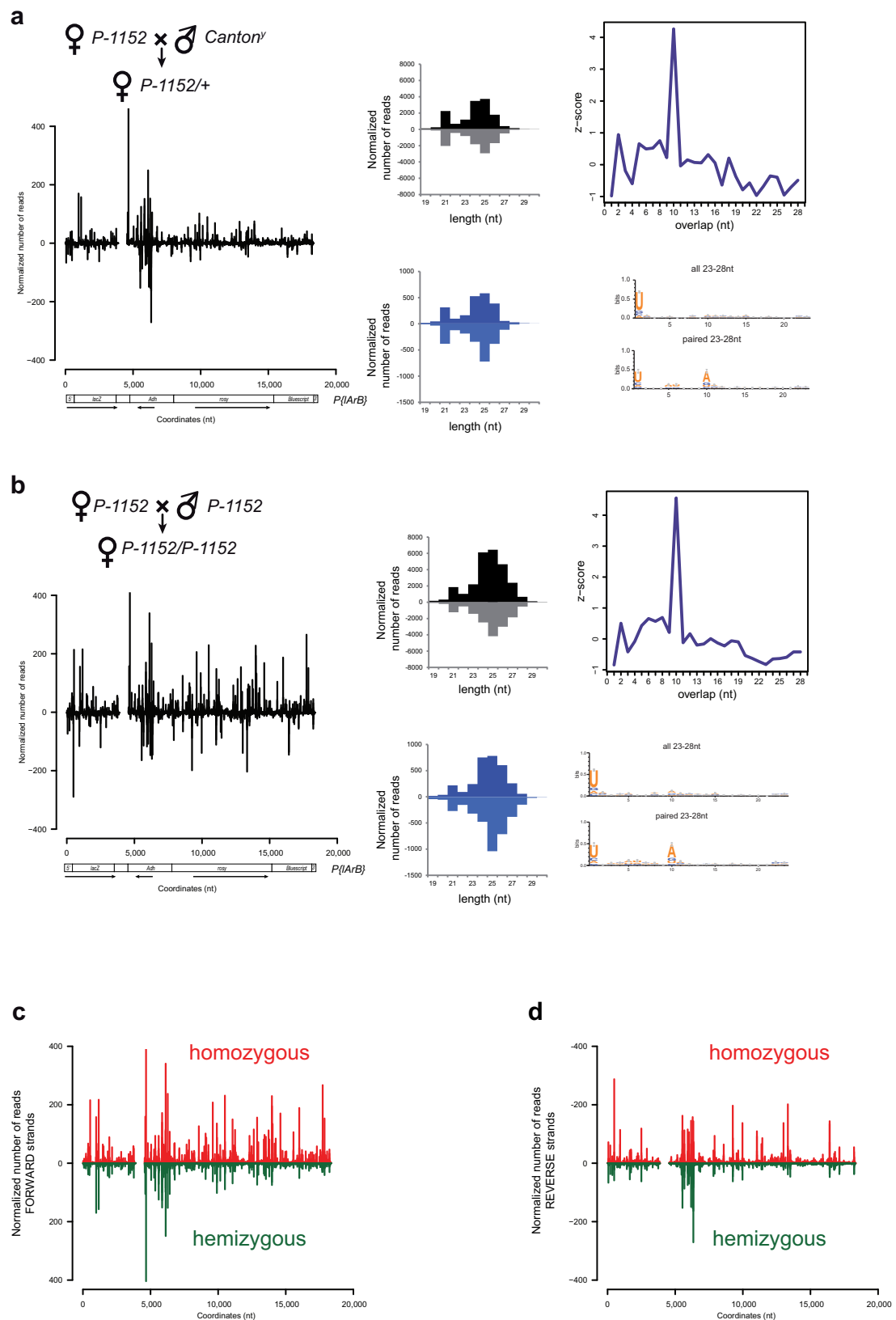


Supplementary Figure 2. Small RNAs matching *P-1152*, *T-1* and *BX2* show the “ping-pong” signature of piRNAs. Small RNAs in ovaries from females having inherited the indicated allele maternally (on the left) were analyzed. Right panels show the relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subsets of 23-28nt small RNAs matching

P{lacW} or *P{larB}*. The sequence composition of all matched 23-28nt small RNAs (middle panels) reveals a strong bias for a U at the first position. The matched 23-28nt small RNAs that formed sense-antisense pairs with complementarity over their 10 first nucleotides (right panels) were strongly enriched in 1U-10A pairs, a typical feature of piRNAs. Small RNAs in ovaries from females having inherited the *BX2* naive allele, as well as in ovaries from females having inherited the *T-1* allele paternally, did not show significant “ping-pong” signatures (not shown).

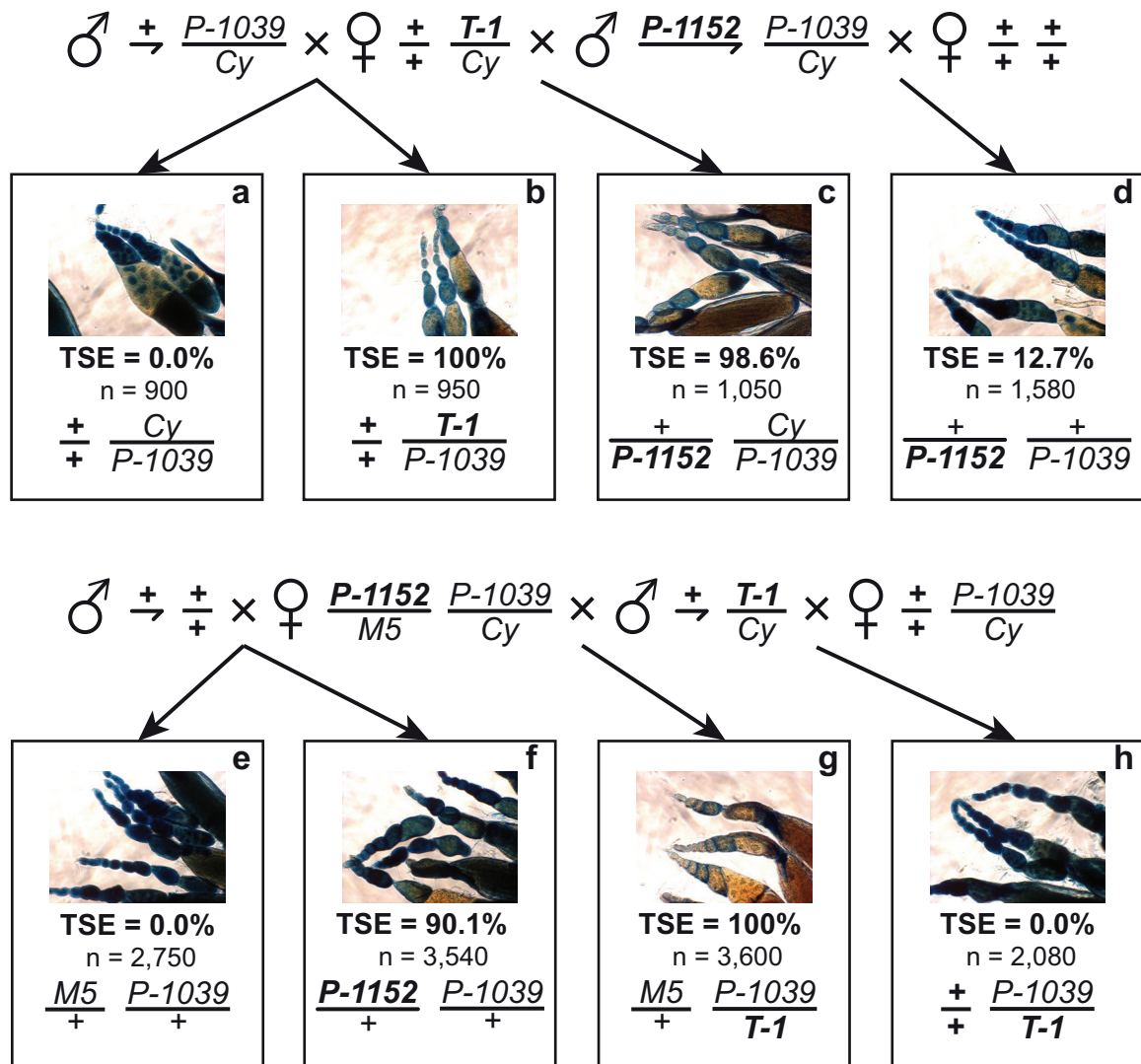


Supplementary Figure 3. *P-1152*, *T-1* and a paramutated *BX2 cluster produce siRNA-like species.** Bar plots show the abundance of 20-22nt small RNAs matching the *P{lacW}* (a, b, e, f) or *P{ArB}* (c, d) sequences in ovaries from female progeny of the indicated cross. Positive and negative values correspond to sense and antisense reads, respectively. z-score plots show, when applicable, the relative frequencies of 20-22nt matching small RNAs that pair with the indicated overlap with 23-28nt small RNAs. Seqlogos show, when applicable, the sequence composition of matched 21nt small RNAs.

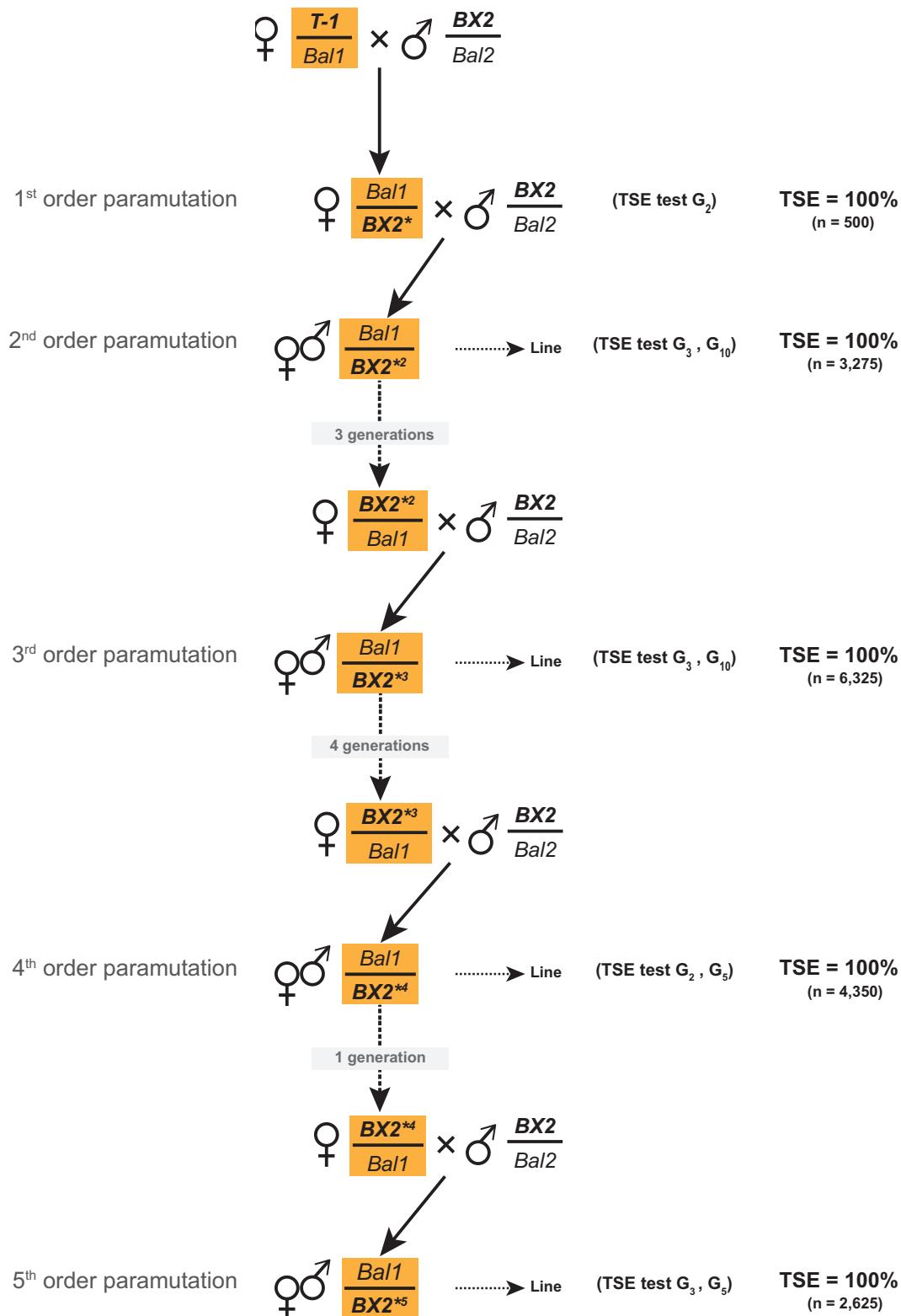


Supplementary Figure 4. Comparison of 19-30nt small RNAs in homozygous and hemizygous *P-1152* females. Bar plots showing the abundance of 19-30nt small RNAs matching the $P\{lArB\}$ sequences in ovaries from hemizygous (a – reprint from Fig.1) or homozygous (b) *P-1152* females obtained from the indicated cross. Positive and negative values correspond to sense and antisense

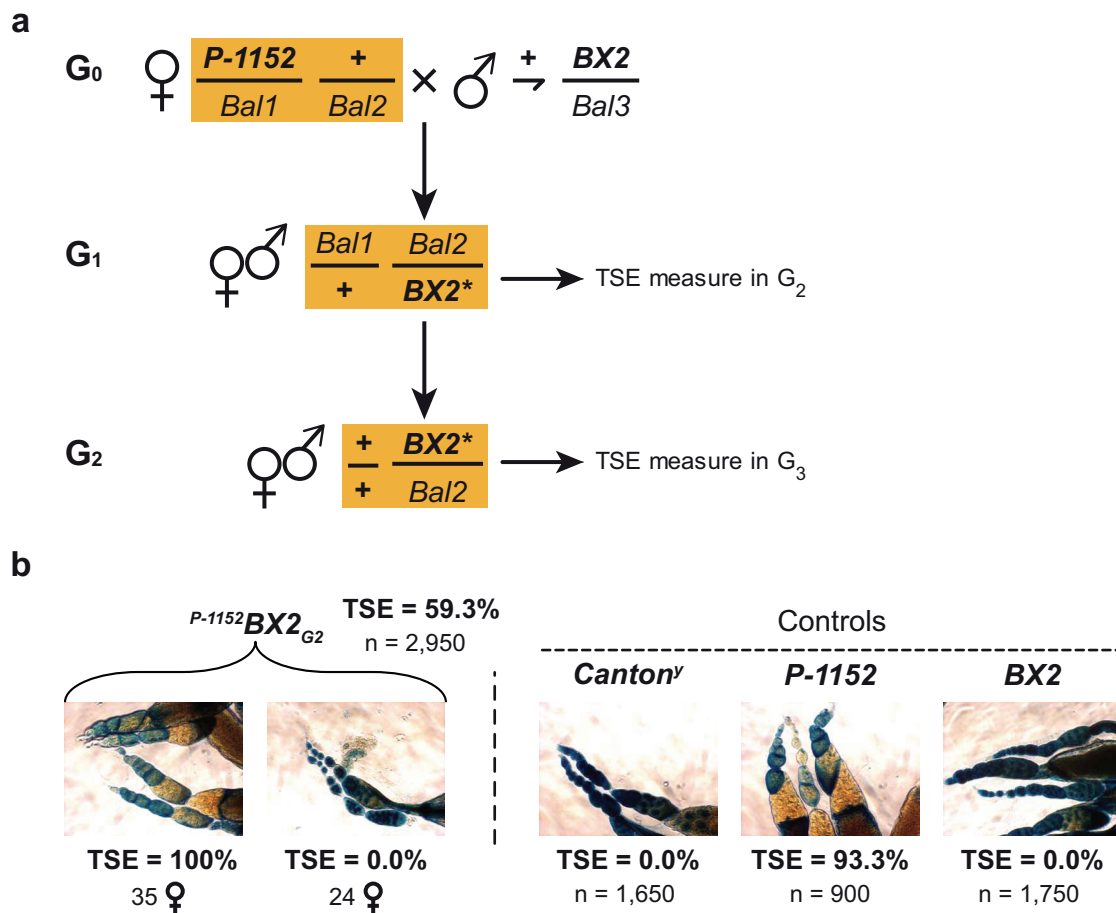
reads, respectively. Middle histograms show the length distributions of small RNAs matching *P{lArB}* (dark) or the *lacZ* sequence only (blue). Graphs to the right show the relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subset of 23-28nt small RNAs matching *P{lArB}*. For direct comparison, sense reads, in **c**, or antisense reads, in **d**, from hemizygous (green) and homozygous (red) *P-1152* females are plotted again together on the same graph. Homozygous *P-1152* females produce about twice as many 20-28nt small RNAs than hemizygous females.



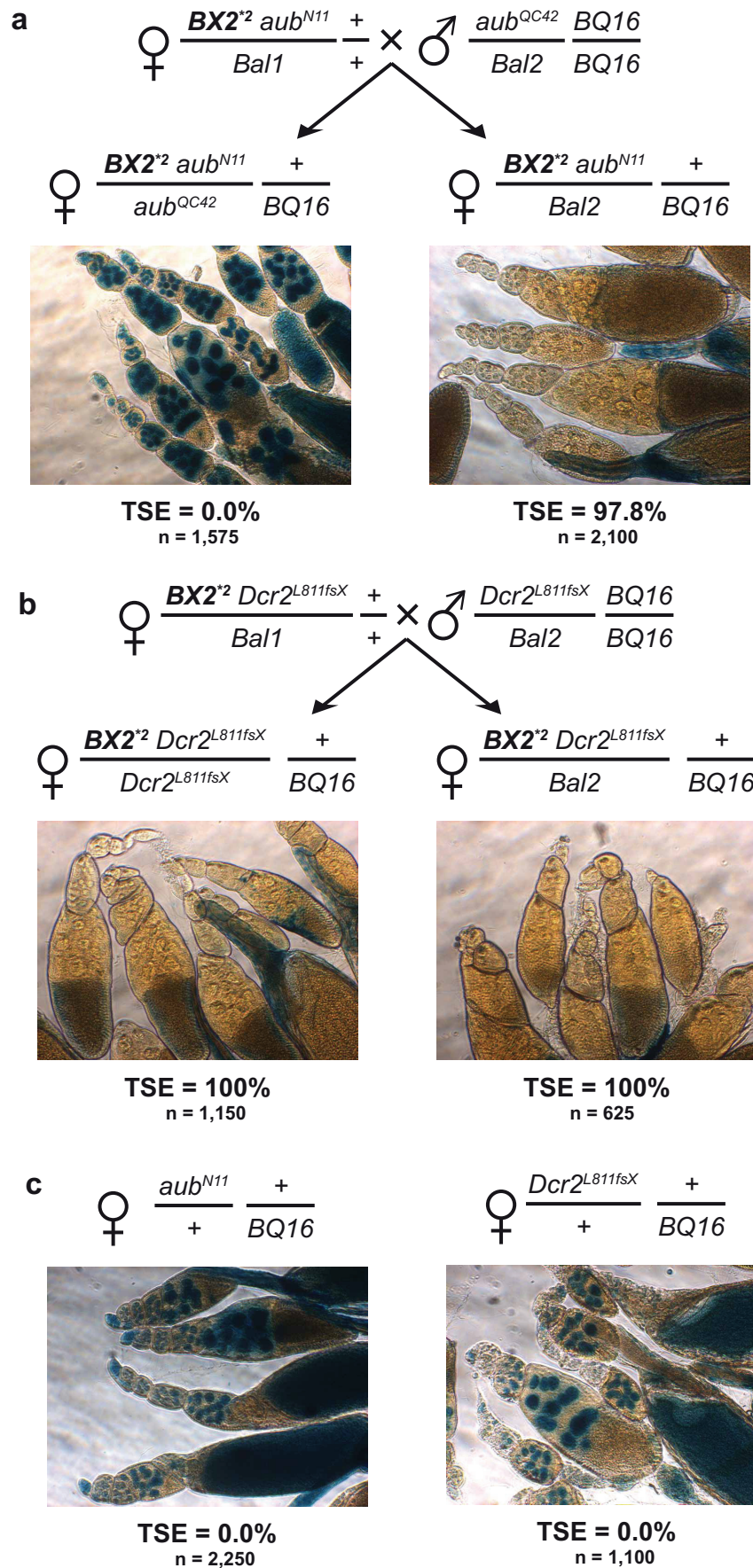
Supplementary Figure 5. Chromosomal and maternally-transmitted components of *T-1* and *P-1152* complement each other. *LacZ* staining of ovaries of G_1 females from the indicated crosses was performed and TSE on the *P-1039* reporter transgene was expressed as the percentage of egg chambers displaying repressed *lacZ* staining of germline cells among the total number (n) of egg chambers analyzed. Note that *lacZ* staining of somatic follicle cells surrounding egg chambers is observed in all ovaries irrespective of the genotype. *P-1152* is carried by the *X* chromosome. *T-1* and the *P-1039* reporter transgenes are carried by chromosome 2. *M5* and *Cy* are balancer chromosomes. Females that inherited cytoplasm from *T-1* mothers, but not the *T-1* chromosome, show no TSE (a), whereas females that inherited both cytoplasm and the *T-1* chromosome from *T-1* mothers show complete TSE (b). Females that only inherited the *P-1152* chromosome from fathers show a weak TSE (d). In contrast, females that inherited cytoplasm from *T-1* mothers and the *P-1152* chromosome from fathers show strong TSE (c). Females that inherited cytoplasm from *P-1152* mothers, but not the *P-1152* chromosome, show no TSE (e), whereas females that inherited both cytoplasm and the *P-1152* chromosome from *P-1152* mothers show strong TSE (f). Females that only inherited the *T-1* chromosome from fathers show no TSE (h). In contrast, females that inherited cytoplasm from *P-1152* mothers and the *T-1* chromosome from fathers show complete TSE (g).



Supplementary Figure 6. BX2 paramutations of third to fifth orders. Following the depicted genetic strategy, $BX2^{*2}$ females were tested for their capacity to be paramutagenic, thereby generating a third order paramutation ($BX2^{*3}$). The procedure was recurrently performed to generate $BX2^{*5}$ females (fifth order paramutation). $BX2^{*2}$, $BX2^{*3}$, $BX2^{*4}$ and $BX2^{*5}$ lines were established. The tests show complete and stable TSE over generations. Measures for TSE in first and last generations, the global percentage for all tests and the corresponding number of egg chambers counted (n) are indicated on the right-hand side.

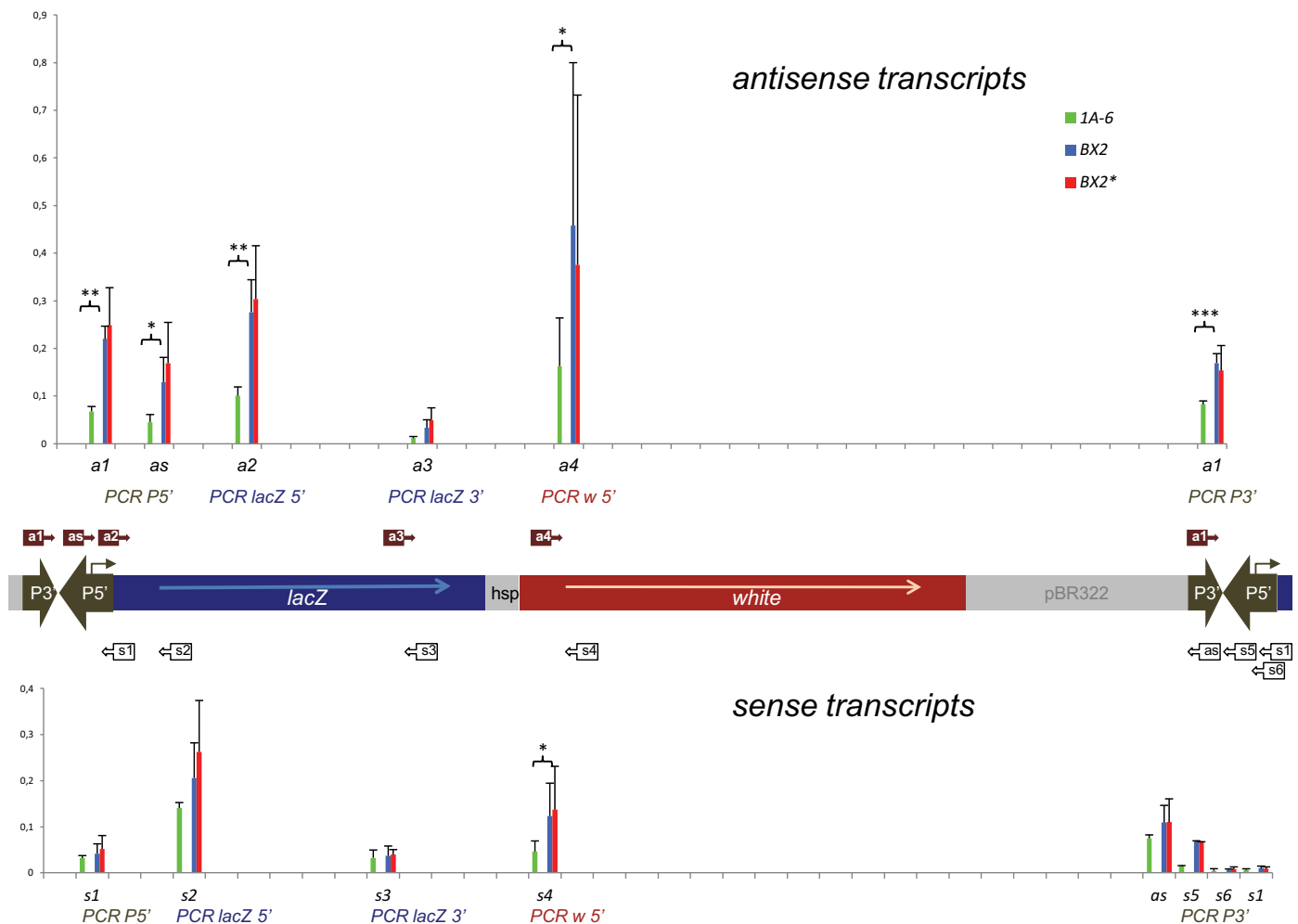


Supplementary Figure 7. Non-allelic "paramutation". We investigated whether an epigenetic conversion similar to paramutation can occur between non-allelic loci. **(a)** Hemizygous females carrying the telomeric *P-1152* silencer locus on the *X* chromosome were mated to *BX2* males in 8 independent crosses. G_1 *BX2*^{*} female progeny were mated to males carrying the *P-1039* reporter transgene and TSE was scored in the G_2 *BX2*^{*} female progeny from these crosses. *Bal1*, *Bal2* and *Bal3* are balancer chromosomes carrying distinct phenotypic markers. **(b)** $P-1152 BX2_{G2}$ females show partial repression capacities (35 females with complete TSE and 24 females with no TSE, resulting in 59.3% TSE). We further derived independent lines from the initial crosses. Out of 8 established *BX2*^{*} lines, 2 still showed TSE capacity after 5 generations. Together, these data indicate that the *P-1152* locus is able to paramutate the non-allelic *BX2* locus, although with incomplete penetrance. This partial penetrance may result from the partial sequence homology between the *P-lacZ-rosy* transgenes of the telomeric *P-1152* locus and the *P-lacZ-white* transgene cluster of the *BX2* locus. TSE in *Canton^y*, *P-1152* and *BX2* females are shown as controls. TSE was scored as indicated in Fig. 1.



Supplementary Figure 8. The repression capacity of a paramutated cluster is abolished by mutations affecting a piRNA pathway gene but not a siRNA pathway gene. Recombinant

chromosomes carrying a *BX2* cluster and a mutant allele of *aubergine* (*aub*) or *Dicer2* (*Dcr2*), involved in the piRNA and siRNA silencing pathways, respectively, were generated. Males carrying these chromosomes were crossed, as in Figure 4, with *BX2** females allowing recovery of paramutated *BX2**² *aub*⁻ or *BX2**² *Dicer2*⁻ (*Dcr2*⁻) chromosomes. Lines were established as in Figure 4 and *BX2**² *aub*⁻ or *BX2**² *Dcr2*⁻ females were crossed with males carrying a *P-lacZ* target (*BQ16*) and a mutant allele of *aub* or *Dcr2*, respectively. This produced females carrying a *BX2** cluster and a target transgene in a heteroallelic or homozygous mutant context for *aub* or *Dcr2* (**a** or **b**, respectively), together with sisters which have the same set of transgenes but are heterozygous for the mutation tested. Expression controls of *BQ16* in heterozygous *aub* and *Dcr2* mutant contexts are shown (**c**). Figure shows that *aub* loss of function results in a complete loss of the *BX2**² silencing capacities (**a**), whereas *Dcr2* loss of function has no effect on these capacities (**b**). In each case, the *BX2**² silencing capacity of heterozygous *aub* or *Dcr2* females is almost complete or complete, showing that the paramutation process was not impaired by a single dose reduction of these genes. In addition, a *BX2**² *Dcr2*⁻ chromosome was maintained over a *Dcr2* mutant chromosome during 4 generations and females were crossed with *BQ16* males at each generation. In each case, complete repression was observed showing that *Dcr2* loss of function does not impair the maintenance of the *BX2* paramutated state (G₄, TSE = 100%, n = 1,310). The same loss of function assay was not possible for *aub* because of the sterility of *aub* homozygous mutant females.



Supplementary Figure 9. Paramutation does not affect the *BX2* cluster transcription profile.

The mean and standard deviation of the measure of the quantity of steady-state sense and antisense transcripts for four to six biological replicates for each genotype (*1A-6* green, *BX2* blue and *BX2** red) are shown. A map of the *P{lacW}* transgene is given with the position and orientation of the RT primers indicated in red boxes-arrow (antisense) or white boxes-arrow (sense). Paired *t*-test analysis reveal significant statistical differences between the quantities of *1A-6* and *BX2* antisense transcripts (*= $p < 0.05$; **= $p < 0.01$; ***= $p < 0.001$) while no difference is observed when *BX2* and *BX2** are compared. The paramutation thus is not correlated to a change in the steady state level of transcription of the *BX2* locus for both sense and antisense transcripts. Interestingly, overlapping antisense transcripts are observed (PCR *P5'* using *a1* RT primer) at comparatively similar levels to antisense transcripts measured by the *as* RT primer. This result suggests that “long antisense transcripts” span over the *BX2* and *BX2** clusters. At the other end (3'*P*), overlapping sense transcripts can also be detected but only with the *s5* RT primer. RT primers located beyond the *P*-promoter (*s6* and *s1*) were unable to generate *P3'* overlapping transcripts. This suggests that, contrary to antisense transcripts, production of sense transcripts depends on transgene structure, in particular, on the position of the promoter.

Row	Parental cross	Genotype analyzed	% of TSE	n
1	♀ <i>P-1039/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>P-1039/Cy</i>	0.0	1,300
2	♀ <i>w¹¹¹⁸</i> X ♂ <i>P-1039/Cy</i>	+/ <i>P1039</i>	0.0	900
3	♀ <i>Canton^y</i> X ♂ <i>P-1039/Cy</i>	+/ <i>P-1039</i>	0.0	2,200
4	♀ <i>P-1152</i> X ♂ <i>P-1039/Cy</i>	<i>P-1152/+</i> ; +/ <i>P-1039</i>	88.6	2,200
5	♀ <i>6-2</i> X ♂ <i>P-1039/Cy</i>	<i>6-2/P-1039</i>	0.0	2,450
6	♀ <i>1A-6/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>1A-6/P-1039</i>	0.0	2,200
7	♀ <i>6-4/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>6-4/P-1039</i>	0.0	2,350
8	♀ <i>DX1/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>DX1/P-1039</i>	0.0	2,100
9	♀ <i>BX2/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>BX2/P-1039</i>	0.0	2,650
10	♀ <i>C-2/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>C-2/P-1039</i>	0.0	2,100
11	♀ <i>T-1/Cy</i> X ♂ <i>P-1039/Cy</i>	<i>T-1/P-1039</i>	100.0	2,600
12	♀ <i>BC69/Cy</i> X ♂ <i>BC69/Cy</i>	<i>BC69/Cy</i>	0.0	1,700
13	♀ <i>w¹¹¹⁸</i> X ♂ <i>BC69/Cy</i>	+/ <i>BC69</i>	0.0	2,550
14	♀ <i>Canton^y</i> X ♂ <i>BC69/Cy</i>	+/ <i>BC69</i>	0.0	2,300
15	♀ <i>P-1152</i> X ♂ <i>BC69/Cy</i>	<i>P-1152/+</i> ; +/ <i>BC69</i>	54.2	2,602
16	♀ <i>6-2</i> X ♂ <i>BC69/Cy</i>	<i>6-2/BC69</i>	0.0	2,200
17	♀ <i>1A-6/Cy</i> X ♂ <i>BC69/Cy</i>	<i>1A-6/BC69</i>	0.0	1,350
18	♀ <i>6-4/Cy</i> X ♂ <i>BC69/Cy</i>	<i>6-4/BC69</i>	0.0	1,150
19	♀ <i>DX1/Cy</i> X ♂ <i>BC69/Cy</i>	<i>DX1/BC69</i>	0.0	1,600
20	♀ <i>BX2/Cy</i> X ♂ <i>BC69/Cy</i>	<i>BX2/BC69</i>	0.0	1,750
21	♀ <i>C-2/Cy</i> X ♂ <i>BC69/Cy</i>	<i>C-2/BC69</i>	0.0	1,900
22	♀ <i>T-1/Cy</i> X ♂ <i>BC69/Cy</i>	<i>T-1/BC69</i>	100.0	3,200
23	♀ <i>BQ16</i> X ♂ <i>BQ16</i>	<i>BQ16/BQ16</i>	0.0	1,400
24	♀ <i>w¹¹¹⁸</i> X ♂ <i>BQ16</i>	+/ <i>BQ16</i>	0.0	2,750
25	♀ <i>Canton^y</i> X ♂ <i>BQ16</i>	+/ <i>BQ16</i>	0.0	2,500
26	♀ <i>P-1152</i> X ♂ <i>BQ16</i>	<i>P-1152/+</i> ; +/ <i>BQ16</i>	85.6	1,800
27	♀ <i>6-2</i> X ♂ <i>BQ16</i>	<i>6-2/+</i> ; +/ <i>BQ16</i>	0.0	1,300
28	♀ <i>1A-6/Cy</i> X ♂ <i>BQ16</i>	<i>1A-6/+</i> ; +/ <i>BQ16</i>	0.0	1,350
29	♀ <i>6-4/Cy</i> X ♂ <i>BQ16</i>	<i>6-4/+</i> ; +/ <i>BQ16</i>	0.0	1,200
30	♀ <i>DX1/Cy</i> X ♂ <i>BQ16</i>	<i>DX1/+</i> ; +/ <i>BQ16</i>	0.0	1,050
31	♀ <i>BX2/Cy</i> X ♂ <i>BQ16</i>	<i>BX2/+</i> ; +/ <i>BQ16</i>	0.0	1,450
32	♀ <i>C-2/Cy</i> X ♂ <i>BQ16</i>	<i>C-2/+</i> ; +/ <i>BQ16</i>	0.0	1,200
33	♀ <i>T-1/Cy</i> X ♂ <i>BQ16</i>	<i>T-1/+</i> ; +/ <i>BQ16</i>	100.0	2,250

Supplementary Table 1. Trans-Silencing by transgene clusters at the 50C site on chromosome

2. TSE induced by the clusters was tested using three different *P-lacZ* TSE reporter transgenes expressed in the germline (*BQ16*, *BC69* and *P-1039*)^{4,6}. After *lacZ* staining of ovaries, TSE was scored as the percentage of repressed egg chambers among the total number (n) of egg chambers analyzed⁶. *LacZ* silencing of the TSE reporters was determined in the presence (*P-1152*) or absence (*Canton^y* and *w¹¹¹⁸*) of the strong telomeric silencer as positive and negative controls, respectively. *T-1* was the only transgene cluster showing TSE capacities. These capacities appear complete on all three TSE reporters and are greater than those of the strongest *P-1152* telomeric silencer. Cy females produced by the cross of *T-1/Cy* females with *BQ16* males did not exhibit silencing (TSE= 0% , n= 2300) showing that inheritance of cytoplasm without the cluster is not sufficient to induce silencing.

Library ID	#1	#2	#3	#4	#5	#6	#7	#8
Parental origin	♀ <i>T-1/+</i> x ♂ <i>Cn^y</i>	♀ <i>Cn^y</i> x ♂ <i>T-1/+</i>	♀ <i>P-1152</i> x ♂ <i>Cn^y</i>	♀ <i>BX2/+</i> x ♂ <i>BX2/+</i>	♀ <i>P-1152</i> x ♂ <i>P-1152</i>	♀ <i>BX2⁺G1</i> x ♂ <i>Canton^y</i>	♀ <i>BX2⁺G41</i> x ♂ <i>Canton^y</i>	♀ <i>BX2⁺G35</i> x ♂ <i>Canton^y</i>
Female genotype	<i>T-1/+</i>	<i>+T-1</i>	<i>P-1152/+</i>	<i>BX2/+</i>	<i>P-1152/P-1152</i>	<i>BX2⁺G2/+</i>	<i>BX2⁺G42/+</i>	<i>BX2⁺G36/+</i>
miRs	8.79%	9.97%	12.16%	9.62%	12.87%	24.30%	20.12%	18.26%
Ribosomal, tRNA, snoRNA	37.95%	32.21%	29.27%	29.67%	27.41%	12.71%	11.98%	21.41%
ncRNA	0.62%	0.67%	0.60%	0.98%	0.79%	0.92%	0.89%	0.71%
Transposons	39.51%	42.85%	43.70%	45.66%	43.38%	42.85%	44.36%	38.98%
Introns	4.94%	5.03%	5.52%	5.99%	6.54%	10.14%	12.98%	12.07%
Transcripts	1.50%	1.63%	1.83%	1.49%	1.89%	2.00%	2.12%	1.85%
Intergenic	5.81%	6.69%	6.18%	5.96%	6.37%	6.36%	6.78%	6.02%
Others	0.87%	0.94%	0.74%	0.64%	0.75%	0.72%	0.77%	0.69%
Dm Matched	3,769,248	3,190,201	4,499,407	5,154,924	3,185,733	15,759,141	15,301,238	17,264,174
Dm UM	121,806	100,174	194,876	166,523	116,695	272,394	322,721	329,525
Total number of reads	3,891,054	3,290,375	4,694,283	5,321,447	3,302,428	16,031,535	15,623,959	17,593,699
Normalization factor	1.5438525	1.6694864	1.1329814	1	1.5630228	0.9794828	1	0.9916008

Supplementary Table 2. Annotation of small RNA libraries. Small RNAs were prepared from ovaries of females of the indicated genotype (Female genotype row). The parental origin of these females is indicated in the Parental origin row. The Canton^y strain (Cn^y) is devoid of any transgene. Values for the different categories of sequences are expressed as their percentage of the total number of sequence reads that matched the *D. melanogaster* genome 5.43 in each library (Dm Matched row). Normalization factors used for library comparisons are indicated in the table. Libraries #1 to #5 and #6 to #8 were treated as 2 separate data subsets because they were prepared using different adapters and sequenced using different Illumina sequencing technologies (see experimental procedures).

Library #4 and library #7 were then taken as references to normalize libraries #1 to #5 and #6 to #8, respectively.

Cluster	Number of transgenes	G ₁	G ₂	G ₅	G _{≥10}
6-2	1	+ ^v (27)	- (50)	- (15)	- (15) [G ₁₀]
1A-6	2	+ (24)	+/- (19/24)	- (10)	- (10) [G ₁₀]
6-4	4	+ (26)	+ (50)	+/- (7/4)	- (10) [G ₁₅]
DX1	6	+ (19)	+ (44)	+/- (71/5)	+/- + ^v (18/11/4) [G ₂₅]
C2	7	+ (14)	+ (8)	+ (10)	+ (10) [G ₁₀]
BX2	7	+ (32)	+ (71)	+ (100)	+ (46) [G ₂₅]

Supplementary Table 3. Paramutation of other *P{lacW}* clusters allelic to *T-1* and *BX2*. Males carrying the indicated *P{lacW}* cluster were crossed with *T-1* females and lines were established as previously done for *BX2** lines (see Fig. 3a). TSE tests were performed at various generations, including G₁, G₂, G₅ and G_{≥10}, and were scored as indicated in Fig. 1. + indicates that all tested females showed complete TSE in ovaries. +^v indicates that all females showed partial TSE resulting in variegation of lacZ staining among egg chambers. +/- indicates co-occurrence of females showing either complete TSE or no TSE. +/-+^v indicates that females showed all three phenotypes, *i.e.* complete TSE, no TSE and partial TSE. Counts for the phenotypic classes are given in parentheses. The results obtained with *BX2* are reported again for comparison. All transgene repeats acquired the capacity to mediate a complete TSE, whereas the *6-2* single copy transgene was only able to mediate partial TSE. The TSE capacity of the 7-copy *C2* repeat was stably transmitted over several generations, whereas those of the *6-2*, *1A-6*, *6-4* and *DX1* clusters, carrying 1, 2, 4 and 6 transgene copies, respectively, were not maintained. The TSE capacity associated with *6-2*, *1A-6* and *6-4* was lost in G₂, G₅ and G₁₀, respectively. The one associated to *DX1* started to segregate in G₅ but still could be observed in individuals at G₂₅.

Line name	Cytological location	Transgene(s) genotype(s)	Transgene structure	Transgene flybase ID	Transgene reference	Insertion flybase ID	Insertion reference
<i>P-1152</i>	1A	<i>P{ry[+t7.2]=lArB}A171.1F1</i>	<i>P{lArB}</i>	FBtp0000160	(6)	FBti0005700	(7)
<i>BC69</i>	35B-C	<i>P{A92}vas[BC69]</i>	<i>P{A92}</i>	FBtp0000154	(8)	FBti0003420	(9, 4)
<i>1A-6</i>	50C	<i>P{w[+mC]=lacW}50C.x2</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0016762	(1)
<i>6-2</i>	50C	<i>P{w[+mC]=lacW}50C</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0016763	(1)
<i>6-4</i>	50C	<i>P{w[+mC]=lacW}50C.x4</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0016765	(1)
<i>BX2</i>	50C	<i>P{w[+mC]=lacW}50C.x7</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0016766	(1)
<i>DX1</i>	50C	<i>P{w[+mC]=lacW}50C.x6</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0022549	(3)
<i>T-1</i>	50C	<i>Ab(2;3)T-1,P{w[+mC]=lacW}50C.x7/Cy</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBti0027547	(2)
<i>C-2</i>	50C	<i>Ab(2LR)C-2,P{w[+mC]=lacW}50C.x7/Cy</i>	<i>P{lacW}</i>	FBtp0000204	(10)	FBab0027548	(2)
<i>P-1039</i>	60B	<i>P{ry[+t7.2]=PZ}elF-5A[01296]/CyO</i>	<i>P{PZ}</i>	FBtp0000210	(11)	FBti0005186	(3)
<i>BQ16</i>	64C	<i>P{A92}fs(3)BQ16[BQ16]</i>	<i>P{A92}</i>	FBtp0000154	(8)	FBti0003435	(9, 4)

Supplementary Table 4. Genotype and references of transgene insertions used. Name of the insertion (column 1); Transgene cytological location on polytene chromosomes (column 2); Insertion symbol (column 3); Transgene structure name as referenced in *Flybase* (column 4); Transgene structure *Flybase* ID (column 5); Reference describing the transgene structure (column 6); *Flybase* ID of the insertion (column 7); Reference describing the insertion (column 8). Grey empty lines separate chromosomes.

Supplementary Discussion

Paramutation reversion rate. Stability of the paramutation at 25°C was measured in *BX2** replicate lines between G₃₂ and G₅₅, using different *P-lacZ* targets. Among 446 females produced by the cross of *BX2** females with males bearing a *P-lacZ* target, 440 females showed 100% TSE, 5 females showed altered TSE (with a phenotype ranging from a single blue egg chamber to a single blue ovariole after *lacZ* staining) and 1 female lost silencing capacities, resulting in an overall TSE of 99.4% (n=22,700). High stability of TSE was similarly observed in other independent *BX2** lines. Taken together, these data indicate a reversion rate below 0.5% *per* generation at 25°C. Some *BX2** lines maintained at 18°C in the stock collections more frequently produced females having lost silencing capacities, suggesting that paramutation stability is temperature-sensitive. Note that *P*-element repression capacities, which have been linked to piRNAs elements, were also found to be temperature-sensitive, being less stable at 18°C than at 25°C¹²⁻¹⁴. *BX2** reversions are consistent with epigenetic modifications and parallel some cases of paramutations in plants¹⁵, but they differ from the *b1* paramutation for which no revertant was observed over decades¹⁶.

Paramutation and *P*-element repression.

Brennecke *et al.*¹⁷ showed that maternal transmission of *P*-element piRNAs appear to stimulate, in the progeny, production of piRNAs by the paternally-inherited *P* strain chromosomes. Without such maternal contribution, this production is very low and three to six generations of maternal transmission of *P*-chromosomes are necessary to restore strong *P*-repression capacities¹⁸. Thus, maternally-inherited *P*-element piRNAs affect mainly the kinetics of *trans*-generational amplification. A similar *trans*-generational kinetic effect was shown for TSE⁵. In striking contrast, the paramutation described here is associated with a new capacity of the paramutated locus to produce a high level of piRNAs and results in a true epigenetic conversion process. The key point is that the naive *BX2* allele has neither the capacity to induce silencing, nor to produce a significant amount of piRNAs on its own, even after a several-year period of maternal transmission. However, upon inheritance of *T-1* cytoplasm alone, without the *T-1* transgene, it takes only a single generation for *BX2* to start producing high levels of piRNAs and showing strong repression capacity stably maintained over many generations. Paramutation could thus play a role, that was not previously suspected, in *P*-element repression dynamics.

Indeed, some natural *Drosophila* populations show strong polymorphism for regulatory telomeric *P*-insertions, which means that females hemizygous for such insertions must be frequent in these populations¹⁹. In these hemizygous females, paramutation of non-telomeric tandem *P*-copies by telomeric copies might be crucial to maintain repression capacities in the progeny that do not inherit the telomeric *P*-elements. Furthermore, in populations with fixed telomeric *P*-elements, paramutation of non-telomeric copies may increase the efficiency and stability of transposable element repression.

Reference

1. Dorer, D. R. & Henikoff, S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* 77, 993-1002 (1994).
2. Dorer, D. R. & Henikoff, S. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* 147, 1181-90 (1997).
3. Ronsseray, S., Boivin, A. & Anxolabehere, D. P-Element repression in *Drosophila melanogaster* by variegating clusters of P-lacZ-white transgenes. *Genetics* 159, 1631-42 (2001).
4. Lemaitre, B., Ronsseray, S. & Coen, D. Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: a model for the P cytotype. *Genetics* 135, 149-60 (1993).
5. Josse, T. et al. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* 3, 1633-43 (2007).
6. Wilson, C. et al. P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes Dev* 3, 1301-13 (1989).
7. Bellen, H. J. et al. P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev* 3, 1288-300 (1989).
8. O'Kane, C. J. & Gehring, W. J. Detection in situ of genomic regulatory elements in *Drosophila*. *Proc Natl Acad Sci U S A* 84, 9123-7 (1987).
9. Higuete, D., Anxolabehere, D. & Nouaud, D. A particular P-element insertion is correlated to the P-induced hybrid dysgenesis repression in *Drosophila melanogaster*. *Genet Res* 60, 15-24 (1992).
10. Bier, E. et al. Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev* 3, 1273-87 (1989).
11. Mlodzik, M. & Hiromi, Y. Enhancer trap method in *Drosophila*: its application to neurobiology. *Gene Expression in Neural Tissues.*, 397-414 (1992).
12. Marin, L. et al. P-Element repression in *Drosophila melanogaster* by a naturally occurring defective telomeric P copy. *Genetics* 155, 1841-54 (2000).
13. Ronsseray, S., Anxolabehere, D. & Periquet, G. Hybrid dysgenesis in *Drosophila melanogaster*: influence of temperature on cytotype determination in the P-M system. *Mol Gen Genet* 196, 17-23 (1984).
14. Ronsseray, S., Lehmann, M., Nouaud, D. & Anxolabehere, D. The regulatory properties of autonomous subtelomeric P elements are sensitive to a Suppressor of variegation in *Drosophila melanogaster*. *Genetics* 143, 1663-74 (1996).
15. Chandler, V. & Alleman, M. Paramutation: epigenetic instructions passed across generations. *Genetics* 178, 1839-1844 (2008).
16. Chandler, V. L. Paramutation: from maize to mice. *Cell* 128, 641-645 (2007).
17. Brennecke, J. et al. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322, 1387-92 (2008).
18. Engels W, R. Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genet. Research* 33, 219-236 (1979).
19. Ronsseray, S., Lehmann, M. & Anxolabehere, D. Copy number and distribution of P and I mobile elements in *Drosophila melanogaster* populations. *Chromosoma* 98, 207-14 (1989).

Ces travaux ont permis de mettre en évidence le premier cas de paramutation stable chez l'animal. Nous avons montré qu'il est possible d'activer *de novo* au locus composé de séquences répétées en tandem et hétérochromatiques, pour la production de piRNAs par transmission maternelle de piRNAs homologues. Ces loci activés *de novo* produisent des piRNAs, de façon dépendante d'Aub mais également des ARNs de 21nt qui pourraient être des siRNAs. Ces piRNAs présentent une signature Ping-Pong et un fort biais de Uracile en position 1 (de Vanssay *et al.* 2012 ; Hermant *et al.* en préparation). Néanmoins, les capacités répressives du locus *BX2** ne sont pas dépendantes de Dicer-2, impliqué dans la biogenèse des siRNAs. Les analyses des petits ARNs ovariens du locus *BX2** en contexte mutant pour *aub* et *Dicer-2* sont incluses dans l'article Hermant *et al.* en préparation.

3/ *Extra-View*: « piRNAs and epigenetic conversion in *Drosophila*. »

Suite à la publication de l'article de Vanssay *et al.* 2012, Stéphane Ronsseray, chef de l'Equipe, a été sollicité afin d'écrire un « *Extra-View* » au sujet du phénomène de paramutation de *BX2*, premier cas de paramutation stable dans le règne animal. J'ai donc participé à l'écriture de ce papier. Le but était de discuter de la paramutation et de proposer de façon plus détaillée un modèle moléculaire.

piRNAs and epigenetic conversion in *Drosophila*

Augustin de Vanssay^{1,†}, Anne-Laure Bougé^{2,‡}, Antoine Boivin¹, Catherine Hermant¹, Laure Teyssset¹, Valérie Delmarre¹, Christophe Antoniewski^{2,§,*}, and Stéphane Ronsseray^{1,*}

¹Epigenetic Repression and Mobile DNA; Laboratoire Biologie du Développement; UMR7622; CNRS-Université Pierre et Marie Curie; Paris, France;

²Drosophila Genetics and Epigenetics; CNRS URA2578; Institut Pasteur; Paris, France; [†]Current affiliation: Epigenetic Regulation of Genome Organisation; Institut Jacques Monod; CNRS; UMR7592; Université Denis Diderot; Paris, France; [‡]Current affiliation: Drosophila Normal and Pathological Neurobiology; INSERM U661; Institut de Génétique Fonctionnelle; Montpellier, France; [§]Current affiliation: Drosophila Genetics and Epigenetics; Laboratoire Biologie du Développement; UMR7622; CNRS-Université Pierre et Marie Curie; Paris, France

Transposable element (TE) activity is repressed in the *Drosophila* germline by Piwi-Interacting RNAs (piRNAs), a class of small non-coding RNAs. These piRNAs are produced by discrete genomic loci containing TE fragments. In a recent publication, we tested for the existence of a strict epigenetic induction of piRNA production capacity by a locus in the *D. melanogaster* genome. We used 2 lines carrying a transgenic 7-copy tandem cluster (*P-lacZ-white*) at the same genomic site. This cluster generates in both lines a local heterochromatic sector. One line (*T-1*) produces high levels of ovarian piRNAs homologous to the *P-lacZ-white* transgenes and shows a strong capacity to repress homologous sequences *in trans*, whereas the other line (*BX2*) is devoid of both of these capacities. The properties of these 2 lines are perfectly stable over generations. We have shown that the maternal transmission of a cytoplasm carrying piRNAs from the first line can confer to the inert transgenic locus of the second, a totally *de novo* capacity to produce high levels of piRNAs as well as the ability to induce homology-dependent silencing *in trans*. These new properties are stably inherited over generations ($n > 50$). Furthermore, the converted locus has itself become able to convert an inert transgenic locus via cytoplasmic maternal inheritance. This results in a stable epigenetic conversion process, which can be performed recurrently—a phenomenon termed paramutation and discovered in Maize

60 y ago. Paramutation in *Drosophila* corresponds to the first stable paramutation in animals and provides a model system to investigate the epigenetically induced emergence of a piRNA-producing locus, a crucial step in epigenome shaping. In this Extra View, we discuss some additional functional aspects and the possible molecular mechanism of this piRNA-linked paramutation.

Piwi-Interacting RNAs are non-coding small RNAs (24–29nt) which repress mobile DNA activity in the *Drosophila* germline,^{1–4} but can also regulate essential genes during development.⁵ In ovary nurse cells, piRNA biogenesis involves 2 main steps. First, a limited number of loci containing fragments of transposable elements (TEs) produce piRNA precursor transcripts, which would be sliced by Zucchini in a surrounding perinuclear structure called the nuage, to produce primary piRNAs loaded on Piwi or Aubergine Argonaute proteins (for reviews see refs. 4, 6, and 7). These primary piRNAs have a strong preference for 5' U. Second, in the nuage, primary piRNAs can initiate an amplification loop that induces slicing of homologous TE transcripts mediated by Aubergine and Ago3.^{8,9} This slicing loop called the ping-pong loop generates complementary sense and antisense secondary piRNAs showing a 10 bp overlap, thus including U1 and A10 positions on complementary strands.^{1,2} In addition, primary piRNAs loaded on Piwi can enter the nucleus and induce Transcriptional Gene Silencing (TGS) of homologous

Keywords: cellular memory, epigenetics, heterochromatin, piRNAs, transposable elements

*Correspondence to: Christophe Antoniewski; Email: christophe.antoniewski@upmc.fr; Stéphane Ronsseray; Email: stephane.ronsseray@upmc.fr

Submitted: 08/30/2013; Revised: 09/16/2013; Accepted: 09/17/2013

<http://dx.doi.org/10.4161/fly.26522>

Extra View to: 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; 490:112–5; PMID:22922650; <http://dx.doi.org/10.1038/nature11416>

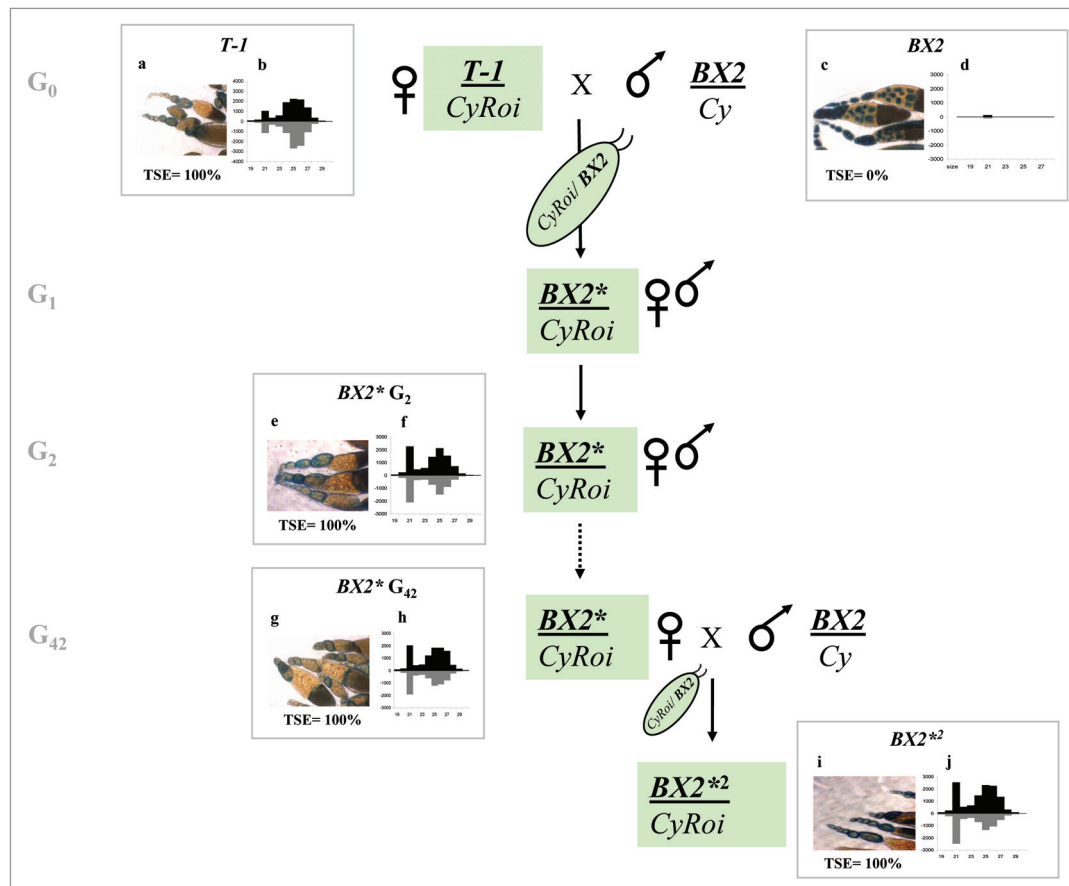


Figure 1. Paramutation linked to maternal transmission of piRNAs. (A, C, E, G, and I) The trans-silencing assay measures the capacity of a tested *P-lacW* transgene cluster to repress a homologous *P-lacZ* transgene target in the female germline. Females carrying the cluster were crossed with males carrying the *P-1039* transgene (not indicated in the mating scheme), which is expressed in both nurse cells and follicle cells. Female progeny of this cross were stained for lacZ activity. Trans-silencing is germline specific and occurs only in the nurse cells and the oocytes, but not in the somatic follicle cells surrounding the egg chamber. The percentage of repressed egg chambers is given as TSE (Trans-Silencing Effect) percentage. In each case, more than 1500 egg chambers were observed. (B, D, F, H, and J) Deep sequencing of small RNAs (19–30nt) from ovaries: histograms show the length-distribution of small RNAs matching the *P-lacW* sequence. Positive and negative values correspond to sense and antisense reads, respectively. The *T-1* and *BX2* lines carry a 7-copy *P-lacZ* transgene cluster, generating a heterochromatin island in the middle of the 2R chromosomal arm (50C). The transgenic locus of the *T-1* line does not show detectable lacZ staining in the germline (not shown) and can silence in trans the expression of a homologous transgene in the germline tissue (A). Moreover, *T-1* produces abundant ovarian piRNAs homologous to the transgenes (B). The transgenic locus of the *BX2* line does neither show lacZ staining in the germline (not shown), nor induce trans-silencing (C), nor produce abundant transgene matching piRNAs (D). These properties are stable over generations. *G2* females, which inherit both a cytoplasm deriving from *T-1* grandmothers (but not the *T-1* cluster) and a *BX2* cluster can induce complete trans-silencing in the germline (E) and produce abundant *P-lacW* homologous piRNAs (F). These properties are stably maintained over generations in derived lines (G and H), called *BX2*/CyRoi*. * symbolizes the epigenetic conversion process and the green background the presence of maternally inherited *P-lacW* homologous small RNAs. The *T-1* cluster thus has paramutagenic properties and the *BX2* cluster is paramutable. Maternal transmission of the cytoplasm from paramutated *BX2*/CyRoi* females can again induce a conversion process similar to that induced by the *T-1* cytoplasm, generating thus a second order paramutation (symbolized by **). *BX2*** females induce trans-silencing (I) and produce abundant ovarian piRNAs (J). We performed the same experimental scheme starting with the reciprocal *G0* cross (Females *BX2* x Males *T-1*). Progeny carrying the *BX2* locus never showed a repression capacity at any generation (data not shown).

sequences, for example euchromatic copies of TEs.^{10–13} Thus, the piRNA pathway appears to be a double mobile DNA repression mechanism that can work in the nucleus via TGS and in the nuage through Post-transcriptional Gene Silencing (PTGS). Extensive work has analyzed the machinery of piRNA biogenesis, but little is known about what makes a piRNA-producing locus, i.e., a locus that is able

to produce primary piRNAs. Two non-exclusive hypotheses are proposed.^{14,26,34} In a genetic determination hypothesis, some sequence motifs could be the causal factor for establishing a piRNA-producing locus. Evidence suggests that this could occur in *C.elegans*.¹⁵ In an epigenetic determination hypothesis, the causal actor could be cytoplasmically transmitted piRNAs. We have tested whatever maternal inheritance

of piRNAs in *Drosophila* can start de novo the production of abundant piRNAs by a locus which was previously quiescent for this production over generations.¹⁶

We have taken advantage of a somatic phenomenon termed Repeat Induced Gene Silencing (RIGS), in which a tandem array of *P-lacZ-white* (*P-lacW*) transgenes can generate a local heterochromatic sector in a previously euchromatic

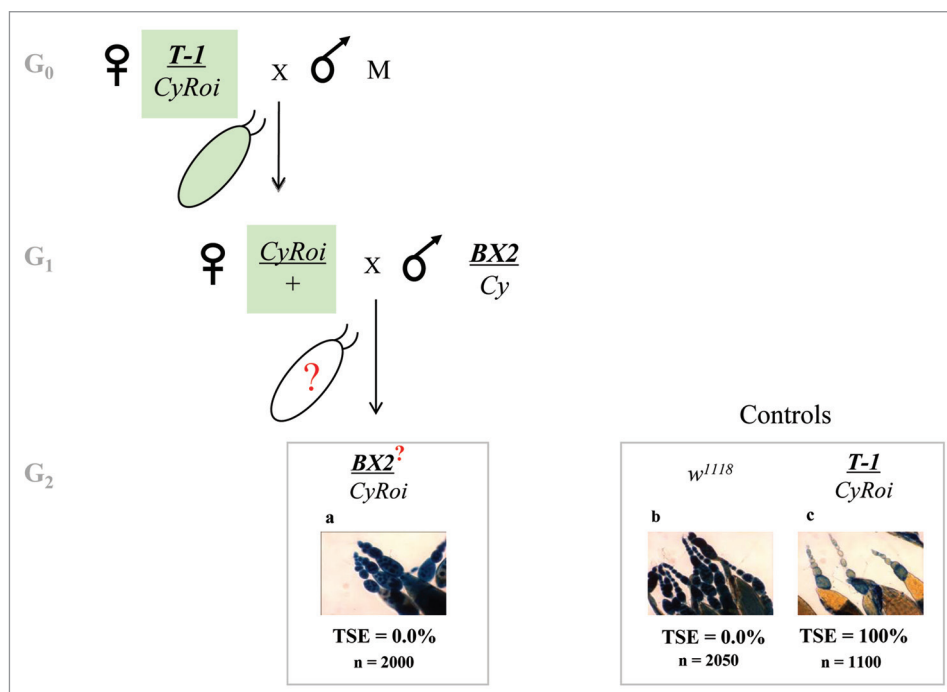


Figure 2. Assay for self-maintaining cytoplasmic components. We generated G₁ females having inherited the cytoplasm of a *T-1* mother but no transgene cluster from any parent. The capacity of these G₁ females to induce paramutation was tested by crossing them with *BX2* males and analyzing silencing capacities of G₂ females carrying the *BX2* cluster. When tested using the TSE assay (cross with *P-1039* males), these G₂ females conferred no silencing capacity to daughter ovaries (A). Therefore, G₁ females, which maternally inherit piRNAs but do not carry a homologous locus allowing zygotic piRNA biogenesis, cannot induce epigenetic conversion of their progeny. TSE was tested as in Figure 1. M lines are devoid of any *P*-transgene. Negative and positive TSE controls were performed by crossing *w¹¹¹⁸* (M line) and *T-1* females with *P-1039* males, respectively, and staining daughter ovaries.

region.^{17,18} Indeed, such clusters generate a strong de novo HP1 binding site on larval polytene chromosomes, and cause a strong variegation of the *white* marker in the adult eye.¹⁹ The intensity of the phenomenon is linked to the number of transgene repeats: 7-copy clusters induce a strong phenomenon whereas a 4-copy cluster shows a weaker effect.^{17,18} In addition, chromosomal rearrangements induced by X-rays can increase the strength of RIGS. We showed previously that the *T-1* line, carrying 7 tandem transgenes, can induce a homology-dependent silencing of a single *P-lacZ* transgene expressed in the female germline.²⁰ This silencing was termed Trans-Silencing Effect (TSE) and was discovered in the course of *P*-element repression studies.^{21,22} TSE was shown to be strongly sensitive to mutants affecting the piRNA pathway and to be linked to maternal piRNA transmission.^{23–26} By contrast, a line called *BX2*, also carrying a 7-copy *P-lacW* tandem cluster at the

same locus, did not induce TSE.²⁰ *T-1* was derived from *BX2* following X-ray treatment, and variegation of *T-1* is stronger than that of *BX2*. We followed silencing capacities of the *T-1* and *BX2* lines over more than 1 decade and these capacities were perfectly stable. We further questioned whether the *BX2* cluster could be induced to produce abundant piRNAs by the cytoplasm of *T-1*. We generated females having maternally inherited the cytoplasm of the *T-1* line—but not the *T-1* cluster—and paternally inherited the *BX2* cluster. Surprisingly, these G₁ females were capable to establish a complete silencing of a homologous *P-lacZ* transgene.¹⁶ By brother-sister mating, we established lines, called *BX2** lines, the * indicating the initial contact with *T-1* cytoplasm. The silencing capacities of *BX2** lines were maintained over generations (n > 50) (Fig. 1). TSE assays performed recently confirmed this stability at the 80th generation (data not shown).

Finally, the *BX2** lines were shown to have themselves the capacity to convert a “naïve” original *BX2* locus into a strong piRNA producing locus, via *BX2** cytoplasmic transmission. Therefore, such an epigenetic conversion can be recurrent. Deep sequencing analysis of ovarian small RNAs showed that abundant piRNA homologous to *P-lacW* transgenes are found in *T-1* ovaries, but not in ovaries from the initial naïve *BX2* line. By contrast, *BX2** ovaries showed abundant piRNAs homologous to *P-lacW* transgenes (Fig. 1). Thus, a *T-1* cytoplasm, which contains transgene derived piRNAs, can confer to the quiescent *BX2* locus the capacity to produce abundant piRNAs, and the conversion process can be repeated (as tested 5 times).¹⁶ Such a stable and recurrent epigenetic conversion is called paramutation as previously defined: an interaction between 2 alleles of a single locus, resulting in a heritable change of 1 allele that is induced by the other allele.^{27–29} *T-1* is the paramutagenic locus and *BX2* is the paramutable locus. Furthermore the paramutated *BX2** locus has become paramutagenic (compare the *BX2**² lines, Fig. 1). Paramutation was discovered in Maize.^{27,28} A pairing step between the paramutagenic and paramutated loci was proposed to be involved in paramutation induction.^{29–31} In the *Drosophila* paramutation described here, no pairing between the 2 loci can take place since *T-1* and *BX2* clusters are never in the same zygote. The cytoplasm is the vector of the phenomenon, showing that paramutation can occur without a pairing step between the partner loci. We have checked that these results could not be caused by a strictly cytoplasmic inheritance of a component, unrelated to the transgene cluster but initially present only in the *T-1* line and not in *BX2*. We have thus performed a cytoplasmic component self-maintenance experiment in which females maternally received *T-1* cytoplasm but inherited neither the *T-1* cluster nor any other homologous cluster (Fig. 2). Such G₁ females were tested for their ability to confer trans-silencing capacities to a naïve *BX2* locus via cytoplasmic

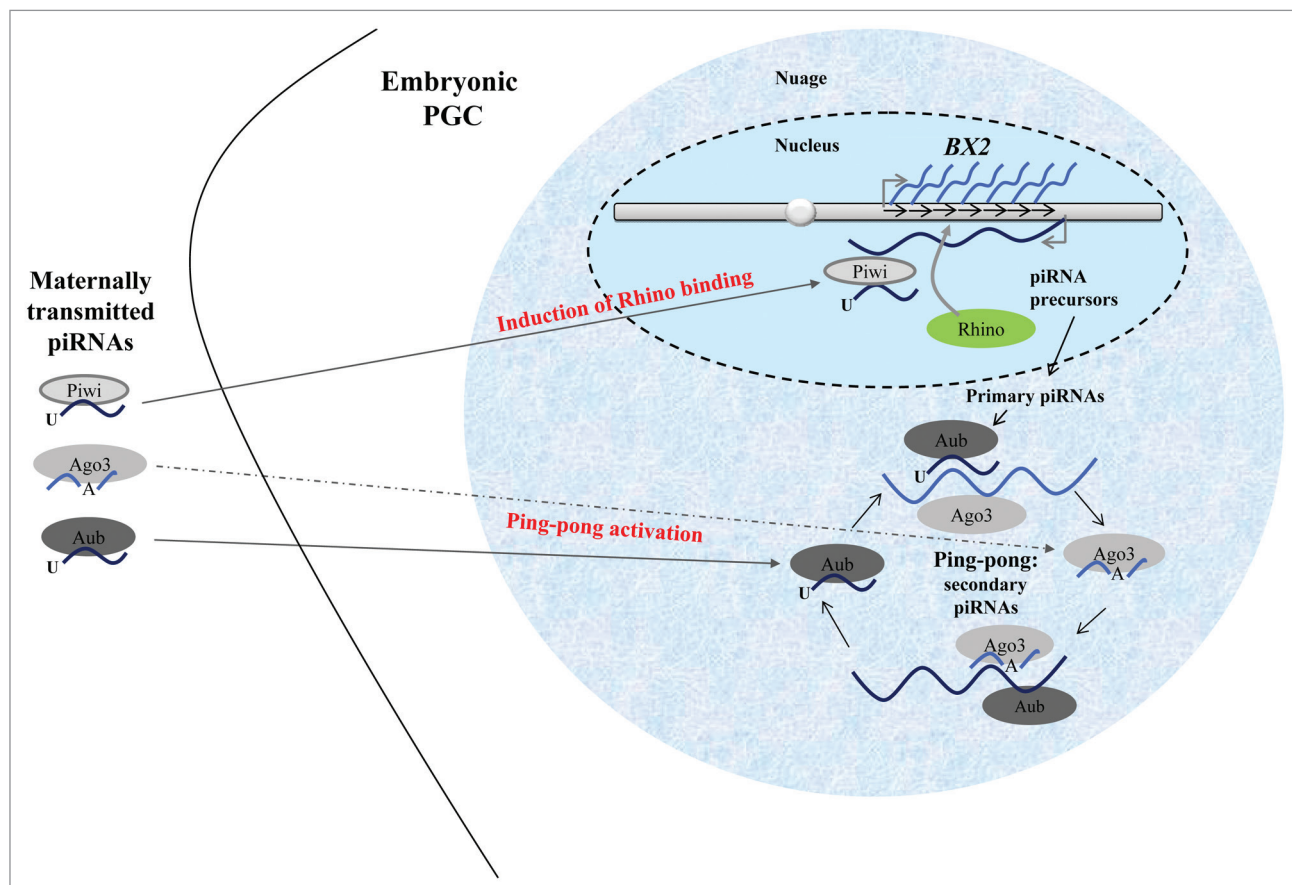


Figure 3. Models for piRNA induced paramutation. Maternal deposition of piRNAs in the embryo activates abundant piRNA production by the *BX2* locus. A first model could be that maternal piRNAs associated to Piwi enter into the nuclei of primordial germ cells (PGC) and interact with *BX2* nascent transcripts. This targeting could modify the chromatin state of the locus (HP1 to Rhino switch), leading to increased production of primary piRNAs and, consequently, of secondary piRNAs in the nuage. Alternatively, maternally transmitted piRNAs associated with Aub or Ago3 could directly activate ping-pong amplification of piRNAs in the nuage. AGO3 maternal transmission is shown with a dashed arrow since some observations suggest that it is less efficient than that of Piwi or Aub.¹

inheritance. G_1 females were crossed to *BX2* males, and G_2 females having paternally inherited a *BX2* cluster were crossed with males carrying a target *P-lacZ* transgene. No repression was observed in female progeny ovaries, showing that without a chromosomal copy of the cluster in G_1 females no persistence of the *T-1* cytoplasmic inheritance was detected in the lineage. Thus, the new properties of *BX2** lines cannot be explained by simple inheritance of auto-replicative cytoplasmic factors not related to the transgene cluster. By contrast, these data are in accordance with a piRNA model since, without piRNA producing loci in G_1 females, piRNAs are expected to be diluted or destroyed in the zygote at the adult stage. Moreover, this result is fully consistent with the fact that when crossing hemizygous *T-1* females with males

carrying a *P-LacZ* transgene expressed in the adult ovaries, female progeny carrying *T-1* show a complete silencing whereas their sisters which have not inherited the cluster do not show any repression (Sup. Fig. 5 of ref. 16).

What could be the mechanism of this epigenetic conversion process? Maternally inherited piRNAs could possibly act in the embryo either on the chromosomal copy of the transgenic cluster or on the piRNA machinery components (Fig. 3). In the first scheme, piRNAs loaded on Piwi, and transported to the posterior pole of the oocyte, could enter the nuclei of embryo pole cells. Furthermore, these piRNAs could interact with nascent transcripts of the cluster in the primordial germ cells (PGC), and modify their chromatin state. This could be mediated by recruitment of the HP1-homolog protein

Rhino, known to be involved in piRNA production,³² which could compete with HP1a for binding methylated histone H3. Following Rhino binding, transcripts from the cluster could be targeted to the piRNA machinery. In second scheme, maternally transmitted piRNAs loaded on Aub or Ago3 could directly activate the ping-pong loop³³ which could target the transcripts produced by the cluster. These 2 epigenetic mechanisms are non-exclusive. One future key point will be to compare the chromatin state of *BX2* and *BX2**. Finally, the epigenetic conversion observed here shows that piRNA-producing loci can de novo be trans-activated by piRNAs and then remain activated. Hence epigenetic induction has a key role in emergence of piRNA producing loci.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

This work was supported by fellowships from the Ministère de l'Enseignement Supérieur et de la Recherche to de Vanssay A and Hermant C; from the Fondation pour la Recherche Médicale to de Vanssay A; from the Association Nationale de la Recherche (ANR) to Bougé AL; and by grants from the Fondation ARC pour la Recherche contre le Cancer to Ronsseray S; and from the ANR (project "Nuclear endosRNAs") to Antoniewski C. We thank Nature Publishing Group for permitting the use of piRNA distribution graphs in **Figure 1**. We thank Neel Randsholt and Anne-Marie Pret for their help in the preparation of the manuscript.

References

- Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 2007; 128:1089-103; PMID:17346786; <http://dx.doi.org/10.1016/j.cell.2007.01.043>
- Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 2007; 315:1587-90; PMID:17322028; <http://dx.doi.org/10.1126/science.1140494>
- Guzzardo PM, Muerdter F, Hannon GJ. The piRNA pathway in flies: highlights and future directions. *Curr Opin Genet Dev* 2013; 23:44-52; PMID:23317515; <http://dx.doi.org/10.1016/j.gde.2012.12.003>
- Ishizu H, Siomi H, Siomi MC. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev* 2012; 26:2361-73; PMID:23124062; <http://dx.doi.org/10.1101/gad.203786.112>
- Rouget C, Papin C, Boureux A, Meunier AC, Franco B, Robine N, Lai EC, Pelisson A, Simonelig M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* 2010; 467:1128-32; PMID:20953170; <http://dx.doi.org/10.1038/nature09465>
- Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, Bonnefond L, Matsumoto N, Nishizawa T, Nakanaga K, Aoki J, et al. Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* 2012; 491:284-7; PMID:23064230; <http://dx.doi.org/10.1038/nature11509>
- Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L, Hannon GJ. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* 2012; 491:279-83; PMID:23064227; <http://dx.doi.org/10.1038/nature11502>
- Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 2009; 137:522-35; PMID:19395010; <http://dx.doi.org/10.1016/j.cell.2009.03.040>
- Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, et al. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 2009; 137:509-21; PMID:19395009; <http://dx.doi.org/10.1016/j.cell.2009.04.027>
- Sienski G, Dönertas D, Brennecke J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* 2012; 151:964-80; PMID:23159368; <http://dx.doi.org/10.1016/j.cell.2012.10.040>
- Wang SH, Elgin SC. *Drosophila* Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proc Natl Acad Sci U S A* 2011; 108:21164-9; PMID:22160707; <http://dx.doi.org/10.1073/pnas.1107892109>
- Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, Perkins EM, Hur JK, Aravin AA, Tóth KF. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* 2013; 27:390-9; PMID:23392610; <http://dx.doi.org/10.1101/gad.209841.112>
- Huang XA, Yin H, Sweeney S, Raha D, Snyder M, Lin H. A major epigenetic programming mechanism guided by piRNAs. *Dev Cell* 2013; 24:502-16; PMID:23434410; <http://dx.doi.org/10.1016/j.devcel.2013.01.023>
- Luteijn MJ, Ketting RF. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet* 2013; 14:523-34; PMID:23797853; <http://dx.doi.org/10.1038/nrg3495>
- Gu W, Lee HC, Chaves D, Youngman EM, Pazour GJ, Conte D Jr., Mello CC. CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. *Cell* 2012; 151:1488-500; PMID:23260138; <http://dx.doi.org/10.1016/j.cell.2012.11.023>
- de Vanssay A, Bougé AL, Boivin A, Hermant C, Teyssier L, Delmarre V, Antoniewski C, Ronsseray S. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* 2012; 490:112-5; PMID:22922650; <http://dx.doi.org/10.1038/nature11416>
- Dorer DR, Henikoff S. Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* 1994; 77:993-1002; PMID:8020105; [http://dx.doi.org/10.1016/0092-8674\(94\)90439-1](http://dx.doi.org/10.1016/0092-8674(94)90439-1)
- Dorer DR, Henikoff S. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* 1997; 147:1181-90; PMID:9383061
- Panti L, Dorer DR, Berloco M, Henikoff S, Pimpinelli S. Heterochromatin protein 1 binds transgene arrays. *Chromosoma* 1998; 107:286-92; PMID:9880761; <http://dx.doi.org/10.1007/s004120050310>
- Ronsseray S, Boivin A, Anxolabéhère D. P-Element repression in *Drosophila melanogaster* by variegating clusters of P-lacZ-white transgenes. *Genetics* 2001; 159:1631-42; PMID:11779802
- Roche SE, Rio DK. Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the *Drosophila* Polycomb group gene, Enhancer of zeste. *Genetics* 1998; 149:1839-55; PMID:9691041
- Ronsseray S, Josse T, Boivin A, Anxolabéhère D. Telomeric transgenes and trans-silencing in *Drosophila*. *Genetica* 2003; 117:327-35; PMID:12723712; <http://dx.doi.org/10.1023/A:1022929121828>
- Josse T, Maurel-Zaffran C, de Vanssay A, Teyssier L, Todeschini AL, Delmarre V, Chaminade N, Anxolabéhère D, Ronsseray S. Telomeric trans-silencing in *Drosophila melanogaster*: tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS One* 2008; 3:e3249; PMID:18813361; <http://dx.doi.org/10.1371/journal.pone.0003249>
- Josse T, Teyssier L, Todeschini AL, Sidor CM, Anxolabéhère D, Ronsseray S. Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* 2007; 3:1633-43; PMID:17941712; <http://dx.doi.org/10.1371/journal.pgen.0030158>
- Todeschini AL, Teyssier 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; PMID:20559422; <http://dx.doi.org/10.1371/journal.pone.0011032>
- Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 2008; 322:1387-92; PMID:19039138; <http://dx.doi.org/10.1126/science.1165171>
- Brink RA. A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics* 1956; 41:872-89; PMID:17247669
- Coe EH. A regular and continuing conversion-type phenomenon at the B locus in Maize. *Proc Natl Acad Sci U S A* 1959; 45:828-32; PMID:16590451; <http://dx.doi.org/10.1073/pnas.45.6.828>
- Chandler VL. Paramutation: from maize to mice. *Cell* 2007; 128:641-5; PMID:17320501; <http://dx.doi.org/10.1016/j.cell.2007.02.007>
- Stam M. Paramutation: a heritable change in gene expression by allelic interactions in trans. *Mol Plant* 2009; 2:578-88; PMID:19825640; <http://dx.doi.org/10.1093/mp/ssp020>
- Suter CM, Martin DIK. Paramutation: the tip of an epigenetic iceberg? *Trends Genet* 2010; 26:9-14; PMID:19945764; <http://dx.doi.org/10.1016/j.tig.2009.11.003>
- Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, et al. The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 2009; 138:1137-49; PMID:19732946; <http://dx.doi.org/10.1016/j.cell.2009.07.014>
- Grentzinger T, Armenise C, Brun C, Mugat B, Serrano V, Pelisson A, Chambeyron S. piRNA-mediated transgenerational inheritance of an acquired trait. *Genome Res* 2012; 22:1877-88; PMID:22555593; <http://dx.doi.org/10.1101/gr.136614.111>
- Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, Gordon A, Hannon GJ, Aravin AA. Production of artificial piRNAs in flies and mice. *RNA* 2012; 18: 42-52; PMID 22096018; <http://dx.doi.org/10.1261/rna.029769.111>

Dans cet « *Extra-View* », nous avons proposé un modèle afin d'expliquer le mécanisme moléculaire de la paramutation qui est ici l'initiation d'un cluster de piRNAs. Nous proposons qu'il y ait un dépôt maternel de piRNAs dans l'embryon par la mère paramutagénique. Ces piRNAs chargés par Piwi, Aub et AGO3 peuvent agir de plusieurs façons. Nous proposons que les piRNAs chargés par Piwi puissent entrer dans le noyau des cellules de la future lignée germinale et permettre le recrutement du complexe RDC au niveau du cluster qui va pouvoir produire ainsi des précurseurs de piRNAs de façon efficace et initier la production de piRNAs primaires puis secondaires. Nous proposons également, et ce de façon non-exclusive, que les piRNAs chargé par Aub et AGO3 puissent entrer dans ces même cellules au niveau du nuage afin d'initier la production de piRNAs secondaires par le Ping-Pong. Il serait donc possible que l'ensemble du phénomène d'activation *de novo* du locus *BX2** ne soit que cytoplasmique. Ainsi, lors de la transmission maternelle de piRNAs dans l'embryon portant uniquement le locus *BX2^{naïf}*, il est possible que la production des piRNAs se fasse uniquement par le mécanisme d'amplification du Ping-Pong. Puis cette production de piRNAs secondaires serait maintenue au cours de générations uniquement par le Ping-Pong. Cela signifie que le locus *BX2** ne produirait pas de piRNAs primaires. Nous avons donc testé l'effet des mutations de gènes impliqués de la biogenèse des piRNAs primaires et secondaires afin de savoir si le locus *BX2** produisait des piRNAs primaires et secondaires (Hermant *et al.* en préparation). Des premiers éléments de réponses nous ont été apportés par l'analyse des piRNAs issus du locus *BX2** qui montre qu'il y a un biais d'Uracile en position 1 des piRNAs (de Vanssay *et al.* 2012). Il s'agit d'un biais observé lors de la production de piRNAs primaires et secondaires, puisque la majorité des piRNAs chargés par Piwi et Aub ont un Uracile en position 1 (Brennecke *et al.* 2007). Or, Piwi charge essentiellement des piRNAs primaires (Brennecke *et al.* 2007).

VII/ Etudes des propriétés fonctionnelles et moléculaires d'un locus producteur de piRNAs paramuté.

Il s'agit donc du premier cas décrit de paramutation stable à travers les générations dans le monde animal. Celle-ci est corrélée à l'acquisition par le locus paramuté de la capacité à produire des piRNAs et à réprimer un transgène cible. Nous avons poursuivi l'analyse des propriétés génétiques et moléculaires de cette paramutation et de ses conséquences sur la régulation de l'expression du génome. C'est l'objet de l'article que je propose que je soumetts en tant que premier auteur. Dans cet article, nous avons analysé le locus *BX2** sur plusieurs aspects fonctionnels et moléculaires. Nous avons continué l'étude du locus *BX2** en contexte mutant pour *aub* et *Dicer-2* par l'analyse de leurs petits ARNs dont j'ai mené l'analyse bio-informatique à l'aide du serveur *Galaxy*, mis à disposition par Christophe Antoniewski. Ensuite, nous avons continué l'analyse fonctionnelle et moléculaire de ce cluster en testant l'effet de mutant de gènes impliqués dans la biogenèse primaire des piRNAs : *rhino*, *cutoff*, *zucchini*. Cette analyse visait à répondre à la question de l'existence d'un phénomène nucléaire dans la paramutation. Autrement dit, est ce que le locus *BX2** paramuté produit des piRNAs primaires ? Un laboratoire concurrent a montré récemment par ChIP Q-PCR, qu'il y a un enrichissement de deux fois de la quantité de la marque H3K9me3 quand *BX2* est paramuté, ce qui signifie qu'il existe une différence chromatienne entre *BX2^{Naïf}* et *BX2** (Le Thomas *et al.* 2014). De plus, il a été montré dans cet article que les capacités répressives de *T-1*, locus paramutagénique, sont associées à une augmentation sur le cluster des marques H3K9me3 mais également des protéines Rhino et Cutoff (Le Thomas *et al.* 2014).

A plusieurs reprises, il a été constaté que les clusters de piRNAs produisaient également des siRNAs. Nous nous sommes demandé si ces siRNAs pouvaient être requis pour la paramutation. J'ai donc testé une paramutation avec uniquement des piRNAs afin de savoir si une paramutation stable est obtenue et quel profil de petits ARNs ovariens allait produire ce locus *BX2**.

Lors de la mise en évidence de la paramutation, il avait été testé une paramutation partiellement homologue avec comme allèle paramutagénique, un silencer télomérique (de Vanssay *et al.* 2012). Cette paramutation s'est montrée être instable au cours des générations. J'ai entrepris de comprendre pourquoi en reproduisant des expériences de paramutations partiellement homologues dans plusieurs conditions afin de déterminer si l'instabilité était due à la localisation génomique des transgènes paramutagéniques ou due au fait que ces transgènes ne sont que partiellement homologue à *BX2*. J'ai également analysé la production

des piRNAs ovariens produits par ces loci *BX2** paramutés par des loci partiellement homologues et testés leurs fonctionnalités.

Enfin, nous avons entrepris d'analyser l'impact somatique d'un héritage maternel de piRNAs dans l'œil. En effet, Augustin de Vanssay, puis nous, avons remarqué que la répression par les piRNAs présentait un effet maternel, non seulement dans les ovaires mais également dans les yeux. La présence de piRNAs et leurs actions dans les têtes de drosophiles ont déjà été mis en évidence (Yan *et al.* 2011 ; Perrat *et al.* 2013). Nous avons analysé l'impact des piRNAs transmis maternellement sur l'œil adulte.

1/ « Paramutation in *Drosophila* involves the primary and secondary piRNA pathways and can induce a rapid spreading within a locus for piRNA production capacity »

Paramutation in *Drosophila* involves the primary and secondary piRNA pathways and can induce a rapid spreading within a locus for piRNA production capacity.

**Catherine HERMANT^{1,2}, Antoine BOIVIN^{1,2}, Laure TEYSSET^{1,2}, Valérie DELMARRE^{1,2},
Amna ASIF-LAIDIN^{1,2}, Christophe ANTONIEWSKI^{1,3}, Stéphane RONSSERAY^{1,2*}**

¹Sorbonne Universités, UPMC Univ. Paris 06, IBPS-UMR7622, Laboratoire Biologie du Développement, F-75005, Paris, France

²CNRS, UMR7622, “Epigenetic repression and mobile DNA”, F-75005, Paris, France

³CNRS, UMR7622, “Drosophila Genetics and Epigenetics”, F-75005, Paris, France

* To whom correspondence should be addressed: stephane.ronsseray@upmc.fr

ABSTRACT

Transposable element (TE) activity is repressed in animal germline by PIWI-Interacting RNAs (piRNAs), a class of small non-coding RNAs produced by discrete genomic loci mostly composed of TE fragments. The mechanism of emergence of piRNAs producing loci is still enigmatic. We have shown that in *Drosophila melanogaster* a tandemly repeated transgenic locus (*P-lacZ-white* cluster), which was previously stably quiescent for ovarian piRNA production, can be activated for such production by maternal inheritance of a cytoplasm containing homologous piRNAs. This new capacity is further long-term stably transmitted over generations and is associated to the capacity to silence in *trans* the expression of a homologous transgenic target in the female germline. Such an epigenetic conversion presents all the functional characteristics of a paramutation, *i.e.*, a heritable epigenetic modification of one allele by the other. At the molecular level, it was shown that such cluster conversion can be correlated to enrichment in histone 3 lysine 9 trimethylation (H3K9me3). We report here that piRNA production and *trans*-silencing capacities of the paramutated cluster depend on the genes *rhino*, *cutoff* and *zucchini* which are involved in primary piRNA biogenesis in the germline and depend on *aubergine* involved in the ping-pong piRNA amplification step. We show that the 21nt small RNAs which are produced, in addition to *bona fide* piRNAs, by the paramutated cluster depend on *Dcr-2* and are not necessary for paramutation to occur. Strikingly, these 21nt siRNAs depend also on all the piRNA genes tested. Moreover, cytoplasmic transmission of small RNAs homologous to only a subset of the transgenic locus results in the production of functional small RNAs by the whole length of the transgene, as soon as the third generation. This shows that a piRNA producing locus can undergo a rapid internal *cis*-spreading for piRNA production capacity. Finally, we have observed a somatic impact of maternally inherited small RNAs. Such maternal inheritance leads to strong repression of *P-lacZ-white* transgene expression in progeny eyes. In conclusion, *Drosophila* paramutation requires the complete piRNA - but not siRNA - machinery, results in occurrence of ovarian siRNAs as pi-machinery by-products, can induce a rapid *cis*-spreading within a locus for piRNA production capacity and can be correlated to repression in adult somatic tissues.

INTRODUCTION

Genomes have to face the presence of a large fraction of mobile DNA whose activity can result in severe deleterious effects on chromosome stability and gametogenesis. In animal germline, a system of genomic traps exists in which any transposable elements (TEs) can insert and consequently integrate a catalogue of the potentially dangerous sequences that have to be repressed [1,2]. These loci undergo a non-canonical transcription and RNA processing which results in production of non-coding small RNAs having the capacity to target the transcripts of the homologous potentially active TE copies scattered through the genome [3,4]. These small RNAs are called PIWI-interacting RNAs (piRNAs) and repress TE activity at both the transcriptional and post-transcriptional levels. Target repression, as well as biogenesis of piRNA by piRNA producing loci, involves heterochromatin formation: repression of euchromatic TE copies involves HP1 binding [5,6] whereas piRNA producing cluster activity requires the presence of the HP1 homolog Rhino associated with Cutoff and Deadlock on the locus [3,7]. These partners are associated in a so-called “RDC” complex [3]. The factors which induce Rhino binding on the cluster are still unknown.

In *Drosophila*, by contrast to *C. elegans* [8,9], the production of piRNA by a piRNA locus does not appear to be simply genetically determined, *i.e.* no specific sequence motif or structure has been identified which is sufficient to license piRNA production by DNA adjacent to this sequence. Conversely, it appears that this capacity is at least partially epigenetically determined. Indeed, piRNA production by a piRNA locus can be induced by maternal inheritance of homologous piRNAs [10]. Such a complex *trans*-generational behaviour of piRNA loci fits with a number of previously described *trans*-generational complex behaviours of TE repression capacities. For example, two TEs have invaded the *Drosophila melanogaster* genome during the last century [11,12]. The *I* element which transposes *via* a RNA intermediate [13] and the *P* element which transposes *via* a DNA intermediate [14]. For both TEs, repression capacities were shown to be short-term maternally determined but long-term chromosomally determined through generations [15-17]. For the *P* element, it was shown that a strictly extra-chromosomally transmitted component can play a key role in repression establishment [18]. After the discovery of piRNA, maternal transmission of piRNAs was interpreted to represent this cytoplasmically transmitted component. In

fact, *I* and *P* element repression by master repressive loci combines genetic and epigenetic transmissions since when introduced by crosses in a cytoplasm devoid of homologous piRNAs, master loci fail in G_1 to establish strong repression capacities but repression progressively increases over generations and reach a strong level after about 5 generations [19]. The genotype determines a kind of equilibrium state but starting a piRNA locus takes time. Interestingly, time within a given generation can sometime also play a role. Indeed, ageing of females can contribute to increase TE master repressive loci activity [20] but this phenomenon can be modulated by temperature [21]. In any case, maternal inheritance of piRNAs plays a key role on the kinetics of piRNA loci activation over generations.

The next step was the finding of a case of strict epigenetic role of maternally transmitted piRNAs on piRNA locus, not kinetics of activation, but more importantly “intrinsic” specification. Indeed, it was found that a transgenic cluster of *P*-element deriving transgenes (called *BX2*) can be inert for repressive and piRNA production capacity, as tested over dozens of generations despite maternal transmission, but can be *de novo* activated by out-crossing males from this line with females from a line (called *T-1*) that produces abundant transgene homologous piRNAs of both orientations [22]. Once activated, this new piRNA cluster, called *BX2**, is stable over generations ($n > 100$) and is able to produce abundant transgenic sense and antisense piRNAs. Cytoplasm of oocytes produced by these females can again *de novo* activate piRNA production by a paternally inherited inactive cluster [22]. This recurrent epigenetic conversion process presents all the features of a “paramutation” process, a phenomenon previously described in plants and described 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*” [23]. This paramutation results in a kind of communication within a genome between homologous clusters for piRNA production capacity. It can play an important role for TE repression establishment following a genetic invasion since master sites of repression can activate piRNA secondary sites which will reinforce their repressive capacities or even replace them further. It was found for example, that *P* element repression master sites located in subtelomeric heterochromatin can potentialize in *trans* repression capacities of some other *P* elements of the genome [24,25]. Investigating the mechanism of *de novo* activation of piRNA producing loci is therefore a key aspect of the epigenome shaping.

We therefore have analyzed functional and molecular properties of the *BX2** paramutated locus. It was proposed that induction of piRNA production by *BX2** may result from two main *trans*-generational processes [26,27]. A first one is the activation by maternal piRNAs of the so-

called primary piRNA pathway in the nucleus of primordial germ cells (PGCs). Maternally inherited piRNAs associated to the PIWI protein could enter the PGC nucleus and induce RDC binding on the transgenic cluster, resulting in piRNA precursor production by the paramutated cluster. Such “nucleic model” has been strongly supported by the fact that *BX2** locus shows an increase of H3K9 methylation with regard to the *BX2* locus and that maternal inheritance of *T-1* piRNAs increases binding of Rhino and Cutoff on the *T-1* locus itself [26]. A second, but not exclusive, process proposed is the activation by maternal piRNAs of the so-called secondary piRNA pathway [1,2,28] which takes place in the nuage, a diffuse cytoplasmic structure surrounding the germ nucleus. Indeed, upon their arrival in the nuage, piRNA precursors produced by piRNA loci are sliced by Zucchini and are loaded onto Piwi or Aubergine [1,2,29]. The primary piRNAs remain in the nuage and are able to target complementary transcripts produced by both piRNA loci and euchromatic TE copies. The step called ping-pong amplification involves the PIWI proteins Aubergine and Ago3 [1,2,29]. In this “cytoplasmic” model, maternally transmitted piRNAs loaded on Aubergine or Ago3 would reach the nuage of PGC and could activate *de novo* the ping-pong process. Therefore, we have first tested the effect of mutations of genes involved in the primary and secondary piRNA pathways on *BX2** silencing and piRNA production capacities. Second, a *BX2** locus produces not only transgenic piRNAs but also 21nt homologous ovarian small RNAs [22]. We have investigated the nature of these 21nt RNAs and their functional implications in the paramutation process using a *Dicer-2* mutation. Third, by combining various sources of maternal piRNAs, we have analyzed the piRNA production capacity spreading within a piRNA locus. Fourth, we have tested a potential impact of maternally inherited piRNAs outside the progeny gonads.

We show in this paper that *BX2** silencing and piRNA production capacities depend on both the primary and secondary piRNA pathways, does not require 21nt RNA production, these last RNAs being *Dicer-2* and surprisingly also piRNA pathway dependant. Further, a rapid spreading for production capacity can be observed within the *BX2* locus and maternally inherited transgenic homologous piRNAs can enhance repression in progeny adult eyes. These results, combined to those of Le Thomas *et al.* [26], support the existence of an effect of maternal piRNAs on the *BX2* locus itself. In addition, they show that spreading of piRNA production capacity within a locus can be rapid.

RESULTS

***BX2** ovarian small RNA dependence on piRNA ping-pong and siRNA pathways.**

*BX2** females were previously shown to produce 23-29nt and 21nt ovarian sense and antisense small RNAs homologous to the *P-lacW* transgene cluster [22]. The 23-29nt, but not 21nt, RNAs present the molecular signatures of the ping-pong amplification step of the piRNA pathway. This strongly suggests that 23-29nt RNAs are dependent on the piRNA pathway whereas 21nt RNAs are dependent on the siRNA pathway. Moreover, the silencing capacities of the *BX2** paramutated locus were shown to be completely impaired by a loss of function of *aubergine* involved in the piRNA ping-pong step, whereas they were insensitive to a loss of function of *Dicer-2* involved in the siRNA pathway [22], suggesting that the 21nt would not be necessary for *BX2** silencing capacities. We have thus performed deep sequencing of ovarian small RNAs of mutant females for *aubergine* and *Dicer-2*. As control, we have analysed small RNAs from ovaries of *BX2** females at generation 83 after the paramutation process and of the initial *BX2* naïve females. *BX2** lines, maintained at 25°C, have presently gone beyond generation 100 and females silencing capacities are still complete (TSE=100%, n= 1150, at G₁₁₅) showing a very strong stability through generations of the *BX2** paramutated state. **Figure 1** shows the analysis of ovarian small RNAs homologous to the *P{lacW}* transgene and to the dual-strand 42AB piRNA producing locus in *BX2* naïve females and in various mutant genotypes of *BX2** females. The *BX2* naïve cluster does not produce significant amounts of *P{lacW}* homologous 23-29nt or 21nt RNAs (A) whereas the *BX2** paramutated cluster still produces high levels of *P{lacW}* homologous 23-29nt and 21nt RNAs, as tested 83 generations after the paramutation process (B). *aub* loss of function results in the complete loss of 23-29nt RNAs homologous to the *P{lacW}* transgene (F) and in a moderate decrease of these RNAs produced by the 42AB locus (H), when compared to heterozygous *aub* females (E, G). Surprisingly, *aub* loss of function also affects the production of 21nt RNAs by the *P{lacW}* cluster (F). This does not result from an uncontrolled genomic background effect affecting the siRNA pathway in our *aub* mutant females since the production of endogenous siRNAs at the *esi-1* locus is not affected (**Figure S3**). *Dcr-2* loss of function results in a strong decrease of 21nt RNAs homologous to the *P{lacW}* transgene but the production of corresponding 23-29nt RNAs is not affected (J). The production of 23-29nt RNAs by the 42AB

locus is not affected by a *Dcr-2* loss of function (L). Thus, *P{lacW}* homologous 23-29nt RNAs depend on *aubergine* and does not depend on *Dicer-2*, the typical picture for piRNAs. *P{lacW}* homologous 21nt biogenesis depends on *Dicer-2*, as expected for siRNAs, but surprisingly depends also on *aubergine*, a more striking picture.

***BX2** transgenic piRNAs, but also 21nt RNAs, depend on the primary piRNA pathway.**

Two main molecular models, one nucleic, one cytoplasmic, can be proposed for the paramutation process [26,27]. In the cytoplasmic one, maternally transmitted piRNAs stimulate the ping-pong process in progeny germline. In the nucleic model, maternally transmitted piRNAs stimulate the production of piRNA precursors in progeny germline, resulting in an increase of primary piRNA production. We have therefore functionally investigated if paramutation involves the part of the piRNA pathway which is upstream the ping-pong step. We have tested the effect of genes involved in primary piRNA production on *BX2** silencing capacities and ovarian small RNAs production. Rhino (*rhi*) and Cutoff (*cuff*) are known to be present on dual strand piRNA producing loci and appear to have a causal role in primary piRNA production *via* a modification of the transcription and splicing pattern of the locus [3,4,7]. Zucchini (*zuc*) appears to be involved in the first cleavage of the precursor RNAs in the nuage [30-32], a preliminary step in primary piRNA biogenesis. Recombinant chromosomes carrying a *BX2* cluster and a mutant allele of *rhino*, *cutoff* or *zucchini* were generated. Males carrying these chromosomes were crossed with *BX2** females allowing recovery of paramutated *BX2* rhi*⁻, *BX2* cuff*⁻ and *BX2* zuc*⁻ lines. Females from these lines were crossed to males carrying a mutant allele of the gene tested and a target transgene to assay their *trans*-silencing capacities. **Figure 2** shows the comparison of females heterozygous or having a loss of function of these genes for *BX2** silencing capacities and for production of ovarian small RNAs homologous to the *P-lacW* transgene and to the 42AB dual-strand piRNA locus. TSE percentages show that *rhi*, *cuff* and *zuc* losses of function result in a complete loss of the *BX2** silencing capacities (B, F, J) whereas no effect is found at the heterozygous state (A, E, I). *rhi*, *cuff* and *zuc* losses of function also result in a complete loss of ovarian piRNAs, and almost complete loss of 21nt RNAs, homologous to the *P{lacW}* transgene (B, F, J). For the 42AB locus, comparison of loss function to heterozygous females show a dramatic effect of each mutation tested on piRNA production (D, H, L vs C, G, K). Strikingly, *zuc* loss of function results in an apparent increase of production of 21nt RNAs by this locus. Therefore, *BX2** silencing capacities and *P{lacW}* homologous piRNAs production depend on the primary piRNA pathway and , as in Figure

1, *P{lacW}* homologous 21nt production strikingly depends also on the piRNA pathway genes tested.

Production of ovarian transgenic siRNAs requires homologous piRNAs, but not siRNAs maternal inheritance.

Previous report [22], combined to Figure 1-J, shows that silencing properties and production of piRNAs by the *P{lacW}* transgenes does not require maternal inheritance of siRNAs. Figures 1 and 2 show that such ovarian siRNAs presence depends, in addition to *Dcr-2*, to all piRNAs genes tested. We have tested if production of *P{lacW}* ovarian siRNAs, like piRNAs, requires their own maternal inheritance or if alternatively they require only piRNA homologous inheritance. We have thus generated *BX2** females which did not inherit *P{lacW}* homologous siRNAs, due to homozygous *Dcr-2* mutant mothers, but having paternally inherited a *Dicer-2* WT allele (**Figure 3**). As expected [22], these *Dcr-2* heterozygous females show complete silencing capacities. Deep sequencing shows that, in addition to ovarian *P{lacW}* piRNAs, they produce *P{lacW}* siRNAs. Further, these *G*₁ females were crossed with *Dcr-2*^{+/+} males and increased ovarian *P{lacW}* siRNAs was found in *BX2** *G*₂ females having recovered a *Dcr-2*^{+/+} genotype. Therefore, production of *P{lacW}* siRNAs can be induced by inheritance of solely homologous piRNAs. This reinforces the idea that these 21nt RNAs are *Dcr-2* dependent byproducts of piRNA biogenesis.

Paramutation of a locus partially covered by maternally inherited piRNAs.

Previous studies have shown that transgenes or natural transposable elements targeted by piRNAs can result in piRNA production, not only by the targeted elements, but also by their adjacent sequences [26,33,34]. This shows the existence of a kind of spreading of piRNA production capacity beyond the sequence targeted by piRNAs. However, the repressive capacity of piRNAs resulting from spreading was not tested. We have tested if such spreading can be observed for silencer loci produced by paramutation and have tested the silencing capacity of piRNAs resulting from such spreading. Telomeric transgenes inserted in subtelomeric heterochromatin can show strong *trans*-silencing capacities [35-38] and produce abundant piRNAs in the germline [22,39]. These transgenes have a different structure than the *P{lacW}* transgene of *BX2* but share some common sequences [22]. It was therefore possible to produce females, having paternally inherited a *BX2* naive locus and maternally inherited a cytoplasm containing piRNAs homologous

to only a subset of the $P\{lacW\}$ transgenes. In order, to transmit the cytoplasm, but not the telomeric transgenes, females heterozygous for these telomeric transgenes were generated. Two telomeric transgenic loci were used (**Figure 4-A**). The $P-1152$ line carries two copies of the $P-lacZ-rosy$ transgene ($P\{larB\}$) in the Telomeric Associated Sequences (TAS) of the X -chromosome[36]. The $P-1152$ telomeric locus was the canonical locus used in previous *Trans*-Silencing Effect studies [35,36,40,41]. $RS3$ carries a $P-frt-white$ transgene in the TAS of the third chromosome ($3R$ arm). It also induces a strong TSE [42]. So the $P\{lacW\}$ shares P - element and $lacZ$ sequences with $P-1152$ and shares P - element and $white$ sequences with $RS3$. In both cases, a part of the $P\{lacW\}$ sequence is missing in the telomeric transgenic construct. We have crossed $BX2$ males with females heterozygous for the $P-1152$ telomeric locus, the $RS3$ telomeric locus and both telomeric loci. G_1 females having inherited $BX2$, but not the telomeric transgenes, were recovered and potentially paramutated $BX2$ lines were established and analyzed as previously for *trans*-silencing and ovarian piRNA production capacities over generations. **Figure 4-B** globally shows that G_1 females carrying a paternally inherited $BX2$ cluster and having recovered a cytoplasm containing piRNAs produced by partially homologous transgenes can show strong or complete silencing capacities. However, the efficiency is not as strong as that of “typical” paramutation experiments in which the $P\{lacW\}$ transgenes of the $BX2$ paramutated locus inherit piRNAs homologous to the entire transgene length. This results always in complete and stable paramutation [22]. For the $BX2^*$ lines activated by a piRNAs homologous to the $P-lacZ$ transgenes (called $P-1152 BX2^*$), two lines show very strong silencing capacities after more than 30-50 generations (one is at 100% at each generation tested), one line falls down after a period of strong silencing capacity and one line shows partial silencing capacities. However, after 50 generations, two $P-1152 BX2^*$ lines are still very strongly paramutated, a picture stronger than those reported previously [22]. For the $BX2^*$ lines activated by a piRNAs homologous to the $P-white$ transgene (called $RS3 BX2^*$), one line appears long-term established at a strong level of repression, while the other falls down. For the $BX2^*$ lines activated by a piRNAs homologous to the $P-lacZ$ plus the $P-white$ transgenes (called $P-1152 + RS3 BX2^*$), the situation resembles to that of lines activated by a $P-1152$ cytoplasm, with a line showing a complete repression capacity and lines showing intermediate repression levels. Surprisingly, these $P-1152 + RS3 BX2^*$ lines did not appear stronger than the $P-1152 BX2^*$ lines despite the fact that small RNAs covering most of the sequence have been maternally transmitted in G_0 . The conclusion of these three sets of paramutation assays is that paramutation by partially homologous transgenes is less efficient than paramutation by fully homologous transgenes but can however produce some long-term stable strong paramutated lines.

For each type of G_0 maternal inheritance, lines showing strong silencing capacities were chosen and ovarian small RNA production was analyzed in G_3 , G_5 and G_{10} by deep sequencing. **Figure 4-C** shows that, as soon as generation 3, the three types of G_0 maternal inheritance can result in strong production of small RNAs covering the entire $BX2$ $P\{lacW\}$ transgene length. Thus, establishment of the capacity to produce piRNAs by the part of the $P\{lacW\}$ which did not inherit homologous piRNAs in G_1 can be rapid. This shows that spreading of the capacity to produce piRNAs by a locus, previously described for target transgenes [34] or transposable elements [33], can be observed in the paramutation process in which, by definition, the stimulating locus is not permanently present in the line, by contrast to previous reports. In addition, it was not reported previously that piRNAs resulting from this spreading were functional, *i.e.* capable to establish silencing. We have therefore established a new *trans*-silencing assay which allows to specifically assay the silencing capacities of piRNAs resulting from spreading within a $^{RS3}BX2^*$ locus. The *faf-lacZ* transgene carries a *faf-lacZ* translational fusion under the control of the *faf* promoter [43] (**Figure 5**). It shows a strong expression in the female germline (**Figure S5**). As expected, the *faf-lacZ* transgene is repressed by *P-1152* (data not shown) but not by *RS3*. Indeed, the *faf-lacZ* transcript shares the *lacZ* sequence with *P-1152* transgenes but has no homology with the *RS3* transgene. *faf-lacZ* is also repressed by a $T^{-1}BX2^*$ cluster because of the *lacZ* homology. Therefore, *faf-lacZ* repression by a $^{RS3}BX2^*$ locus can demonstrate that *lacZ* homologous piRNAs, produced by spreading within this $BX2^*$ locus, are functional for inducing repression. Figures 5 and S5 show that lines paramutated by a *RS3* cytoplasm can elicit a strong repression of the *faf-lacZ* transgene. As control, these lines show also repression of the classical *BQ16* target. Therefore, *lacZ* homologous piRNAs resulting from spreading within a piRNA producing locus can be functional for homologous *trans*-silencing. Finally, the paramutagenicity of $P^{-1152}BX2^*$, $^{RS3}BX2^*$ and $P^{-1152} + ^{RS3}BX2^*$ females has been tested. **Figure S6** shows that maternal transmission of a cytoplasm produced by these females, combined to paternal transmission of a *BX2* naïve locus, result in female progeny showing strong *trans*-silencing properties. These capacities are transmitted over generations. Thus, paramutation mediated *via* inheritance of piRNAs homologous to a subset of the targeted locus can be strong and stable.

Somatic impact of maternally transmitted piRNAs

piRNAs are maternally transmitted not only at the posterior pole of *Drosophila* embryos but also at more anterior regions where they can be implicated in gene regulation. Indeed, piRNAs homologous to transposable elements (*roo*, *412*) have been shown to be involved in *nanos* mRNA

decay at the anterior part of the embryo by a mechanism which involves mRNA deadenylation[44]. In addition, piRNAs can have an effect in gonad somatic cells. Indeed, in *Drosophila simulans*, expression of the *tirant* transposable element in ovarian follicle cells appears to be affected by *tirant* homologous piRNAs maternal transmission[45]. Finally, maternal depletion of Piwi was shown in *Drosophila melanogaster* to impact *white* variegation in adult eyes[46]. Therefore, piRNAs maternal deposition can impact not only progeny germline but also progeny somatic tissues. We have tested if piRNAs deposited by the paramutagenic $P\{lacW\}$ clusters can affect $P\{lacW\}$ expression in progeny adult eyes. The *T-1* initial paramutagenic line carry seven $P\{lacW\}$ copies inserted in tandem orientation [47,48] and produces abundant ovarian piRNAs including piRNAs homologous to the *white* gene [22], used as transgene transformation marker. *white* is involved in eye pigment transport and loss of function results in white eyes instead of red ones. The *T-1* cluster shows strong variegation for the *white* marker, *i.e.* stochastic ON-OFF *white* repression from cell to cell. This results from a phenomenon termed *Repeat Induced Gene Silencing* (RIGS [47,48]) which was further enhanced by an X-ray treatment. RIGS is linked to heterochromatin formation on tandem repeats, which includes *de novo* binding of HP1 of the locus, as tested in salivary glands [49]. In addition in the *T-1* line, variegation was further enhanced by an X-ray treatment which induced chromosomal rearrangements. In order to investigate a possible somatic impact of $P\{lacW\}$ homologous piRNAs maternal inheritance, we have performed the two reciprocal crosses between *T-1/Cy* individuals and *white* mutant individuals from three different genetic backgrounds. Female progeny eyes were compared for eye phenotype (picture and pigment dosages).

Figure 6-A shows that for the three backgrounds tested (w^{1118} , $w^{1118} Cy/Xa$, yw^c), a very clear difference is observed between the two reciprocal cross progenies for the *white* phenotype. Females having maternally inherited the *T-1* cluster present stronger *white* repression than females having paternally inherited the cluster. Note that the two types of females have, in each case, strictly the same genotype. These differences are confirmed by dosage pigments.

We have further tested the effect of $P\{lacW\}$ homologous piRNAs inheritance on eye phenotype of another $P\{lacW\}$ cluster called *DXI*. This cluster has six $P\{lacW\}$ copies, one of which is in opposite orientation[47,48]. This cluster shows a low level of variegation at 25°C but variegation increases at higher temperatures to reach almost the white phenotype at 29°C (data not shown). *DXI* has no *trans*-silencing capacities [50]. The two reciprocal crosses were first performed

between *DX1* and *w*¹¹¹⁸ individuals to test for the existence of maternal effect of *DX1* inheritance on its eye phenotype and no significant difference was observed between the two types of progeny for *white* expression (**Figure 6-B**). Further *DX1* individuals were crossed with *T-1* individuals and female progenies having inherited *DX1*, but not *T-1*, were analyzed. Figure 6-B show that *DX1* females having maternally inherited a *T-1* cytoplasm shows more repressed eyes than *DX1* females having inherited a *DX1* cytoplasm. Thus, *T-1* maternal cytoplasm induces *DX1* repression in adult heads. Moreover, females produced by the *DX1* females x *T-1* males cross show significantly more repressed eyes than those produced by the *DX1* x *w*¹¹¹⁸ cross. This likely results from background effects since no paternal inheritance from *T-1 P{lacW}* piRNAs was previously detected [22]. By contrast, the difference observed between crosses between *DX1* and *T-1* individuals can be attributed to piRNAs maternal transmission since these two types of female progeny have the same genotype. Consequently, a repressive effect of a *T-1* cytoplasm which carries a high level of *P{lacW}* piRNAs can be observed in progeny adult eyes with both the *T-1* and *DX1 P{lacW}* clusters as read out.

DISCUSSION

Implication of the primary piRNA pathway in *BX2** capacities.

The present analysis shows that all the piRNA genes tested are required for a paramutated locus to exhibit *trans*-silencing and piRNA production capacities. In a paramutation cytoplasmic phenomenon model, it can be proposed that a *trans*-generational nuage to nuage transfer of informations can be the only factor which explains the acquisition of silencing capacities by the *BX2** paramutated locus. It was previously shown by Q-RT-PCR that the *BX2* naïve locus is transcribed in both orientations [22]. Thus, during paramutation, piRNAs loaded on Aub or Ago3, produced in the nuage of *T-1* or *BX2** female nurse cells, could be transferred to the nuage of zygotic PGC having paternally inherited a naïve *BX2* locus, resulting in a *de novo* targeting of *BX2 P{lacW}* transcripts by the ping-pong machinery. In a paramutation nucleic phenomenon model, maternally transmitted piRNAs, loaded on Piwi, could enter into PGC nuclei and activate a *de novo* targeting of *BX2* naïve transcript by the primary piRNA pathway *via* promoting the binding of the Rhino-Deadlock-Cutoff complex on the *P{lacW}* cluster. This last model is strongly supported by ChIP analysis which showed that a *BX2** cluster, as tested in G₁ ovaries of the paramutation

experiment, shows an increase of the H3K9 methylation level when compared to a naïve *BX2* locus [26]. H3K9me3 residues can indeed be targeted by the chromodomain of Rhino, resulting in the RDC stabilization on the cluster. The mutant analysis reported here provides a functional assay which shows that *rhino*, *cutoff* and *zucchini* are necessary for *BX2** *trans*-silencing and piRNAs production capacities (Figure 2). Therefore, paramutation-linked repressive properties do not solely depend on a cytoplasmic ping-pong activation resulting from nuage to nuage *trans*-generational transfer of informations. These properties depend also on the primary piRNA pathway which is linked to the binding of the RDC complex on the locus [3,4]. In addition, analysis performed with the initial paramutagenic locus called *T-1*, which produces a high level of ovarian *P{lacW}* piRNAs, strongly reinforces this model. Indeed, ChIP analysis performed on ovaries of G₁ females produced by the two crosses between *T-1* and WT (devoid of transgene) individuals showed that maternal inheritance of the *T-1* cytoplasm is associated to an increase of Rhino and Cutoff binding on the *T-1 P{lacW}* cluster in progeny ovaries [26]. Paramutation associated repressive properties depend also on the Zucchini and on *aub* activities which take place in the nuage (Figure 1 and 2). In conclusion, combination of chromatin analysis together with mutant analysis converge to show that *trans*-generational induction of piRNA-mediated silencing by maternally inherited piRNAs involves all the piRNA machinery, including the germline primary piRNA pathway. Thus, during paramutation, the paramutated locus is, and remains, modified over generations.

***BX2** transgenic 21nt RNAs are actual siRNAs which depend also on the piRNA machinery.**

Deep sequencing of ovarian small RNAs associated to both telomeric *P*-transgene insertions and to the *T-1* paramutagenic or *BX2** paramutated clusters showed, in addition to 23-29 piRNAs, the presence of 21nt abundant transgene homologous RNAs [22]. Sense and antisense 21nt RNAs were found but no ping-pong signatures was detected between them. These 21nt RNAs were proposed to be *Dicer-2* dependent. Here, ovarian small RNAs of *BX2**, *Dcr-2*⁻ / *Dcr-2*⁻ female have been analyzed by deep sequencing. As expected, these *Dcr-2* mutant females lacks transgenic 21nt RNAs showing that they are actually siRNAs (Figure 1). Strikingly, all the piRNA mutations tested result also in the loss of these transgenic 21nt RNAs (Figure 2). Thus both primary and secondary piRNA pathways are required for 21nt RNAs biogenesis suggesting that they could appear downstream the piRNA pathway. It was shown further that these transgene homologous 21nt require piRNA, but not 21nt, maternal inheritance to be produced, confirming that their appearance is downstream of piRNA. 21nt production appears to be a common feature for a number of piRNA

producing loci [51,52]. *cuff* mutant analysis was shown to strongly affect 21nt RNA production by four major double-strand piRNA producing loci including the 42AB locus [26]. What can be the mechanism of the 21nt dependence from piRNA pathway? 21nt biogenesis depends on all the piRNA actors tested including for example Rhino which is necessary to produced piRNA precursors, Zucchini which makes the piRNA precursor first slicing in the nuage and Aubergine which is an actor of the ping-pong amplification step. The fact that 21nt RNAs depend from Aubergine likely results from the fact that *aubergine* mutants induce delocalize of Piwi within nurse cell nuclei [53], possibly reducing a positive feedback in the nucleus of piRNAs associated to Piwi on primary piRNA production. One possibility is that 21nt *P{lacW}* RNA production results from the fact that both sense and anti-sense RNAs, produced by this duals-strand cluster, undergo a local concentration due to the piRNA machinery, for example in the nuage, leading to occurrence of double-strand RNAs which can further be processed by *Dcr-2*. piRNA mutations would abolish this concentration and thus 21nt RNA production.

Spreading of the piRNA production capacity within a piRNA producing cluster can be completed within 3 generations.

piRNA producing loci constitute genomic traps in which insertions of new TEs are followed by *de novo* production of piRNA homologous to this new insertion. The kinetics of appearance of this new piRNA production is not clear but several evidences come from previous studies concerning the repression capacity inheritance of the *I* factor (called reactivity) and of the *P* element (called cytotype). In both systems, it appeared that establishment of repression capacities, following paternal inheritance of *I* or *P* elements carrying genomes in a cytoplasm devoid of *I* or *P* homologous piRNAs, is progressive and require several generations to be completed [11,12]. For example, experiments were performed with a line (*Lk-P(1A)*) carrying only two *P* elements inserted in a telomeric piRNA producing locus (Telomeric Associated Sequences, TAS). It was shown that paternal transmission of the telomeric *P*-element locus in a cytoplasm devoid of *P*-homologous piRNAs, followed by maternal inheritance of the telomeric *P* elements in successive generations, requires four generations to progressively establish a complete *P*-repression[19]. So, a spreading must occur within the TAS locus for piRNA production capacity which finally results in strong production of piRNAs homologous, not only to the TAS, by also to the *P*-element. Similar spreading process for piRNA production was observed for the *flamenco* piRNA producing locus which works in the somatic follicle cells which surround egg chambers [39]. Further, analysis of old transgenic lines which carry a *I* fragment within a *P* transgene have shown that piRNA production

capacity can spread within transgenes [34]. These transgenes behaves as a relay, for *I*-element piRNA production, of a stronger master locus. The capacity to produce piRNAs invaded the whole transgene and reached in some cases flanking sequences. Similar spreading process until flanking sequences was also found for natural transposons insertions [33]. In these last experiments, the functionality of the piRNAs resulting from the spreading process, *i.e.* the capacity to induce *trans*-silencing, was not reported.

We have used the paramutation experiment to observe such a spreading process within a paramutated cluster, to test if it can be rapid and finally to test the *trans*-silencing capacities of piRNAs resulting from the spreading. The *BX2* cluster which carries *P{lacW}* transgenes were introduced by crosses in a cytoplasm containing *P* and *lacZ* homologous piRNAs or containing *P* and *white* piRNAs. So, a part of the paternally inherited *P{lacW}* transgenes was not in contact to homologous maternally inherited piRNAs during the paramutation process. Despite these partial “covering” of the *P{lacW}* length in *G*₁ females, a strong repression was found in the *BX2** lines as soon as the first generation. This repression capacity was correlated to the production of abundant piRNAs corresponding to the complete *P{lacW}* length, as tested at generation 3 (Figure 4). These *BX2** lines were also shown to induce be fully paramutagenic (Figure S6). Finally, *lacZ* homologous piRNAs produced by *BX2** females, paramutated by a cytoplasm missing *lacZ* piRNAs, were shown to be functional for *trans*-silencing (Figure 5). Therefore, they are likely canonical piRNAs. This shows that such spreading process can result in rapid production of functional piRNAs.

What can be the mechanism of this process? It can be proposed that the RDC complex binding results in a tagging of a complete transcript covering the whole *P{lacW}* transgene. After, transfer to the nuage, this transcript is driven toward destruction by the piRNA pathway and no memory persists in the nuage concerning the part of the sequence which was targeted by maternal piRNAs in the nucleus. Some observations however do not perfectly fit with this model. First, in that case, establishment of repression should require only one generation to be completed and in fact it was shown to be progressive over generations [19]. Second, we have made an additional control experiment in which a paternally inherited *BX2* locus was introduced in a maternal cytoplasm produced by *Lk-P(1A)* females. In that experiment, absolutely no paramutation of the paternally transmitted *BX2* cluster was observed (**Figure S7**) despite maternal inheritance of *P*-element homologous piRNAs corresponding to 820bp out of the 10691bp of the *P{lacW}* transgene. It remains however possible that such tagging process of transcripts is quantitative and could be linked to the fact that the probability to fix the RDC complex on the locus is dependent on the

length of the locus in contact with maternal homologous piRNAs in embryos. Consequently this spreading could be able to fill a gap of 4 kb but not a gap of almost 10 kb.

Maternally inherited piRNAs have a phenotypic effect in somatic tissues.

Maternal deposition of piRNAs in oocytes is not restricted to the pole plasm and maternally transmitted piRNAs were shown to play a key role for *nanos* expression patterning in embryos. Indeed, it was shown that *nanos* maternally deposited mRNA are targeted by piRNAs homologous to two TEs (*roo* and *412*) in the anterior part of the embryo [44]. These piRNA target a short homologous motif in the 3'UTR of *nanos* and promotes RNA instability *via* deadenylation of the *nanos* transcript. Aubergine and Ago3 are involved in this process which is mechanistically completely distinct from the piRNA mediated target degradation occurring in the gonads. Moreover, in *D.simulans*, maternal deposition of piRNAs homologous to the *tirant* TE affects *tirant* expression at the adult stage in the somatic follicle cells of the ovaries [45]. Therefore, maternally transmitted piRNAs can have an effect in progeny somatic tissues including at the adult stage. We have investigated the effect of maternally transmitted transgenic piRNAs on transgene expression in adult eyes and have found a clear effect with the use of two *P{lacW}* clusters as phenotypic read-out. Females that maternally inherited *P{lacW}* piRNAs show more repressed *white* expression in eyes (Figure 6). We have investigated the presence of *P{lacW}* piRNAs and 21nt RNAs in heads and found a small but significant amount of both RNAs in female progeny from *T-1* mothers. These small RNAs may be responsible for the increased repression of the cluster expression in the eyes. Alternatively, these head small RNAs can have no functional role and a chromatin based memory of the early presence of small RNAs can be responsible for a developmental epigenetic transmission of *P{lacW}* repressed state during development. To understand this piRNA somatic impact in progeny, it will be necessary to test it in zygotes which are mutants for small RNA silencing pathways or chromatin based- repression pathways.

MATERIAL AND METHODS

Experimental conditions. All crosses were performed at 25°C and involved 3-5 couples in most cases. *lacZ* expression assays were carried out using X-gal overnight staining as described in Lemaitre *et al.* 1993 [54], except that ovaries were fixed for 6 min. *Trans*-silencing was quantified as previously described by determining the percentage of egg chambers with no *lacZ* expression in the germline (TSE%) [36].

Dissection and RNA extraction.

Total RNAs were extracted using TRIzol® (Life Technologies™) as described in the Reagent manual (http://tools.lifetechnologies.com/content/sfs/manuals/trizol_reagent.pdf). For the RNA precipitation step, 100% ethanol was used instead of isopropanol. For ovarian small RNAs, 25 pairs of ovaries were dissected for each sample in PBS 1X.

Transgenes and strains. *P-lacZ* fusion enhancer-trap transgenes: *P-1152*, and *BQ16*, contain an in-frame translational fusion of the *E. coli lacZ* gene to the second exon of the *P* transposase gene and a *rosy* transformation marker [55]. The *P-1152* insertion was mapped to the telomere of the *X* chromosome (cytological site 1A) and consists of two *P-lacZ* insertions in the same TAS unit and in the same orientation [36]. *P-1152* is homozygous viable and fertile. *BQ16* is located at 64C in euchromatin of the third chromosome and is homozygous viable and fertile. *P-1152* shows no *lacZ* expression in the ovary, *BQ16* is strongly expressed in the nurse cells and in the oocyte [36]. *RS3* is a *P-FRT-white* transgene (FBms0003945). It is inserted in the Telomeric Associated Sequences (TAS) of the 3R chromosomal arm (site 100E3). It is homozygous viable and fertile (Bloomington #123282). *P{faf-lacZ.F}JF1* is a *P*-transgene which carries a translational *faf-lacZ* fusion containing 4.2Kb of the *faf* gene upstream the *lacZ* sequence [43]. It is driven by the *faf* promoter and the transcript does not include *P* element sequence. It is strongly expressed in the female germline (Bloomington #4921). *P-1152* and *RS3* induce a strong TSE and produced abundant ovarian transgenic piRNAs ([22,56] and data not shown).

***P{lacW}* clusters.** The *BX2* line carries seven *P-lacZ-white* transgenes, including at least one

defective copy, inserted in tandem orientation at cytological site 50C on the second chromosome. The transgene insertion site is located near the *mRpL53* gene, in an *Ago1* intron [22]. This site is not a piRNA-producing locus, as observed for instance in the deep-sequencing dataset from *P-1152* ovaries (data not shown). The *P-lacZ-white* construct contains the *P-lacZ* translational fusion and is marked by the mini-*white* gene (*P{lacW}*, FBtp0000204, 10691bp). *T-1* derives from *BX2* following X-ray treatments [47,48]. *T-1* has chromosomal rearrangements including translocations between the second and the third chromosomes. *DX1* carries six *P-lacZ-white* transgenes, one of which is in opposite orientation, at the same genomic site than *BX2*. After overnight staining, weak *lacZ* expression is detected in the follicle cells of *BX2*, *T-1* and *DX1* female ovaries, presumably because of a position effect at 50C, but no staining is observed in the germline (data not shown). The three *P{lacW}* clusters show *white* variegated repression in the eyes due to *Repat Induced Gene Silencing* (RIGS). Variegating repression levels are different between the clusters (*T-1* >> *DX1* > *BX2*).

Lines carrying transgenes have M genetic backgrounds (devoid of *P* transposable elements), as do the multi-marked balancer stocks used in genetic experiments. The *w*¹¹¹⁸ and Canton^y line were used as control completely devoid of any *P* element or transgene. Crosses performed with *BX2*, *T-1* or *DX1* were performed with females carrying the cluster at the heterozygous state (referred to as hemizygous in case of insertions) because of the sterility (*BX2*, *DX1*) and lethality (*T-1*) induced by transgene clusters.

Two strong hypomorphic mutant alleles of *aubergine* induced by EMS were used. Both of them are homozygous female sterile and TSE was previously shown to be abolished by a heteroallelic combination of these alleles [36]. *aub*^{QC42} comes from the *Bloomington Stock Center* (stock #4968) and has not been characterized at the molecular level [57]. *aub*^{N11} has a 154bp deletion, resulting in a frameshift which is predicted to add 16 novel amino acids after residue 740. *Dicer2*^{L811fsX} is a loss of function allele induced by EMS which has a sequence variant at residue 811 resulting in a stop codon [58]. It is homozygous viable and fertile. *rhino*⁰²⁰⁸⁶ (*rhi*²) results from a *P{PZ}* insertion (*P-lacZ-rosy*) at nucleotide 267. The *P{PZ}* transgene is oriented in the opposite direction to that of *rhi* transcription [59] ((Bloomington #1226) . *rhi*^{KG00910} (*rhi*^{KG}) is due to the insertion of a *P{SUPor-P}* transgene which carries the *yellow* and *white* transformation markers (Gene disruption project 2001). (Bloomington #13161). *cutoff*^{N1WM25} and *cutoff*^{QQ37} alleles were isolated from an EMS screen [57]. *zucchini* alleles were isolated from an EMS screen [57]. *zuc*^{HM27} contains a stop codon at residue 5, *zuc*^{SG63} a substitution of histidine 169 with a Tyrosine in the

conserved HKD domain presumably involved in nuclease activity. TSE was previously shown to be abolished by a heteroallelic combination of these alleles³⁹. The crosses performed to generate hetero-allelic mutant *BX2** females are indicated in **Table S1**. When informations were available, the most severe allele was maternally introduced. Lines carrying mutations of *zuc* and *cuff* were kindly provided by Attilio Pane and Trudi Schüpbach and *aub*^{N11} was kindly provided by Paul Macdonald.

All the alleles described above being located on the second chromosome are maintained over a *Cy* balancer chromosome. Additional information about mutants and stocks are available at flybase: <http://flybase.bio.indiana.edu/>

Deep sequencing analyses

Libraries were sequenced using an Illumina Hi-Seq 2000. Sequence reads in fastq format were trimmed from the adapter sequence 5'-TGGAATTCTCGGGTGCCCCAAG-3' and matched to the *D. melanogaster* genome release 5.49 using Bowtie [60] as well as to the sequences of the *P*-element constructs *P{lacW}* (Flybase ID FBtp0000204). Only 19-29nt reads matching the reference sequences with 0 or 1 mismatch were retained for subsequent analysis. For global annotation of the libraries (**Table S1**), we used release 5.49 of fasta reference files available in Flybase, including transposon sequences (dmel-all-transposon_r5.49.fasta) and release 18 of miRNA sequences from miRBase (www.mirbase.org).

Sequence length distributions, small RNA mapping and frequency maps were generated using in-house Pythonscripts and R (www.r-project.org/) to analyze Bowtieoutputs. Scripts were integrated and run in a Galaxy instance hosted by the laboratory. Analysis pipelines and script codes are accessible from <http://www.drosophile.org/galaxy> upon request. For library comparisons, read counts were normalized (Table S1) to the total number of small RNAs that matched the *D. melanogaster* genome and did not correspond to abundant cellular RNAs (rRNAs, tRNAs and snoRNAs). For small RNA mapping, we matched each individual RNA sequence to *P{lacW}* or the 42AB locus and gave to each matched position a weight corresponding to the normalized occurrence of the sequence in the small RNA library. When RNA sequences matched *P{lacW}* or the 42AB locus repeatedly, the weight was divided by the number of hits to these *P*-element constructs.

Distributions of piRNA overlaps (ping-pong signatures, **Figure S1**) were computed by collecting, for each sequencing dataset, all the 23-28nt RNA reads matching *P{lacW}* or the 42AB

locus whose 5' ends overlapped with another 23-28nt RNA read on the opposite strand (**Table S2**). Then, for each possible overlap of 1 to 28nt, the number of read pairs was counted. Distributions of siRNA overlaps (**Figure S2**) were computed using a similar procedure, except that 20-21nt RNA reads were collected instead of the 23-28nt RNA reads. The distributions of piRNA/siRNAs overlaps (Figure S2) were computed by collecting separately the 20-21nt and 23-28nt RNA reads matching *P{lacW}* or the 42AB locus and counting for each possible overlap of 1 to 21nt the number of read pairs across these two distinct read datasets. 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 it using the formula $z(i) = (O(i) - \text{mean}(O)) / \text{standard deviation}(O)$.

Supplementary Information is linked to the online version of the paper at www.

Acknowledgements: We thank Doug DORER, Steve HENIKOFF, Attilio PANE, Trudi SCHÜPBACH, Paul MACDONALD and the *Bloomington Stock Center* for providing stocks, and flybase.org for providing databases. We thank Ritha ZAMY for technical assistance. We thank Neel RANDSHOLT, Anne-Marie PRET, for critical reading of the manuscript. This work was supported by fellowships from the *Ministère de l'Enseignement Supérieur et de la Recherche* to CH and by grants from the *Association de la Recherche contre le Cancer* (Fondation ARC), from the *Fondation pour la Recherche Médicale* (FRM) and from the *Association Nationale de la Recherche* (ANR, project "platisipi") to SR.

Author Information Small RNA sequences have been deposited at the NCBI under accession numbers Reprints and permissions information is available atThe authors declare no competing financial interests. Correspondence and requests for materials should be addressed to SR (stephane.ronssey@upmc.fr).

LEGEND TO FIGURES

Figure 1. Effect of mutations affecting *aubergine* and *Dicer-2* on small RNA production by a *BX2 cluster**

(A-L) Deep sequencing of small RNAs from ovaries of the indicated genotypes. Histograms show the length distributions of small RNAs matching *P{lacW}* or the 42AB locus. Positive and negative values correspond to sense and antisense reads, respectively. Plots show the abundance of 19-29nt small RNAs matching *P{lacW}* or the 42AB piRNA producing locus (unique mappers). *aub* loss of function has a dramatic effect on 23-29nt RNA production by the *P{lacW}* cluster and a moderate effect on such production by the 42AB locus. Surprisingly *aub* loss of function has also a strong effect on 21nt RNA production by the *P{lacW}* cluster. *Dcr-2* loss of function strongly affects production of *P{lacW}* homologous 21nt RNAs without affecting 23-29nt RNA production. *P{lacW}* 21nt biogenesis therefore depends on both *Dicer-2* and *aubergine*.

Figure 2. Effect of mutations affecting *rhino*, *cutoff* and *zucchini* on *trans*-silencing and small RNA production capacities of a *BX2 cluster**

BX2 rhi⁻*, *BX2* cuff⁻* or *BX2* zuc⁻* females were crossed with males carrying a mutant allele of *rhi*, *cuff* or *zuc* respectively to generate heterozygote or loss of function females for deep sequencing of ovarian small RNAs. *BX2* rhi⁻*, *BX2* cuff⁻* or *BX2* zuc⁻* females were also crossed with males carrying a mutant allele of *rhi*, *cuff* or *zuc* respectively and a *P-lacZ* target transgene (*BQ16*) to measure the effect of the tested mutant on *BX2** silencing capacities. This last cross produced females carrying a *BX2** cluster and a target transgene in a heteroallelic or homozygous mutant context for the gene tested, together with heterozygous sisters. *Trans*-silencing (TSE) capacities were measured in these females by counting the percentage of egg chambers showing target transgene repression. (A-L) In each case, the genotype tested is indicated, ovarian small RNAs are shown and the percentage of TSE is given below the genotype with the total number of egg chambers counted in parenthesis. Histograms show the length distributions of small RNAs matching *P{lacW}* or the 42AB locus. Positive and negative values correspond to sense and antisense reads, respectively. Plots show the abundance of 19-29nt small RNAs matching *P{lacW}* or the 42AB piRNA producing locus (unique mappers). *rhi*, *cuff* and *zuc* mutations show that *BX2** silencing capacities, ovarian *P{lacW}* piRNA, but also 21nt RNA, production depend on the primary piRNA pathway.

Figure 3. Paramutation by a cytoplasm devoid of siRNAs

BX2 Dcr-2⁻/Dcr-2⁻* females were crossed to *BX2 Dcr-2⁺/Dcr-2⁺* naïve males to generate G₁ females carrying a paternally inherited *BX2* locus and having maternally inherited piRNAs but no 21nt RNAs homologous to the *P{lacW}* cluster. In addition these females have paternally inherited a *Dcr-2⁺* allele. These G₁ females were crossed with *Dcr-2⁺* males to establish a *BX2* Dcr-2⁺/Dcr-2⁺* line in which no maternal *P{lacW}* 21nt RNAs have been initially introduced. At generations 1 and 2, deep sequencing of ovarian small RNAs was performed and the silencing capacity was controlled by crossing females with males carrying the *BQ16* target transgene and scoring TSE in female progeny. The genotype tested is indicated with maternally inherited chromosome written above the bar and ovarian 19-29nt RNAs and TSE percentages are presented as in Figure 2. As expected, G₁ and G₂ females show complete silencing capacities (as well as G₆ and G₁₀ females tested similarly, data not shown). Despite the lack of maternal *P{lacW}* 21nt transmission in G₀, such ovarian small RNAs appear as soon as G₁ females and increase in G₂ in *BX2* Dcr-2⁺/Dcr-2⁺* females. Therefore, maternal inheritance of solely *P{lacW}* piRNAs results in immediate production of both ovarian piRNA and 21nt RNAs in a female carrying a *Dcr-2⁺* allele and in a strong and stable paramutated *BX2** locus.

Figure 4. Paramutation of a *BX2* locus by piRNAs produced by partially homologous transgenes

BX2 males carrying a *P{lacW}* naive cluster were crossed to females carrying partially homologous *P*-transgenes inserted in telomeric piRNA producing loci at the heterozygous state. Female progeny were recovered which carry the *BX2* locus but did not carry the telomeric transgenes. These G_1 females inherited a cytoplasm carrying piRNA homologous to the telomeric transgenes, which therefore cover only partially the *BX2* *P{lacW}* transgene length. Lines were established and the *trans*-silencing capacities of these putative *BX2** lines were tested in subsequent generations, as in Figures 2 and 3. In addition, deep sequencing of ovarian small RNAs was performed in G_3 , G_5 and G_{10} . **(A-left-middle)** Structure of the telomeric *P*-transgenes used and of the *BX2* *P{lacW}* transgenes. Colored arrows below the transgenes allow visualization of the common sequences between *P{lacW}* and telomeric transgenes. The *P-1152* line carries two *P-lacZ-rosy P{larB}* transgenes inserted in the Telomeric Associated Sequences (TAS) of the *X*-chromosome. *RS3* carries a *P-frt-white* transgene in the TAS of the *3R* chromosomal arm. These two lines induce a strong TSE. **(A-right)** Ovarian piRNA production in females that carry a maternally inherited *P-1152* or *RS3* locus at the heterozygous state (G_1 females from crosses between *P-1152* females and Canton males and between *RS3* females and Canton males, respectively) is shown, mapped on the *P{lacW}* transgene. **(B)** Three types of replicate *BX2** sublines were generated. Sublines which inherited in G_1 : 1- a cytoplasm from females heterozygous for the *P-1152* locus ($P^{1152}BX2$ sublines, in orange); 2- a cytoplasm from females heterozygous for the *RS3* locus ($RS3BX2$ sublines, in green); 3- a cytoplasm from females heterozygous for both the *P-1152* and *RS3* loci ($P^{1152+RS3}BX2$ sublines, in purple). Repression capacities of various *BX2** lines are shown with regard to generations. Names of replicate sublines are given below the graphs. **(C)** Deep sequencing of ovarian small RNAs in various *BX2** sublines at generations 3, 5 and 10. Plots show the abundance of 19-29nt small RNAs matching *P{lacW}*. TSE assays show that silencing capacities vary among replicate *BX2** lines but strong and stable paramutated lines can be recovered with the three types of G_0 cytoplasmic inheritance (B). Ovarian small RNAs were analyzed for a subset of *BX2** sublines showing strong silencing capacities ($P^{1152}BX2$: BB6 for G_3 and G_5 and QA5 for G_{10} ; $RS3BX2$: BB7-4 for G_3 and G_5 and QA8 for G_{10} ; $P^{1152+RS3}BX2$: JD1-2 for G_3 and G_5 and QA1 for G_{10}). Deep sequencing analysis shows that piRNAs corresponding to the whole length of *P{lacW}* can be produced as soon as generation 3 (C), showing that piRNA production by a paramutated locus can rapidly go beyond the part of the locus which received maternally inherited piRNAs.

Figure 5. Functional properties of piRNAs resulting from spreading of piRNA production capacity within a piRNA producing cluster

A *BX2** locus paramutated by a *RS3* cytoplasm produces piRNAs homologous to the *lacZ* sequence despite such piRNAs were not inherited in G_0 from *RS3* females (Figure 4). This formally shows the existence of a spreading of the piRNA production capacity within the $P\{lacW\}$ transgenes. The silencing capacities of these *lacZ* homologous piRNAs was specifically tested by crossing ^{RS3}*BX2** females with *faf-lacZ* males and by staining progeny ovaries. The *faf-lacZ* is strongly expressed in the female germline (**Figure S5**) due to the *faf*-promoter. The *faf-lacZ* transcript is expected to be targeted by *lacZ* homologous piRNAs but not by *P*-element homologous piRNAs. As controls, *faf-lacZ* males were also crossed with *RS3*, *BX2* naïve and ^{T-1}*BX2** females. In addition, all females were also crossed with *BQ16* males whose *lacZ* transcript can be targeted by piRNAs homologous to the *P* element. The parental crosses are indicated in the Table and TSE percentages are given with the number of egg chambers counted in parenthesis. For the ^{RS3}*BX2** lines tested, the name of the line is given with the generation at which the test was performed. As expected *RS3* females, producing *P*-element, but no *lacZ* homologous piRNAs, can repress *BQ16* but not *faf-lacZ*. *BX2* naïve females show no repression capacity and ^{T-1}*BX2** repress both targets. ^{RS3}*BX2** lines show *BQ16* repression capacities at each generation tested except one case. The null value obtained for BB7-4 at G_7 probably results from variations which occurred within the line since sisters of the females tested repressed *faf-lacZ* and BB7-4 showed further repression in G_{11} . Finally all ^{RS3}*BX2** lines show in each case repression of the *faf-lacZ* target transgene, demonstrating that *lacZ* homologous piRNAs of these paramutated lines can induce *trans*-silencing. Thus piRNAs resulting from spreading of piRNA production capacity within the $P\{lacW\}$ cluster are functional.

Figure 6. Somatic impact of maternally transmitted piRNAs

The effect of *P{lacW}* small RNA maternal inheritance on *P{lacW}* cluster expression in adult eyes was tested using two different transgene clusters, *T-1* (**A**) and *DX1* (**B**), as read-out. *T-1*, but not *DX1*, has *trans*-silencing capacities linked to the presence of abundant ovarian piRNAs. The G₀ cross is indicated above the picture and was performed at 25°C. Green female symbol indicates cytoplasmic inheritance containing *P{lacW}* piRNAs. The genotype of the G₁ females analyzed is given below the picture. The maternally inherited chromosome is written above the bar. (**A** and **B**) Pigments dosages are shown below the pictures. Numbers in abscissa indicate the genotype analyzed (reported directly on the head picture). Canton^y and *w*¹¹¹⁸ female heads were used for WT and *w*⁻ controls respectively. Bars indicate the standard error. Maternal inheritance of *P{lacW}* small RNA produced by a *T-1* female results in an increased of silencing of both *T-1* and *DX1* expression in the eye.

Table S1. Annotation of small RNA libraries

Small RNAs were prepared from ovaries of females of the indicated genotype (*Female genotype* row). The parental origin of these females is indicated in the *Parental origin* row. The Canton^y and the *w*¹¹¹⁸ strains are devoid of any transgene. Values for the different categories of sequences are expressed as their percentage of the total number of sequence reads that matched the *D. melanogaster* genome 5.49 in each library (Dm Matched row). Normalization factors used for library comparisons are indicated in the Table. Libraries were prepared using Illumina sequencing technologies (see experimental procedures). Library #5 was then taken as reference to normalize other libraries.

Table S2. Small RNA z-scores and pi/si ratios

Relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subsets of 24-29nt small and 20-21nt RNAs matching *P{lacW}* or the 42AB locus. Note that the cross between *BX2*/CyRoi* females and *w*¹¹¹⁸ males can produce G₁ females carrying the *CyRoi* chromosome (and thus devoid of *BX2** chromosomal) which were analyzed as controls for ovarian small RNAs (Table S1, Library ID3, Table S2, library ID3). No significant *P{lacW}* small RNAs were found in their ovaries showing that *P{lacW}* maternally deposited piRNAs have been degraded-diluted before the adult stage.

Figure S1. Effect of mutants on 23-28nt ping-pong signatures

A-AB. Panels show the relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subsets of 23-28nt small RNAs matching *P{lacW}* or the 42AB locus. The genotype is indicated above the graph. Y-AB correspond to G₁ and G₂ females shown on Figure 3.

Figure S2. Effect of mutants on 20-21nt RNAs

A-AB. Bar plots show the abundance of 20-21nt small RNAs matching *P{lacW}* or the 42AB locus in ovaries from female of the indicated genotype. Positive and negative values correspond to sense and antisense reads, respectively. Panels show the relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subsets of 20-21nt small RNAs (left) and of 20-21nt RNAs that pair with the indicated overlap with 23-28 (right), matching *P{lacW}* or the 42AB locus. Y-AB correspond to G₁ and G₂ females shown on Figure 3.

Figure S3. Effect of mutants on ovarian endo-siRNAs

A-N. Length distribution of 19-29nt small RNAs matching the matching *esi-1* endosiRNA producing locus. Bar plots show the abundance of 19-29nt small RNAs on the *esi-1* locus. Positive and negative values correspond to sense and antisense reads, respectively. Panels show the relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subsets of 20-21nt small RNAs matching the *esi-1* locus.

Figure S4. piRNA distribution on partially homologous paramutations

(Graphs to the left) Length distributions of small RNAs matching *P{lacW}*. Positive and negative values correspond to sense and antisense reads, respectively. **(Graphs to the right)** Relative frequency (z-score) of overlapping sense-antisense small RNA pairs in the subset of 23-28nt small RNAs matching *P{lacW}*. The nature of the cytoplasm inherited in G₁ females by the *BX2* cluster is indicated at the top of the columns. The generation of the test is indicated on the right.

Figure S5. Trans-silencing assay using different target transgenes structures

The transcript produced by the *faf-lacZ* transgene can be targeted by *lacZ* homologous piRNAs but not by *P*-homologous piRNAs whereas the *P-lacZ* fusion transcript produced by the *BQ16* transgene, can be targeted by both types of piRNAs (see Figure 5). *BX2** females having maternally inherited various cytoplasm were tested, together with control females, for their capacities to repress both target transgenes in the female germline (**A**, *faf-lacZ* ; **B**, *BQ16*). The parental cross is indicated with the genotype of female progeny stained. Maternal chromosomal complement is written above the bar. Overnight ovary lacZ staining is shown. A strong lacZ staining is observed for both target with w^{1118} and no lacZ repression is detected with a *BX2* naïve cluster which produces no transgenic piRNAs. *RS3* represses *BQ16* but, as expected, failed to repress *faf-lacZ*. A *BX2** cluster, paramutated by *T-1* completely represses both targets. A strong repression is also observed for both targets with $^{RS3}BX2*$ females despite it is not complete at all generations (Figure 5). Thus, *lacZ* homologous piRNAs produced by $^{RS3}BX2*$ females, which result from the spreading of piRNA production capacity within the cluster, can repress *faf-lacZ* and are therefore functional.

Figure S6. Paramutagenicity of partially homologous paramutation

*BX2** lines having maternally inherited in G_0 piRNA homologous to a subset of the *BX2 P{lacW}* transgene length, and showing strong silencing capacities at the 20th generation, were tested for their ability to *de novo* paramutate a naïve *BX2* locus. These $P^{-1152}BX2*$, $^{RS3}BX2*$, and $P^{-1152+RS3}BX2*$ females were crossed with males carrying a *BX2* naïve cluster and G_1 progeny having paternally, but not maternally, inherited a *BX2* cluster were crossed to establish lines (**A**). These lines carry a second order paramutated locus. These lines were subsequently tested for *trans*-silencing capacities at generations 1, 4 and 36 (**B**). The TSE percentage is indicated with the number of egg chambers counted in parenthesis. Strong silencing capacities are observed over generations showing that $P^{-1152}BX2*$, $^{RS3}BX2*$ and $P^{-1152+RS3}BX2*$ capacities are strongly paramutagenic.

Figure S7. Effect of maternal transmission of piRNA limited to the *P{lacW}* 820bp *P*-element sequence on a *BX2* naïve cluster

BX2 males carrying a *P{lacW}* naïve cluster were crossed to females *Lk-P(1A)* which carry complete *P*-elements inserted in a telomeric piRNA producing locus at the heterozygous state. *Lk-P(1A)* females have a very strong capacity to repress *P*-element transposition and *P*-induced hybrid dysgenesis. Female progeny were recovered which carry the *BX2* locus but did not carry the telomeric *P*-copies. These G₁ females inherited a cytoplasm carrying piRNA homologous to the complete *P*-element, which therefore cover only the first 587bp and the last 233bp of the *BX2 P{lacW}* transgene length (10691bp). Lines were established and the silencing capacities of these putative of ^{Lk-P(1A)}*BX2** lines were tested in subsequent generations using *BQ16* and *faf-lacZ* targets as in Figure 5 and are presented similarly. In addition, deep sequencing of ovarian small RNAs was performed in G₃ and G₅. TSE controls tests performed with *RS3*, *BX2* naïve and ^{T-1}*BX2** are reprinted from Figure 5. TSE assays show that silencing capacities of ^{Lk-P(1A)}*BX2** lines are null at all generations tested. Accordingly, no *P{lacW}* homologous piRNAs were detected at generations 3 and 5. Inheritance of piRNAs homologous to 820bp out of the 10691bp of the *P{lacW}* transgene does not result in *BX2* paramutation.

REFERENCES

1. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, et al. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128: 1089-1103.
2. Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, et al. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315: 1587-1590.
3. 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.
4. Zhang Z, Wang J, Schultz N, Zhang F, Parhad SS, et al. (2014) The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* 157: 1353-1363.
5. Langmead B, Schatz MC, Lin J, Pop M, Salzberg SL (2009) Searching for SNPs with cloud computing. *Genome Biol* 10: R134.
6. Le Thomas A, Rogers AK, Webster A, Marinov GK, Liao SE, et al. (2013) Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* 27: 390-399.
7. Klattenhoff C, Xi H, Li C, Lee S, Xu J, et al. (2009) The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* 138: 1137-1149.
8. Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, et al. (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol Cell* 31: 67-78.
9. Ruby JG, Jan C, Player C, Axtell MJ, Lee W, et al. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 127: 1193-1207.
10. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, et al. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322: 1387-1392.
11. Engels WR (1989) *P* elements in *Drosophila* Berg DE, Howe, M. M., editor. American Society for Microbiology, Washington, D.C. 437-484 p.
12. Finnegan DJ (1989) The I factor and I-R Hybrid Dysgenesis in *Drosophila melanogaster*. *Mobile DNA Chp.* 18: 503-518.
13. Bucheton A, Paro R, Sang HM, Pelisson A, Finnegan DJ (1984) The molecular basis of I-R hybrid dysgenesis in *Drosophila melanogaster*: identification, cloning, and properties of the I factor. *Cell* 38: 153-163.
14. Karess RE, Rubin GM (1984) Analysis of P transposable element functions in *Drosophila*. *Cell* 38: 135-146.
15. Bucheton A (1978) Non-Mendelian female sterility in *Drosophila melanogaster*: influence of ageing and thermic treatments. I. Evidence for a partly inheritable effect of these two factors. *Heredity* 41: 357-369.
16. Engels W R (1979) Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. *Genet Research* 33: 219-236.
17. Picard G (1976) Non-mendelian female sterility in *Drosophila melanogaster*: hereditary transmission of I factor. *Genetics* 83: 107-123.
18. Ronsseray S, Lemaitre B, Coen D (1993) Maternal inheritance of P cytotype in *Drosophila melanogaster*: a "pre-P cytotype" is strictly extra-chromosomally transmitted. *Mol Gen Genet* 241: 115-123.
19. Coen D, Lemaitre B, Delattre M, Quesneville H, Ronsseray S, et al. (1994) *Drosophila* P element: transposition, regulation and evolution. *Genetica* 93: 61-78.
20. Khurana JS, Wang J, Xu J, Koppetsch BS, Thomson TC, et al. (2011) Adaptation to P Element Transposon Invasion in *Drosophila melanogaster*. *Cell* 147: 1551-1563.

21. Ronsseray S, Anxolabehere D, Periquet G (1984) Hybrid dysgenesis in *Drosophila melanogaster*: influence of temperature on cytotype determination in the P-M system. *Mol Gen Genet* 196: 17-23.
22. de Vanssay A, Bouge AL, Boivin A, Hermant C, Teyssset L, et al. (2012) Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature AOP*: Aug 26. doi: 10.1038/nature11416.
23. Chandler VL (2007) Paramutation: from maize to mice. *Cell* 128: 641-645.
24. Simmons MJ, Ragatz LM, Sinclair IR, Thorp MW, Buschette JT, et al. (2012) Maternal enhancement of cytotype regulation in *Drosophila melanogaster* by genetic interactions between telomeric P elements and non-telomeric transgenic P elements. *Genet Res (Camb)* 94: 339-351.
25. Ronsseray S, Marin L, Lehmann M, Anxolabehere D (1998) Repression of hybrid dysgenesis in *Drosophila melanogaster* by combinations of telomeric P-element reporters and naturally occurring P elements. *Genetics* 149: 1857-1866.
26. Le Thomas A, Marinov GK, Aravin AA (2014) A transgenerational process defines piRNA biogenesis in *Drosophila virilis*. *Cell Rep* 8: 1617-1623.
27. de Vanssay A, Bouge AL, Boivin A, Hermant C, Teyssset L, et al. (2013) piRNAs and epigenetic conversion in *Drosophila*. *Fly (Austin)* 7: 237-241.
28. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, et al. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313: 320-324.
29. Ishizu H, Siomi H, Siomi MC (2012) Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev* 26: 2361-2373.
30. Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, et al. (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* 491: 284-287.
31. 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.
32. Huang H, Gao Q, Peng X, Choi SY, Sarma K, et al. (2011) piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. *Dev Cell* 20: 376-387.
33. 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.
34. Olovnikov I, Ryazansky S, Shpiz S, Lavrov S, Abramov Y, et al. (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.
35. Josse T, Maurel-Zaffran C, de Vanssay A, Teyssset L, Todeschini AL, et al. (2008) Telomeric trans-silencing in *Drosophila melanogaster*: tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS ONE* 3: e3249.
36. Josse T, Teyssset L, Todeschini AL, Sidor CM, Anxolabehere D, et al. (2007) Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* 3: 1633-1643.
37. 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.
38. Ronsseray S, Josse T, Boivin A, Anxolabehere D (2003) Telomeric transgenes and trans-silencing in *Drosophila*. *Genetica* 117: 327-335.
39. Muerdter F, Olovnikov I, Molaro A, Rozhkov NV, Czech B, et al. (2011) Production of artificial piRNAs in flies and mice. *Rna*.
40. Todeschini AL, 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.
41. Pöyhönen M, de Vanssay A, Delmarre V, Hermant C, Todeschini AL, et al. (2012) Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3 (Bethesda)* 2: 331-338.

42. Dufourt J, Dennis C, Boivin A, Gueguen N, Theron E, et al. (2013) Spatio-temporal requirements for transposable element piRNA-mediated silencing during *Drosophila* oogenesis. *Nucleic Acids Res* 42: 2512-2524.
43. Fischer-Vize JA, Rubin GM, Lehmann R (1992) The fat facets gene is required for *Drosophila* eye and embryo development. *Development* 116: 985-1000.
44. Rouget C, Papin C, Boureux A, Meunier AC, Franco B, et al. (2010) Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* 467: 1128-1132.
45. Akkouche A, Grentzinger T, Fablet M, Armenise C, Burlet N, et al. (2013) Maternally deposited germline piRNAs silence the tirant retrotransposon in somatic cells. *EMBO Rep* 14: 458-464.
46. Gu T, Elgin SC (2013) Maternal depletion of Piwi, a component of the RNAi system, impacts heterochromatin formation in *Drosophila*. *PLoS Genet* 9: e1003780.
47. Dorer DR, Henikoff S (1994) Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* 77: 993-1002.
48. Dorer DR, Henikoff S (1997) Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. *Genetics* 147: 1181-1190.
49. Fanti L, Berloco M, Piacentini L, Pimpinelli S (2003) Chromosomal distribution of heterochromatin protein 1 (HP1) in *Drosophila*: a cytological map of euchromatic HP1 binding sites. *Genetica* 117: 135-147.
50. Ronsseray S, Boivin A, Anxolabehere D (2001) P-Element repression in *Drosophila melanogaster* by variegating clusters of P-lacZ-white transgenes. *Genetics* 159: 1631-1642.
51. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, et al. (2009) Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137: 522-535.
52. Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, et al. (2008) An endogenous small interfering RNA pathway in *Drosophila*. *Nature* 453: 798-802.
53. Zhang F, Wang J, Xu J, Zhang Z, Koppetsch BS, et al. (2012) UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* 151: 871-884.
54. Lemaître 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.
55. O'Kane CJ, Gehring WJ (1987) Detection in situ of genomic regulatory elements in *Drosophila*. *Proc Natl Acad Sci U S A* 84: 9123-9127.
56. Dufourt J, Brasset E, Desset S, Pouchin P, Vaury C (2011) Polycomb Group-Dependent, Heterochromatin Protein 1-Independent, Chromatin Structures Silence Retrotransposons in Somatic Tissues Outside Ovaries. *DNA Res*.
57. Schupbach T, Wieschaus E (1991) Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* 129: 1119-1136.
58. Lee YS, Nakahara K, Pham JW, Kim K, He Z, et al. (2004) Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* 117: 69-81.
59. Volpe AM, Horowitz H, Grafer CM, Jackson SM, Berg CA (2001) *Drosophila* rhino encodes a female-specific chromo-domain protein that affects chromosome structure and egg polarity. *Genetics* 159: 1117-1134.
60. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.

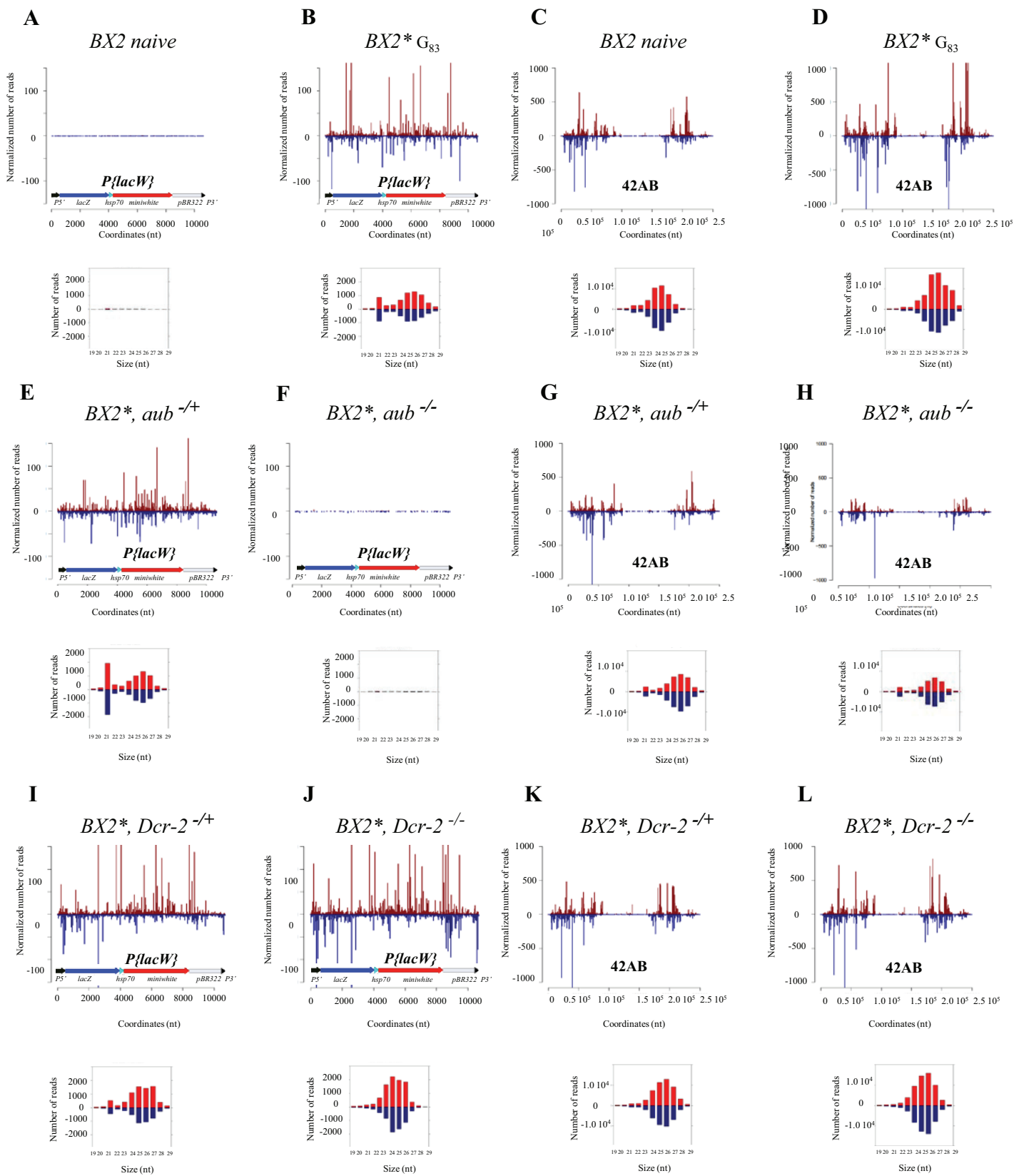


Fig. 1 - Hermant *et al.*

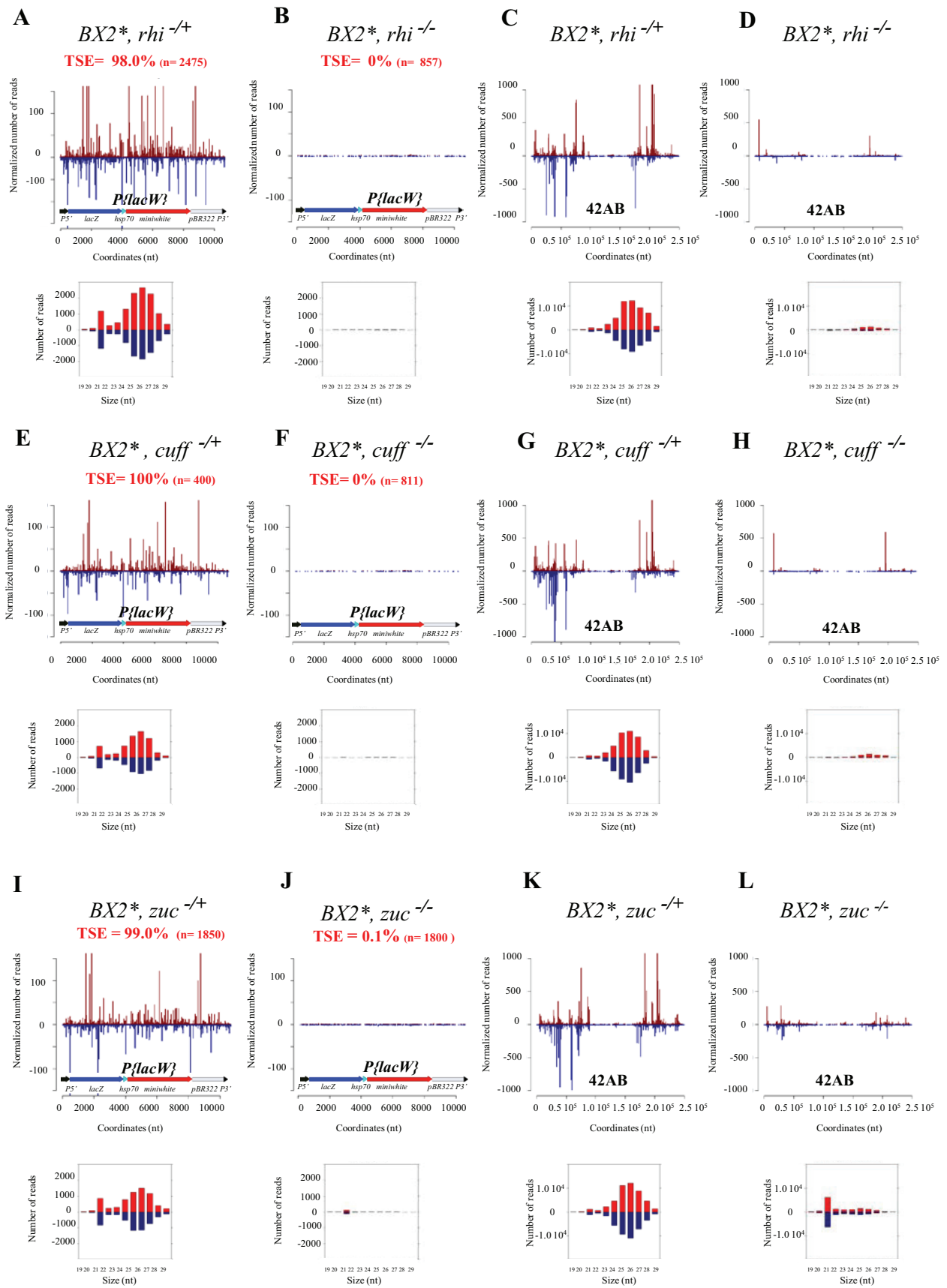


Fig. 2 - Hermant *et al.*

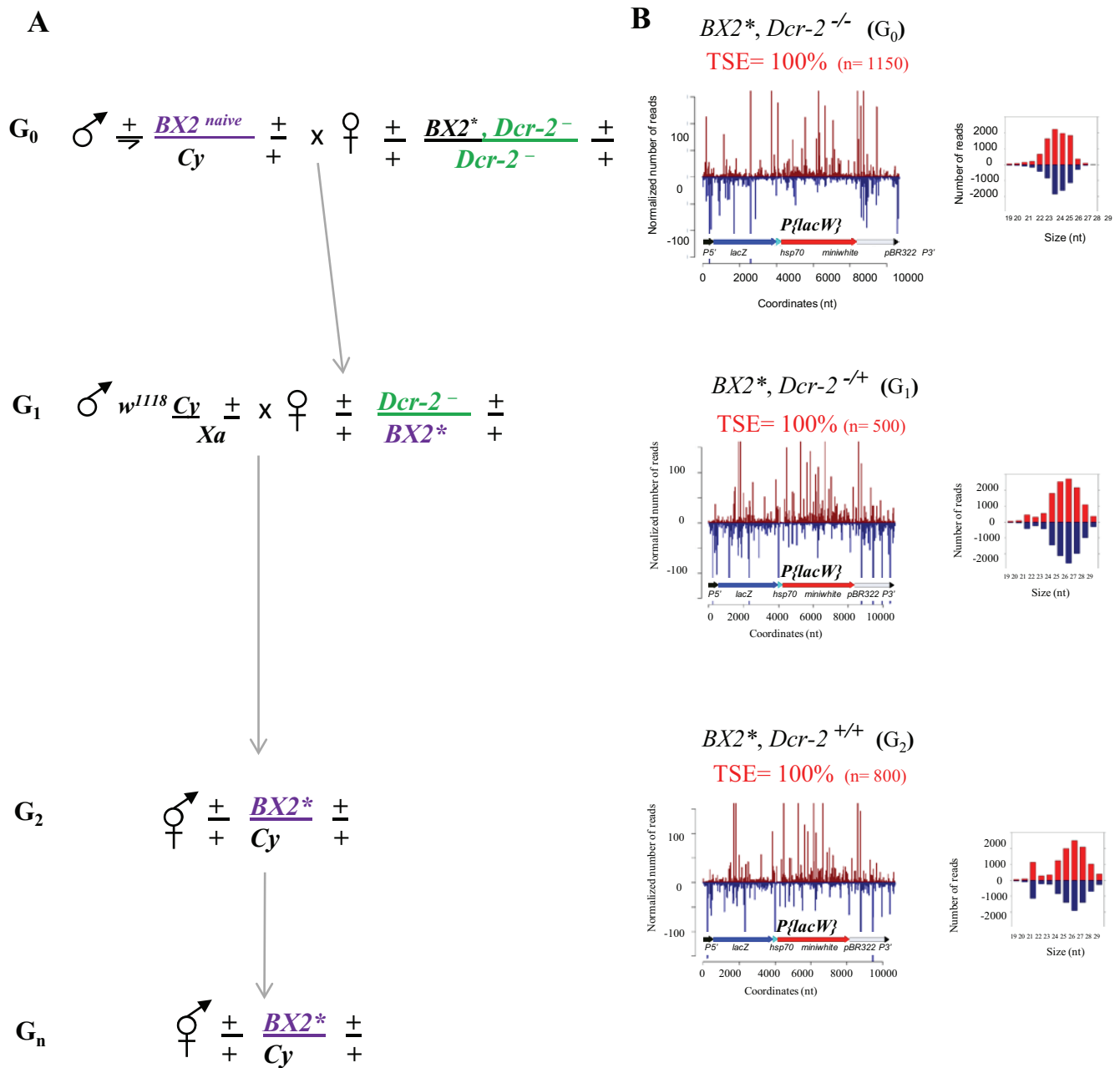


Fig. 3 - Hermant *et al.*

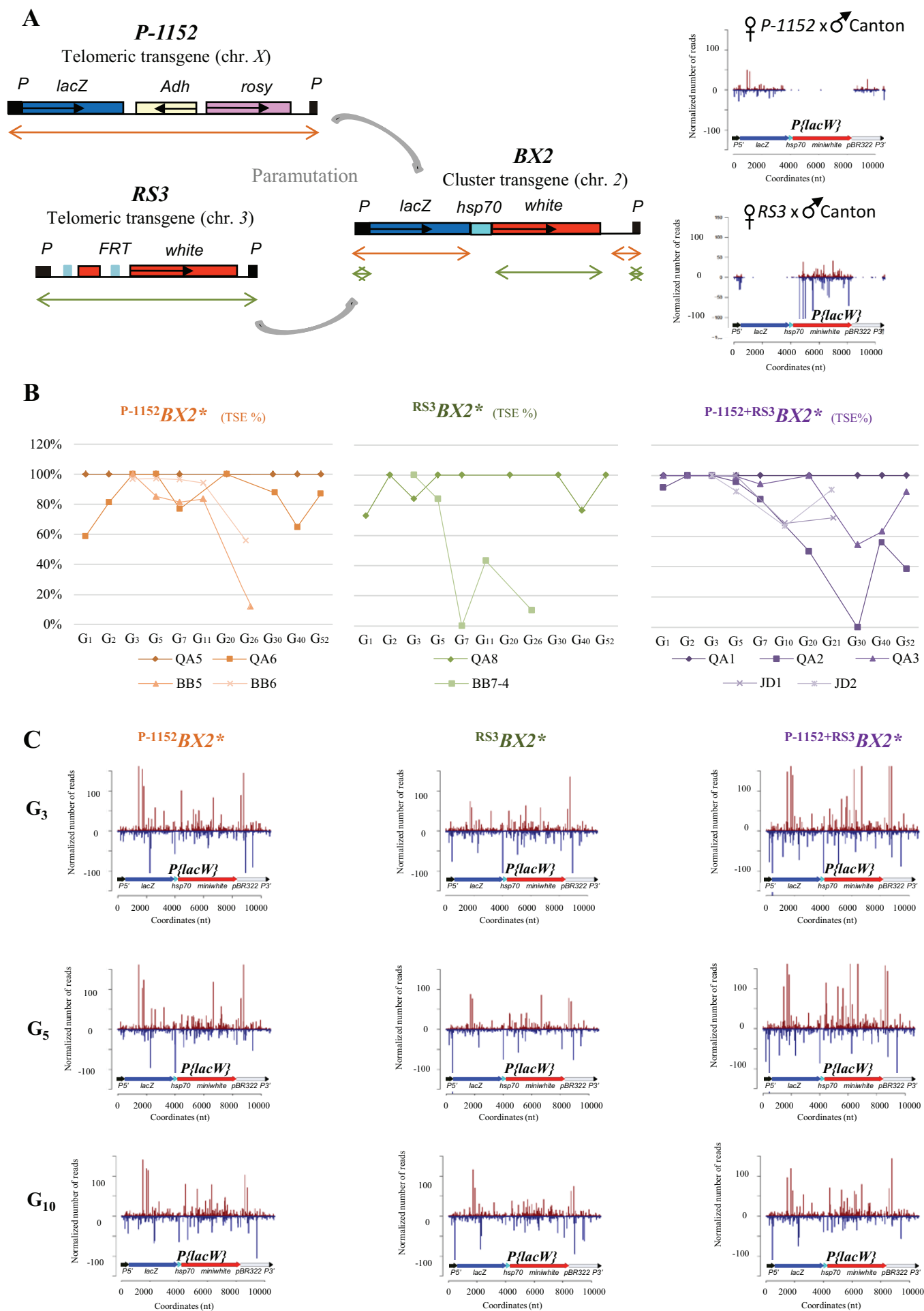
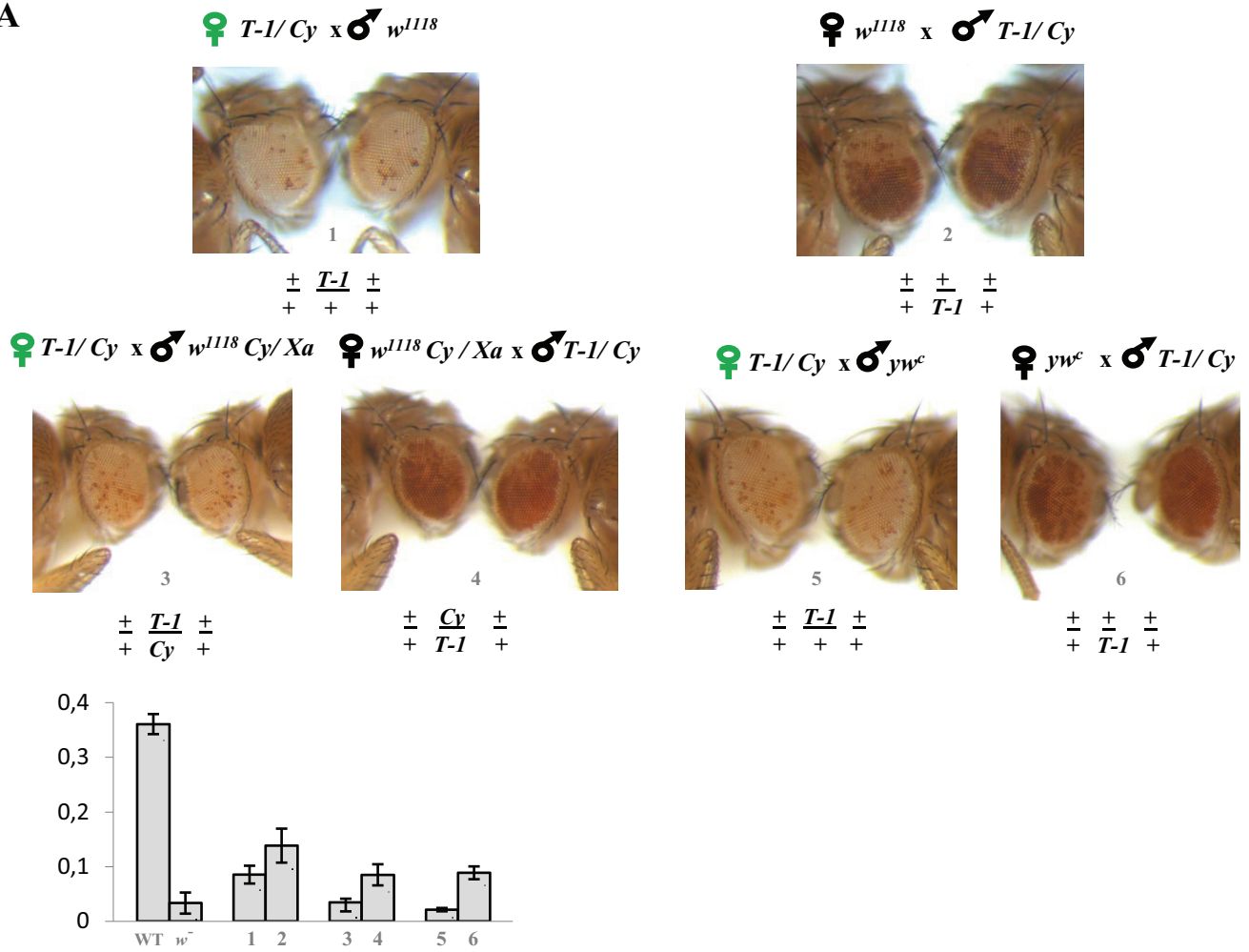


Fig. 4 - Hermant *et al.*

♀ ♂		<i>faf-lacZ</i> target	<i>BQ16</i> target
		P white faf lacZ P	P lacZ rosy P
RS3 P FRT white P		0% (n=275)	100% (n> 2000)
BX2 naive P lacZ white P piRNA		0% (n>2000)	0% (n> 2000)
T-1 BX2* P lacZ white P		100% (n>2000)	100% (n> 2000)
RS3 BX2* P lacZ white P Functional piRNAs?	QA8 G ₂₁	100% (n= 600)	100% (n= 1550)
	QA8 G ₄₀	43.8% (n= 400)	76.6% (n= 1175)
	BB7-4 G ₃	87.5% (n= 800)	100% (n= 850)
	BB7-4 G ₅	100% (n= 1000)	84.2% (n= 950)
	BB7-4 G ₇	34.4% (n= 800)	0% (n= 1100)
	BB7-4 G ₁₁	51.2% (n= 1025)	43.2% (n= 925)

Fig. 5 - Hermant *et al.*

A



B

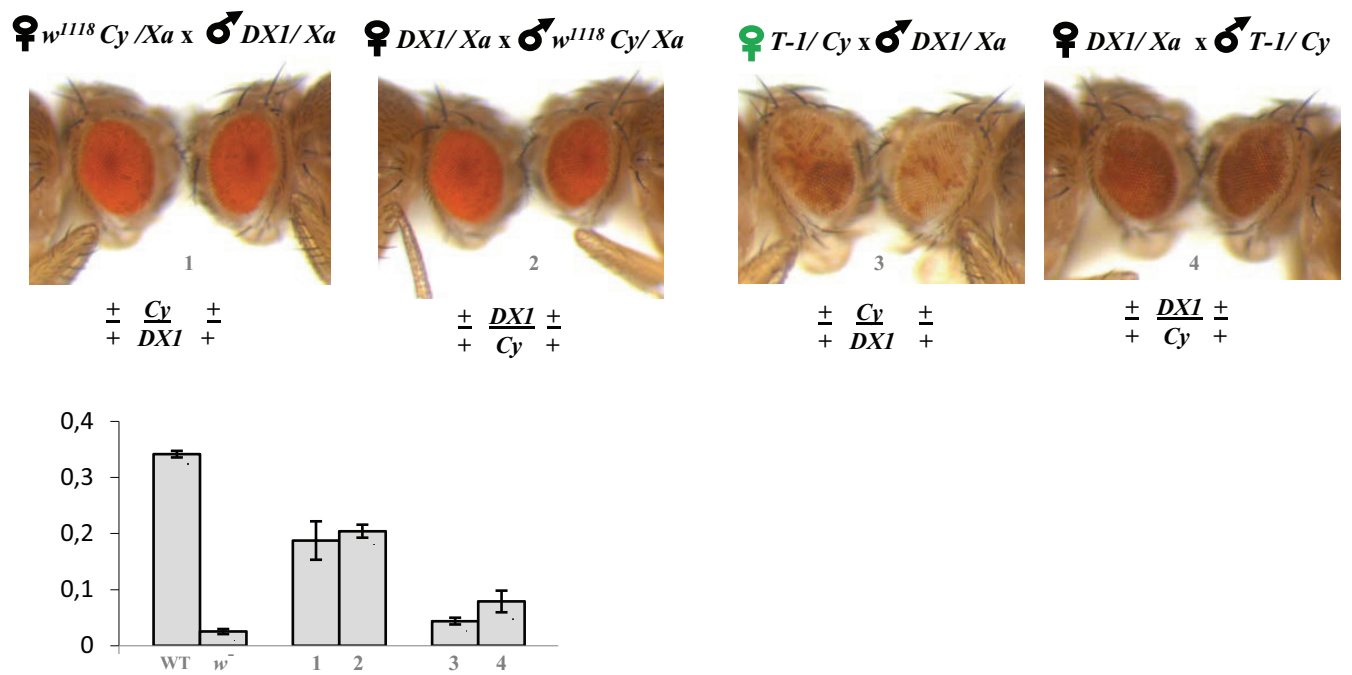


Fig. 6- Hermant *et al.*

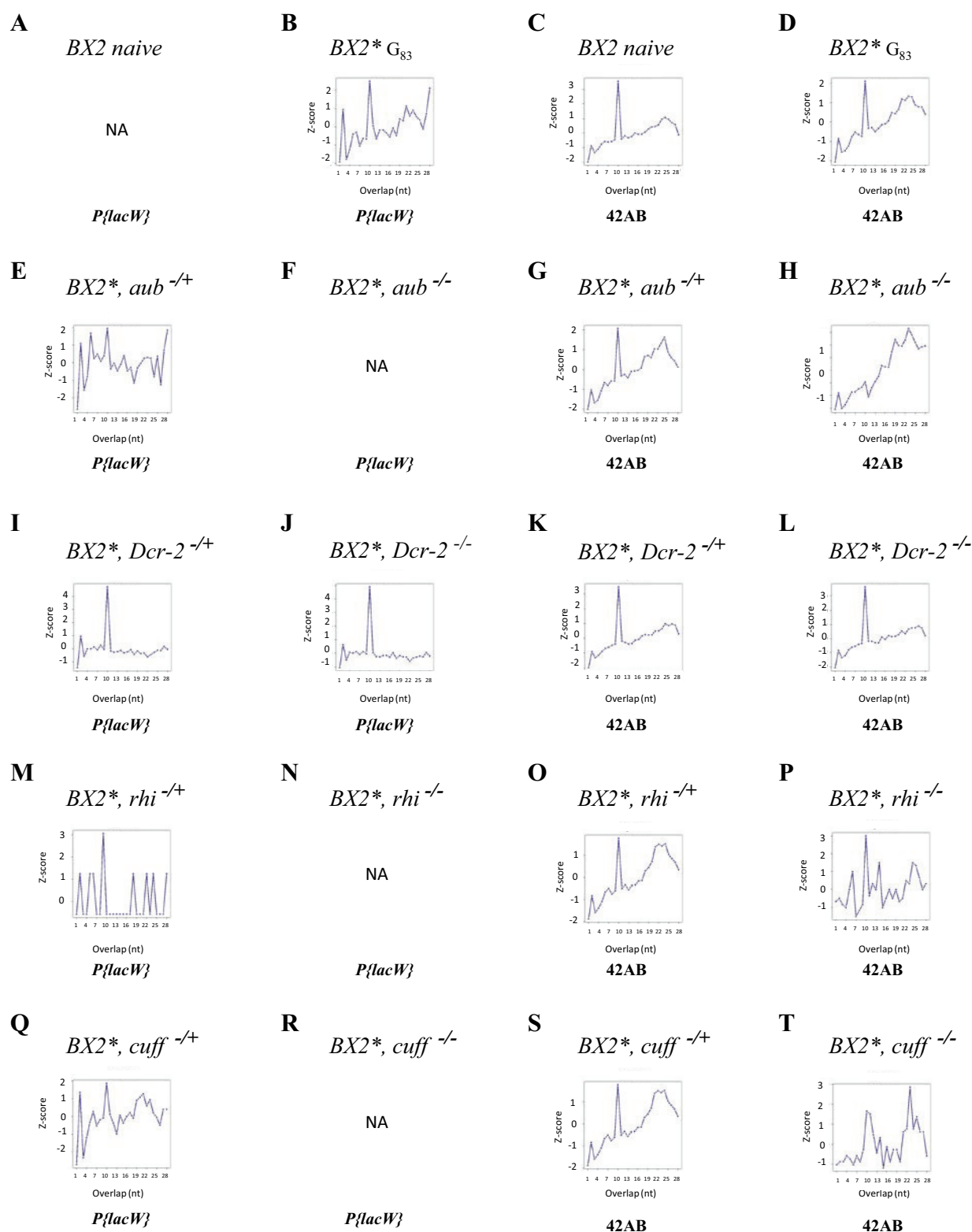
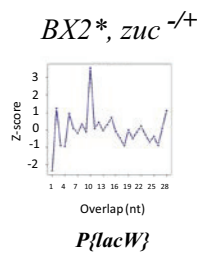
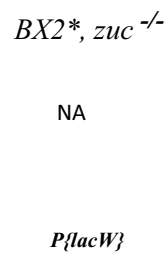


Fig. S1 - Hermant *et al.*

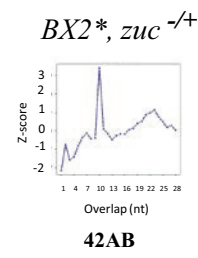
U



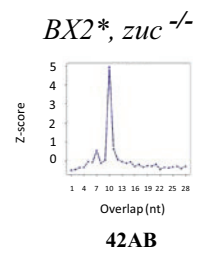
V



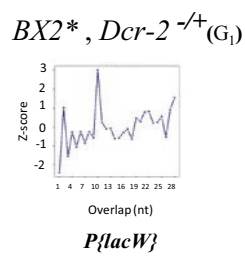
W



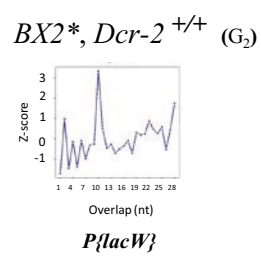
X



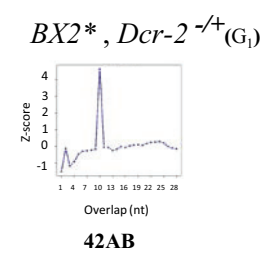
Y



Z



AA



AB

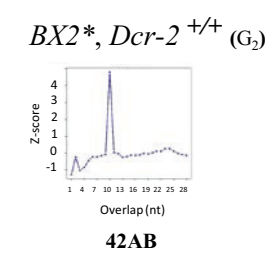


Fig. S1 - continued - Hermant *et al.*

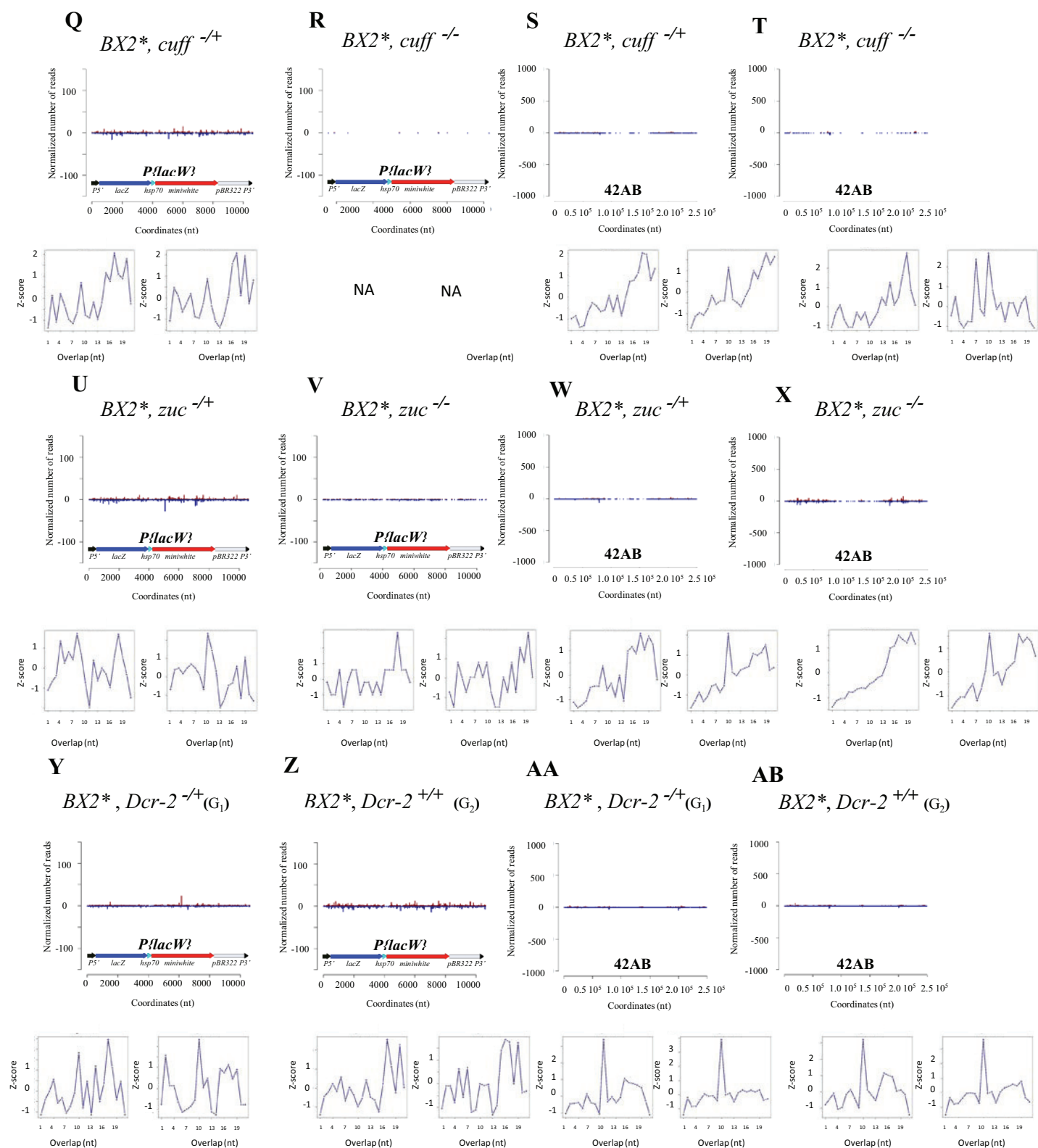


Fig. S2 – continued - Hermant *et al.*

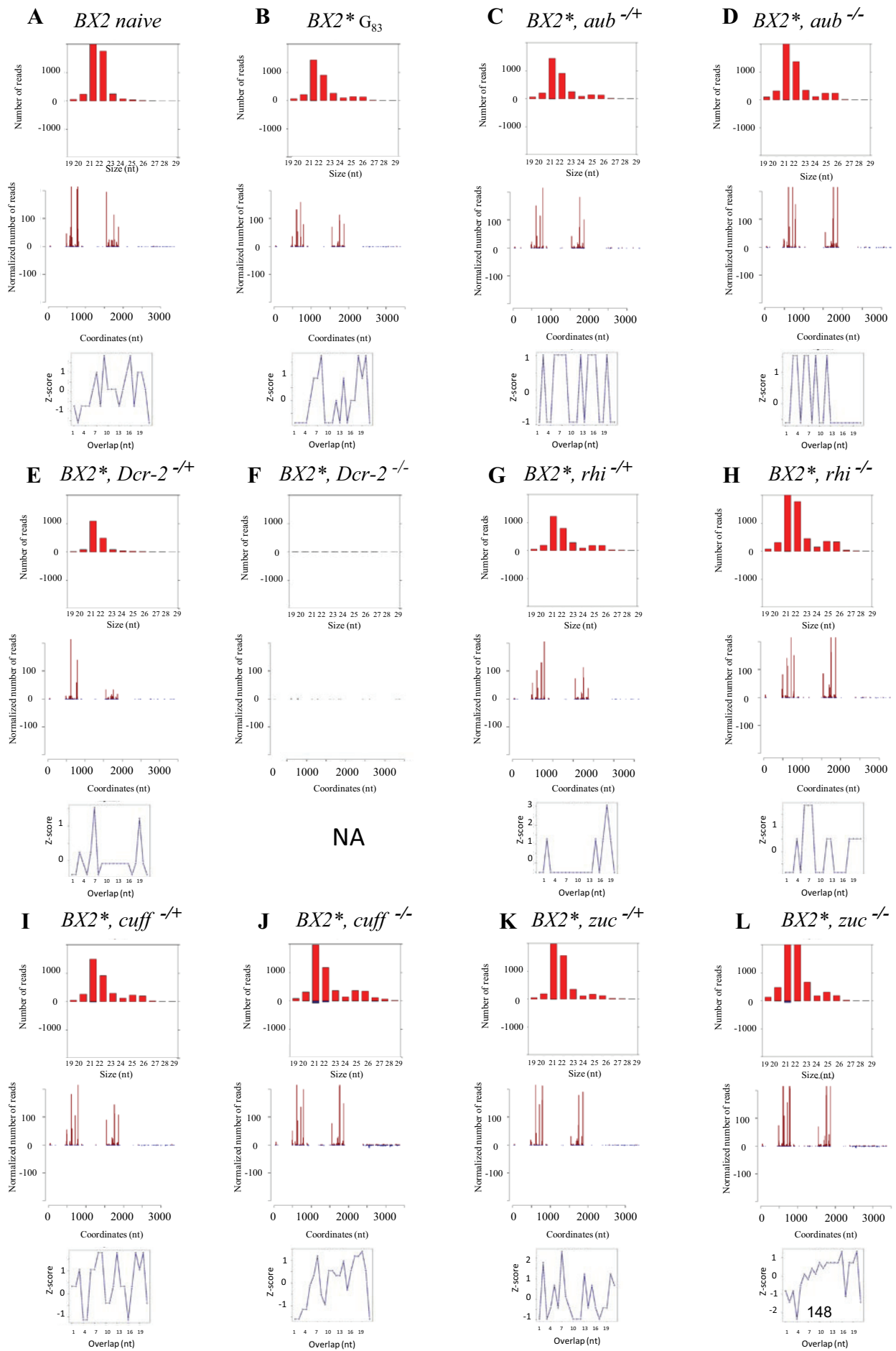


Fig. S3 - Hermant *et al.*

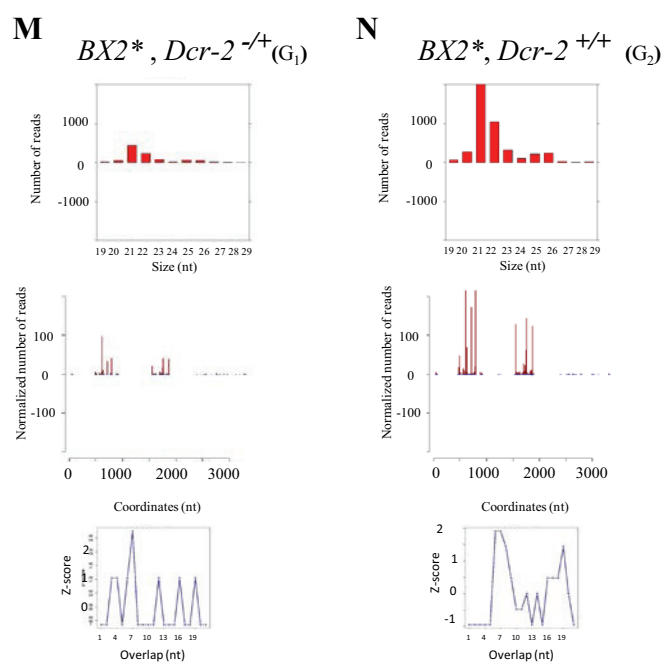


Fig. S3 - continued - Hermant *et al.*

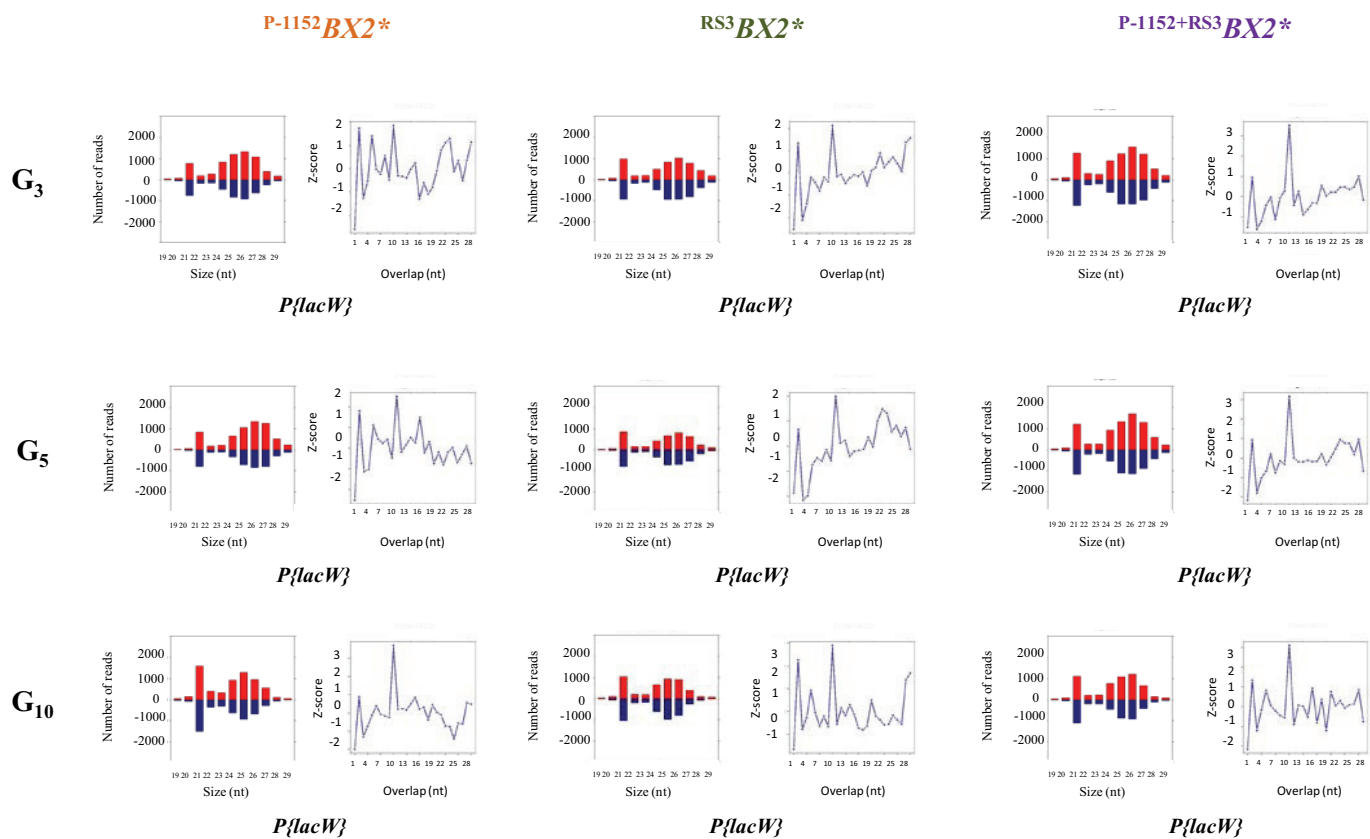
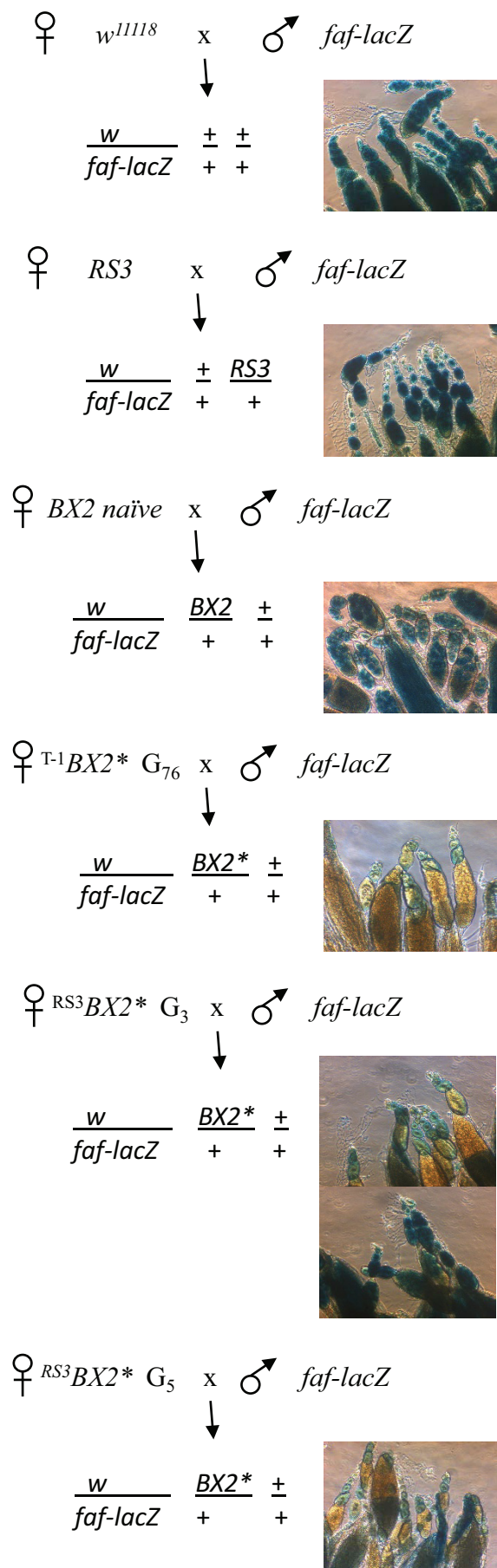


Fig. S4 - Hermant *et al.*

A



B

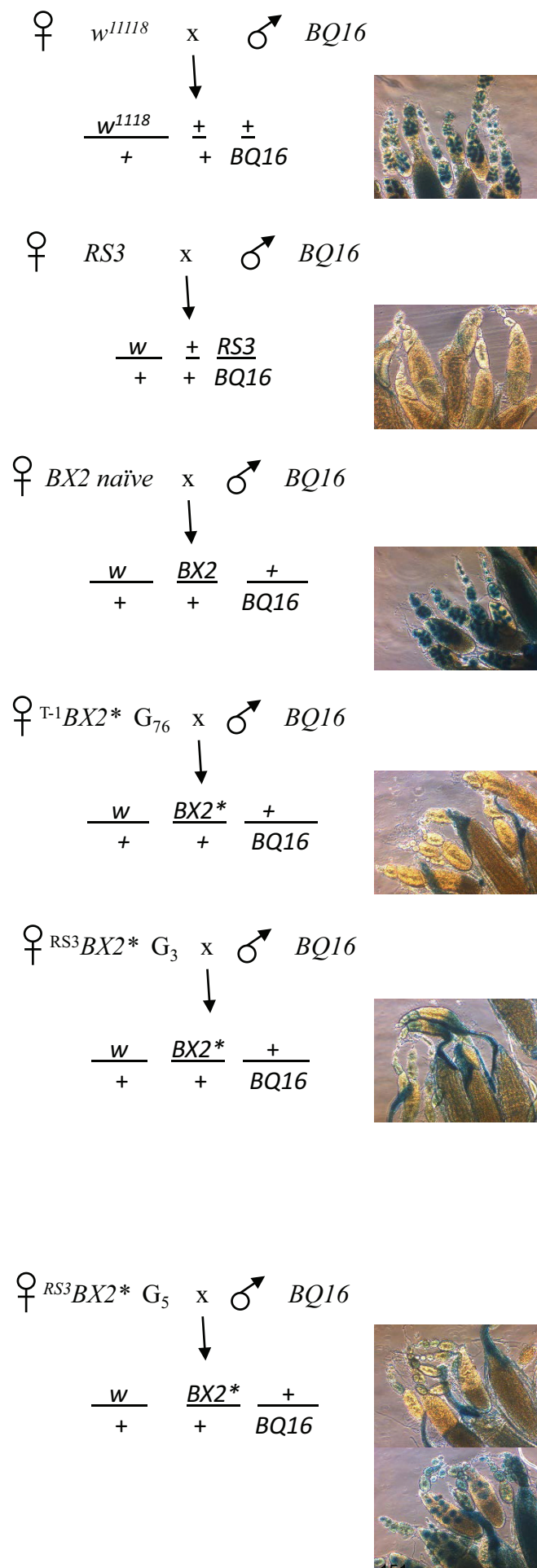
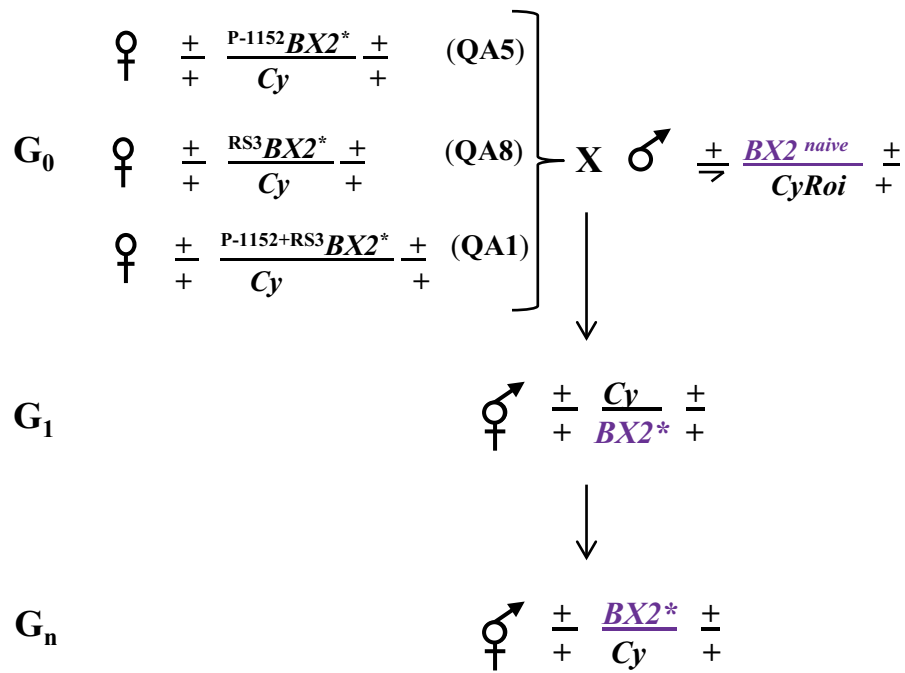


Fig. S5 - Hermant *et al.*

A



B

		TSE percentage
QA5 <i>BX2*</i>	G ₁	100% (n= 950)
	G ₄	100% (n= 1350)
	G ₃₆	100% (n= 1450)
QA8 <i>BX2*</i>	G ₁	100% (n= 1025)
	G ₄	100% (n= 1775)
	G ₃₆	95.0% (n= 1300)
QA1 <i>BX2*</i>	G ₁	100% (n= 800)
	G ₄	100% (n= 1175)
	G ₃₆	92.1% (n= 1275)

Fig. S6 - Hermant *et al.*

A

♀ ♂		<i>faf-lacZ</i> target	<i>BQ16</i> target
		<i>P</i> <i>white</i> <i>faf</i> <i>lacZ</i> <i>P</i>	<i>P</i> <i>lacZ</i> <i>rosy</i> <i>P</i>
<i>RS3</i> <i>P</i> <i>FRT</i> <i>white</i> <i>P</i>		0% (n= 275)	100% (n> 2000)
<i>BX2</i> naive <i>P</i> <i>lacZ</i> <i>white</i> <i>P</i> ♂ piRNA		0% (n> 2000)	0% (n> 2000)
<i>T-1BX2*</i> <i>P</i> <i>lacZ</i> <i>white</i> <i>P</i>		100% (n> 2000)	100% (n> 2000)
<i>Lk-P(1A)BX2*</i> <i>P</i> <i>lacZ</i> <i>white</i> <i>P</i> piRNA production and trans-silencing capacities?	<i>Lk-P(1A)BX2</i> G ₃	0% (n> 2000)	0% (n> 2000)
	<i>Lk-P(1A)BX2</i> G ₅	0% (n> 2000)	0% (n> 2000)
	<i>Lk-P(1A)BX2</i> G ₇	0% (n> 2000)	0% (n> 2000)
	<i>Lk-P(1A)BX2</i> G ₁₁	0% (n> 2000)	0% (n> 2000)
	<i>Lk-P(1A)BX2</i> G ₂₆	ND	0% (n> 2000)

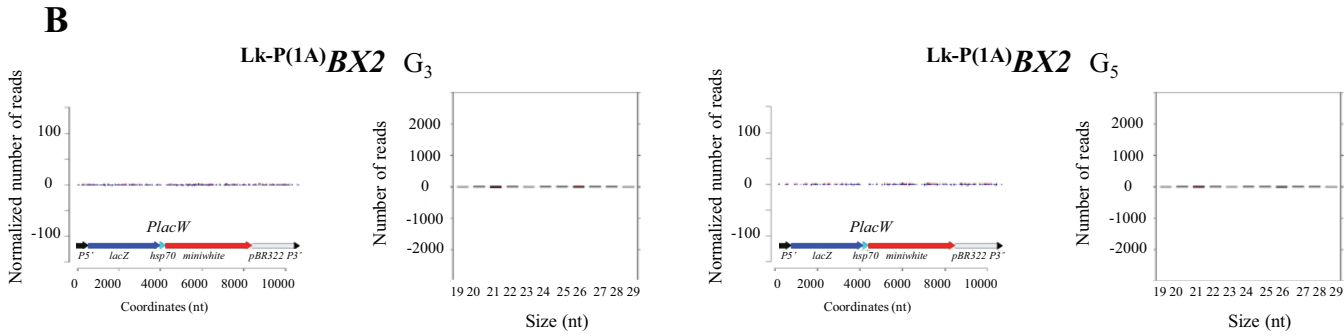


Fig. S7 - Hermant *et al.*

Library ID	1	2	3	4	5	6	7	8	9	10	11	12	13
Parental origin	$BX2^*/Cy$ x Contop ^r	$BX2^*/CyRoI$ x w ¹¹⁸	$BX2^*/CyRoI$ x w ¹¹⁸	$BX2^*, au\beta^{Gc12}/Cy$ x au β^{N1}/Xa	$BX2^*, au\beta^{Gc12}/Cy$ x au β^{N1}/Xa	$BX2^*, Der-2^{m10X}/Cy$ x Der-2 ^{m10X} /Cy	$BX2^*, Der-2^{m10X}/Cy$ x Der-2 ^{m10X} /Cy	$BX2^*, rh1^2/Cy$ x rh1 ^{SG} /Cy	$BX2^*, rh1^2/Cy$ x rh1 ^{SG} /Cy	$BX2^*, cu\beta^{G037}/Cy$ x cu β^{N1m23}/Cy	$BX2^*, cu\beta^{G037}/Cy$ x cu β^{N1m23}/Cy	$BX2^*, zue^{m27}/Cy$ x zue ^{SG63} /Cy	$BX2^*, zue^{m27}/Cy$ x zue ^{SG63} /Cy
Female genotype	$BX2^+$	$BX2^*/C_{GS}/+$	$CyRoI/+$	$BX2^*, au\beta^{Gc12}/Xa$	$BX2^*, au\beta^{Gc12}/Xa$	$BX2^*, Der-2^{m10X}/Cy$	$BX2^*, Der-2^{m10X}/Cy$	$BX2^*, rh1^2/Cy$	$BX2^*, rh1^2/Cy$	$BX2^*, cu\beta^{G037}/Cy$	$BX2^*, cu\beta^{G037}/Cy$	$BX2^*, zue^{m27}/Cy$	$BX2^*, zue^{m27}/Cy$
Dmel_matched	20 491 721	15 675 824	14 222 359	12 756 621	11 489 237	26 206 014	17 956 124	16 245 668	11 825 256	13 389 291	13 685 445	20 131 028	15 216 228
miRNAs	34.15 %	24.67 %	22.55 %	12.03 %	18.24 %	17.79 %	20.79 %	20.88 %	23.26 %	25.98 %	31.16 %	17.49 %	24.26 %
tRNAs	5.41 %	2.33 %	1.47 %	1.01 %	1.79 %	4.05 %	5.39 %	3.47 %	3.23 %	1.86 %	7.65 %	4.80 %	6.37 %
ncRNAs	0.85 %	0.62 %	0.66 %	0.82 %	0.86 %	0.50 %	0.64 %	1.69 %	2.12 %	0.63 %	1.44 %	0.72 %	2.09 %
miscRNAs	6.93 %	4.87 %	6.50 %	62.25 %	60.31 %	28.34 %	18.42 %	11.94 %	37.56 %	5.94 %	22.81 %	16.45 %	42.58 %
Transposons	38.18 %	51.17 %	51.44 %	17.11 %	12.41 %	36.60 %	40.55 %	44.61 %	19.85 %	49.48 %	22.49 %	45.40 %	16.07 %
Introns	4.88 %	5.01 %	5.30 %	2.10 %	2.04 %	4.53 %	5.02 %	5.01 %	3.19 %	4.17 %	2.46 %	4.79 %	1.88 %
Transcripts	2.27 %	2.24 %	2.57 %	1.36 %	1.51 %	1.77 %	2.19 %	3.34 %	7.41 %	4.25 %	8.37 %	2.18 %	2.79 %
Intergenic	6.34 %	8.26 %	8.56 %	2.86 %	2.46 %	5.64 %	6.17 %	8.13 %	2.71 %	6.58 %	2.69 %	7.28 %	2.85 %
Remaining	0.99 %	0.83 %	0.95 %	0.45 %	0.38 %	0.79 %	0.84 %	0.93 %	0.68 %	1.09 %	0.93 %	0.89 %	1.12 %
Normalisation factor	0.2424497	0.2993644	0.3327195	0.9293427	1	0.245758	0.3183141	0.316918	0.6219892	0.3527794	0.4576677	0.274712545	0.56064162

Library ID	14	15	16	17	18	19	20	21	22	23	24	25	26
Parental origin	$BX2^*/Der-2^{m10X}/Cy$ x Contop ^r	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸	$BX2^*/Der-2^{m10X}/Cy$ x w ¹¹⁸
Female genotype	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$	$BX2^*/Der-2^{m10X}/Cy$
Dmel_matched	17 877 800	16 197 143	17 350 013	17 979 952	15 409 205	16 279 059	18 406 856	13 852 712	18 297 477	58 044 195	111 170 832	21 495 883	18 739 203
miRNAs	22.41 %	21.50 %	16.23 %	13.60 %	26.32 %	24.06 %	22.04 %	19.61 %	13.72 %	15.61 %	14.34 %	22.48 %	16.82 %
tRNAs	4.19 %	2.15 %	0.64 %	0.99 %	1.24 %	1.86 %	3.42 %	1.07 %	3.38 %	3.64 %	0.89 %	0.60 %	0.89 %
ncRNAs	0.88 %	0.82 %	0.57 %	0.70 %	0.92 %	0.96 %	1.70 %	0.77 %	2.39 %	2.31 %	0.57 %	0.50 %	0.55 %
miscRNAs	10.84 %	8.35 %	51.56 %	53.83 %	27.77 %	9.10 %	14.59 %	49.06 %	28.52 %	23.75 %	59.30 %	30.86 %	52.35 %
Transposons	45.66 %	48.39 %	22.66 %	22.77 %	32.74 %	47.41 %	40.81 %	22.00 %	35.22 %	37.94 %	18.54 %	34.51 %	21.78 %
Introns	5.01 %	6.31 %	2.53 %	2.42 %	3.62 %	5.18 %	6.11 %	2.46 %	5.61 %	5.72 %	2.04 %	3.29 %	2.23 %
Transcripts	3.30 %	3.20 %	1.33 %	1.45 %	1.60 %	2.36 %	2.86 %	1.24 %	3.58 %	3.33 %	0.93 %	1.56 %	1.23 %
Intergenic	6.98 %	8.29 %	3.89 %	3.73 %	5.02 %	8.08 %	7.54 %	3.24 %	6.81 %	6.85 %	2.94 %	5.40 %	3.59 %
Remaining	0.73 %	1.01 %	0.59 %	0.50 %	0.76 %	1.01 %	0.93 %	0.54 %	0.76 %	0.84 %	0.45 %	0.80 %	0.57 %
Normalisation factor	0.2866747	0.3003971	0.5251579	0.5360289	0.3981389	0.3037089	0.2885683	0.6304115	0.3494993	0.1033353	0.979384	0.2955765	0.4970375

Table S1 - Hermant *et al.*

Line	Library ID	total piRNAs (24-29nt)	Normalized total piRNAs (24-29nt)	Forward piRNA ratio	piRNA density	z-score piRNA	Ufreq (forward reverse)	Total siRNAs (20-21nt)	Normalized total siRNAs (20-21nt)	Forward siRNA ratio	siRNA density	z-score siRNA	piRNA/siRNA ratio
<i>BX2</i> naïve	1	214	51.88	0.4953271	0.0200187	1.7128952	65.094396 65.444444	339	82.19	0.5221238	0.0317118	2.3419732	0.38698010
<i>BX2*</i> <i>G</i> ₈₃	2	27 353	8 188.51	0.5882352	2.5587464	2.5181246	66.183965 71.179568	6 313	1 889.89	0.4953271	0.5905519	0.5790182	0.8124814
<i>CyRoi/+</i>	3	21	6.99	0.71428571	0.0019645	0	53.333333 83.333333	3	0.99	1	0.00028064	0	0.875
<i>BX2*</i> , <i>aub</i> ^{+/+}	4	7 929	7 368.75	0.5860764	0.7417212	1.7671161	67.807187 64.990892	4 353	4 045.43	0.5065472	0.4072029	-0.2981412	0.64557889
<i>BX2*</i> , <i>aub</i> ^{-/-}	5	85	85.00	0.3176470	0.0079513	-0.278543	62.962962 68.965517	45	45.00	0.5777777	0.00420954	-0.6236095	0.6538461
<i>BX2*</i> , <i>Dcr-2</i> ^{+/+}	6	40 917	10 056.40	0.6158564	3.8275958 84	4.4648073	69.978967 71.898460	4 587	1 127.37	0.5410943	0.4290926	2.0445663	0.89919567
<i>BX2*</i> , <i>Dcr-2</i> ^{-/-}	7	43 954	13 991.17	0.5814715	4.1116931	4.6378718	67.794819 74.075860	1 170	372.43	0.6196581	0.1094480	4.2014314	0.9740714
<i>BX2*</i> , <i>zuc</i> ^{+/+}	8	33 987	10 771.09	0.5653043	3.1793264	3.3958016	63.420606 73.430580	6 616	2 096.73	0.5078597	0.6188961	2.4810759	0.8370563
<i>BX2*</i> , <i>zuc</i> ^{-/-}	9	79	49.13	0.531645	0.0073900	-0.4685212	52.380932 54.054054	458	284.87	0.5	0.0428437	0.8125	0.1471135
<i>BX2*</i> , <i>cuff</i> ^{+/+}	10	10 765	3 797.67	0.6863910	1.0070159	0.3553217	68.615095 72.727272	1 997	704.50	0.5032548	0.1868101	1.5890444	0.8435198
<i>BX2*</i> , <i>cuff</i> ^{-/-}	11	99	45.30	0.3333333	0.0092609	-0.6324555	62.218928 59.090909	12	5.49	0.5833333	0.0011225	-0.2041241	0.8918918
<i>BX2*</i> , <i>rhl</i> ^{+/+}	12	52 653	14 646.43	0.5938313	4.9254443	1.8035804	62.218928 69.275320	8 096	2 224.07	0.4990118	0.7573433	0.0772238	0.866730
<i>BX2*</i> , <i>rhl</i> ^{-/-}	13	173	96.99	0.4219653	0.0161833	-0.7071067	65.753426 62.0	19	10.65	0.4736842	0.0017773	-0.2041241	0.9010416
<i>BX2</i> ^{+/+} <i>Dcr2</i> ^{-/-} <i>BX2*</i> <i>G</i> ₁	14	70 357	20 169.57	0.5322569	6.5815715	2.9468850	64.433347 70.971467	3 616	1 036.62	0.5306969	0.3382600	1.9479326	0.9511172
<i>BX2</i> ^{+/+} <i>Dcr2</i> ^{-/-} <i>BX2*</i> <i>G</i> ₂	15	52 644	15 814.10	0.5846250	4.9246024	3.2468979	66.799880 70.416609	8 269	2 483.98	0.4961905 9	0.7735266	-0.3077964	0.8642490
<i>P-1152</i> <i>BX2*</i> <i>G</i> ₃	16	15 451	8 114.21	0.6086337	1.4453695	1.9379303	61.814216 70.927327	3 160	1 659.50	0.5037974	0.2956033	0.6967465	0.8302079
<i>P-1152</i> <i>BX2*</i> <i>G</i> ₅	17	15 222	8 159.43	0.6129286	1.4239476	2.5077637	63.708463 69.025797	3 216	1 723.87	0.5071517	0.3008419	0.8655784	0.8255776
<i>P-1152</i> <i>BX2*</i> <i>G</i> ₁₀	18	16 368	6 516.74	0.5960410	1.5311506	3.6543741	61.767176 69.797381	8 341	3 320.88	0.5191224	0.7802619	2.77637213	0.6624306
<i>RS3</i> <i>BX2*</i> <i>G</i> ₃	19	24 607	7 473.36	0.5007924	2.3018709	2.1843067	64.351231 73.184630	6 805	2 066.74	0.5108008	0.6365762	0.0430727	0.7833630
<i>RS3</i> <i>BX2*</i> <i>G</i> ₅	20	18 827	5 432.88	0.5176608	1.7611786	1.8166883	66.324646 76.605054	6 039	1 742.66	0.5129988	0.5649204	0.4567986	0.7571382
<i>RS3</i> <i>BX2*</i> <i>G</i> ₁₀	21	9 270	5 843.91	0.5214670	0.8671655	3.0750769	63.198195 73.128949	3 560	2 244.26	0.5019662	0.3330215	1.13155775	0.7225253
<i>P-1152-RS3</i> <i>BX2*</i> <i>G</i> ₃	22	28 751	10 048.45	0.5558067	2.6895229	3.2824855	64.618278 75.162474	7 603	2 657.24	0.5070366	0.7112254	0.2334350	0.7908620
<i>P-1152-RS3</i> <i>BX2*</i> <i>G</i> ₅	23	99 917	10 324.95	0.5849254	9.3467726	2.9894773	67.397809 72.878740	24 586	2 540.60	0.5097616	2.2999064	-0.029205	0.8025268
<i>P1152-RS3</i> <i>BX2*</i> <i>G</i> ₁₀	24	6 950	6 806.72	0.5797122	0.6501403	2.7979603	63.167089 70.284149	2 392	2 342.69	0.5041806	0.2237605	0.0550304	0.7439520
<i>LK-P1A</i> <i>BX2*</i> /cy <i>G</i> ₃	25	299	88.38	0.5384615	0.0279701	0.8333333	65.217391 57.971014	196	57.93	0.5102041	0.0183349	0.0348367	0.6040404
<i>LK-P1A</i> <i>BX2*</i> /cy <i>G</i> ₅	26	171	84.99	0.4269006	0.0159963	0.5000000	60.275972 66.326536	97	48.21	0.5154639	0.0090739	0.1818182	0.6380597

Table S2 - Hermant *et al.*

42AB

Line	Library ID	Total piRNAs (24-29nt)	Normalized total piRNAs (24-29nt)	Forward piRNA ratio	piRNA density	z-score piRNA	Ufreq (forward reverse)	Total siRNAs (20-21nt)	Normalized total siRNAs (20-21nt)	Forward siRNA ratio	siRNA density	z-score siRNA	piRNA/siRNA ratio
<i>BX2^{naive}</i>	1	251 380	60 947,01	0.5277746	1.0055159	3.2609650	75.4597805 72.2402870	18 528	4 492,11	0.5104706	0.0741117	2.0277587	0.9313543
<i>BX2* G₈₃</i>	2	358 209	107 235,02	0.5983629	1.4328302	1.9976951	79.6765870 70.4441509	7 648	2 289,54	0.5338650	0.0305918	1.4049114	0.9790956
<i>CyRo1/+</i>	3	209 599	69 737,67	0.5671163	0.8383926	2.8232125	75.5053867 66.8970153	4 923	1 637,98	0.4834451	0.0196919	1.1345525	0.9770513
<i>BX2*, aub -/+</i>	4	65 515	60 885,89	0.4795695	0.2620589	1.7100998	70.5146567 67.4067339	5 607	5 210,82	0.4920634	0.0224279	10.1894181	0.9211636
<i>BX2*, aub -/-</i>	5	44 209	44 209,00	0.4756723	0.1768329	0.4341606	67.3593688 67.7653149	4 908	4 908,00	0.4680114	0.0196319	0.6119080	0.9000753
<i>BX2*, Dcr-2 -/+</i>	6	326 304	80 197,63	0.5554053	1.3052107	3.3756072	75.8657794 71.193473	8 996	2 211,00	0.4869942	0.0359838	2.7157112	0.9731702
<i>BX2*, Dcr-2 -/-</i>	7	309 077	98 383,57	0.5392766	1.2363030	3.5546984	76.1528216 72.2020519	5 527	1 759,32	0.4933960	0.0221079	3.6898731	0.9824318
<i>BX2*, zuc -/+</i>	8	293 082	92 882,96	0.5303430	1.1723233	3.1127857	76.1641397 68.5829071	9 582	3 036,71	0.4862241	0.0383278	1.8364065	0.9683411
<i>BX2*, zuc -/-</i>	9	14 717	9 153,82	0.5259224	0.0588677	4.7124807	55.2842377 70.7162764	23 954	14 899,13	0.4897303	0.0958156	1.5687799	0.3805694
<i>BX2*, cuff -/+</i>	10	205 849	72 619,29	0.5229075	0.8233927	2.5202851	70.7162764 67.6648779	5 103	1 800,23	0.4852047	0.0204119	0.7719812	0.9758096
<i>BX2*, cuff -/-</i>	11	11 510	5 267,76	0.8339704	0.0460398	1.6992683	71.0027591 66.7189962	945	432,50	0.5470899	0.0037799	2.7339671	0.9241268
<i>BX2*, rhl -/+</i>	12	255 956	70 314,31	0.57845489	1.0238199	1.6707586	77.1984885 67.4783895	5 817	1 598,00	0.5081657	0.0232679	0.7568642	0.9777784
<i>BX2*, rhl -/-</i>	13	8 502	4 766,58	0.7215949	0.0340078	2.8922345	75.2241238 62.8221377	848	475,42	0.3030660	0.0033919	2.3237900	0.9093048
<i>BX2^{Dcr2-/-} BX2* G₁</i>	14	1 069 586	306 623,25	0.4755036	4.2783268	4.2313534	74.4431885 74.1922285	29 905	8 573,01	0.4699214	0.1196195	4.2567205	0.9728010
<i>BX2^{Dcr2-/-} BX2* G₂</i>	15	510 105	153 234,06	0.5166857	2.0404118	4.5187635	74.1922285 70.7521264	19 209	5 770,33	0.5342808	0.0768356	3.9249952	0.9637096
<i>P-1152 BX2* G₃</i>	16	111 163	58 378,13	0.5554905	0.4445022	2.8138294	75.9271255 74.5300001	3 938	2 068,07	0.4855256	0.0157519	1.5213884	0.9657865
<i>P-1152 BX2* G₅</i>	17	118 411	63 471,72	0.5395022	0.4736421	3.1491431	65.9514377 66.6383340	2 918	1 564,13	0.4746401	0.0116719	1.8368471	0.9759496
<i>P-1152 BX2* G₁₀</i>	18	130 333	51 890,64	0.5550551	0.5213299	3.5805824	74.8527826 67.4294287	7 994	3 182,72	0.5056292	0.0319758	1.3433085	0.9422094
<i>RS3 BX2* G₃</i>	19	206 259	62 642,69	0.4883423	0.8250326	2.2707430	72.0675105 70.1404286	6 483	1 968,94	0.4872744	0.0259318	0.6003600	0.9695264
<i>RS3 BX2* G₅</i>	20	414 879	119 720,93	0.4964266	1.6595093	4.6203198	73.9945716 72.3768148	16 407	4 734,54	0.5242884	0.0656277	4.0592789	0.9619579
<i>RS3 BX2* G₁₀</i>	21	75 715	47 731,61	0.55228158	0.3028587	2.8301637	75.0023914 64.9889675	3 968	2 501,47	0.4931955	0.0158719	1.2626127	0.9502026
<i>P-1152+RS3 BX2* G₃</i>	22	307 821	107 583,22	0.5247172	1.2312790	4.5609108	72.9838466 69.7628140	12 636	4 416,27	0.5260367	0.0505437	3.8346680	0.9605688
<i>P-1152+RS3 BX2* G₅</i>	23	659 798	68 180,42	0.5406351	2.6391814	2.1888619	74.5646445 69.7895709	22 409	2 315,64	0.4900709	0.0896356	0.3950173	0.9671521
<i>P-1152+RS3 BX2* G₁₀</i>	24	58 566 1	57 358,60	0.54359184	0.2342630	2.6814524	75.0949520 66.8237934	2 515	2 463,15	0.5121272	0.0100599	1.2869472	0.9588251
<i>LK-P1A BX2*/Cy G₃</i>	25	184 436	54 514,96	0.5179087	0.7377410	2.5125870	73.1388809 69.1143226	6 516	1 925,98	0.4771332	0.0260639	1.3844381	0.9658762
<i>LK-P1A BX2*/Cy G₅</i>	26	98 781	49 097,86	0.5140867	0.3951224	2.7303362	73.4492361 68.4035984	3 776	1 876,81	0.4759004	0.0151039	1.1558218	0.9631815

Table S2 - continued - Hermant *et al.*

Dans cet article, nous avons montré que le locus *BX2** requiert les deux voies, primaires et secondaires, de la machinerie piRNA. En effet, les capacités de répression et de production des piRNAs du locus *BX2** sont dépendantes des protéines Rhino et Cutoff, impliquées dans la production des précurseurs de piRNAs, de Zucchini, impliquée dans la biogenèse des piRNAs primaires, et de Aub, impliquée dans la biogenèse des piRNAs primaires et secondaires. Le locus *BX2** produit également des siRNAs dépendant de Dicer-2, impliquée dans la biogenèse des siRNAs. Ces siRNAs ne sont nécessaires ni à la répression, ni à la mise en place de la paramutation (de Vanssay *et al.* 2012 ; Hermant *et al.* en préparation). Nous avons constaté la production d'une très faible quantité de siRNAs indépendante de Dicer-2 et présentant une certaine signature Ping-Pong avec les piRNAs. De plus, nous avons montré qu'il est possible d'obtenir une activation *de novo* du locus *BX2* par transmission maternelle de piRNAs homologues uniquement sur une partie du transgène. L'analyse des petits ARNs ovariens de ce locus montre qu'il produit deux types de petits ARNs : piRNAs et siRNAs, tout le long de sa séquence alors qu'il n'en a reçu seulement que sur une partie. Par ailleurs, les piRNAs du locus *BX2**, produits par les régions qui n'ont pas reçu de piRNAs maternels, sont fonctionnels car ils sont capables de réprimer un transgène cible (*faf-lacZ*). Enfin, nous avons montré qu'il y a un impact somatique d'un héritage maternel des piRNAs du locus *T-1* sur deux systèmes « read-out », le cluster de transgènes *T-1* lui-même et le cluster de transgène *DXI* dans les yeux adultes.

Donc, nous avons activé *de novo* un locus *BX2* par transmission de piRNAs maternels pour la production de piRNAs primaires et secondaires chez le zygote. Cela signifie qu'il est possible qu'à chaque génération, la transmission maternelle de piRNAs active l'ensemble de clusters de piRNAs, ainsi que de nouveaux loci composés de séquences répétées et transcrits dans les deux sens.

VIII/ Emergence d'un locus producteur de piRNAs par un phénomène environnemental

1/ Activation du locus *BX2* par la chaleur

La paramutation liée aux piRNAs chez la drosophile correspond à un transfert entre loci homologues de la capacité à produire de piRNAs. Ceci peut avoir une implication importante dans l'augmentation du nombre de loci producteurs de piRNAs au sein d'un génome. Mais comment le premier locus producteur de piRNAs a-t-il pu démarrer ? A-t-il pu par exemple démarrer par un stress?

Des lignées *BX2^{naïf}* et *T-1^{naïf}* ont été utilisées pour répondre à cette question. Il a été testé l'effet de la température. Il a été montré en effet dans le passé que la chaleur augmente les capacités répressives des éléments *P* télomériques insérés dans des loci producteurs de piRNAs (Marin *et al.* 2000). L'hypothèse est que la chaleur pourrait induire un stress qui augmenterait la production de piRNAs.

Stéphane et Valérie Delmarre au laboratoire ont donc entrepris la construction de souches portant un locus *BX2^{naïf}* avec une cible (*BX2*, *P1039/Cy* et *BX2/Cy*; *BQ16*, *nosgal4/Sb*) et un locus *T-1^{naïf}* avec une cible (*T-1*, *P1039/Cy*). Ces lignées ne présentent pas de capacités répressives. Il est en effet possible de générer une souche *T-1^{naïf}* car le passage par voie paternelle du locus *T-1* lui fait perdre de façon stable ses capacités répressives. Toutes ces lignées ont été passées successivement de 25°C (pour la ponte) à 29°C (pour le développement) pendant plusieurs mois et ont été testées en coloration *LacZ*. La plupart des lignées sont devenues répressives (10 lignées sur 14). Il a donc été possible d'activer *de novo* par la chaleur des clusters *BX2^{naïf}* et *T-1^{naïf}*. Par contre, les mêmes lignées toujours maintenues à 25°C n'ont montré aucune activité répressive sur cette même période. Les lignées activées ont été, ensuite, replacées à 25°C et j'ai repris leur analyse en établissant 5 sous-lignées de chaque condition.

La question biologique est : est-ce que ces clusters que nous appellerons *BX2^θ* et *T-1^θ*, activés *de novo* par une température de 29°C et ramenés ensuite à 25°C, sont semblables du point de vue fonctionnel et moléculaire aux souches *BX2** et *T-1* ? Cela implique en premier lieu leur stabilité au cours des générations après leur retour à 25°C.

2/ Propriétés fonctionnelles et moléculaires d'un locus activé par la chaleur.

Dans un premier temps, j'ai donc testé la stabilité des lignées $BX2^{\theta}$ et $T-I^{\theta}$ au cours du temps en effectuant régulièrement des tests TSE (**Figure 26 et 27**). Jusqu'en G_{15} , une seule lignée (sur 15) s'est montrée instable. Cela signifie que toutes les autres lignées $BX2^{\theta}$ et $T-I^{\theta}$ montrent des capacités répressives aussi stable qu'une souche $BX2^*$ paramutée par héritage maternel de piRNAs ($T-I^{\theta} BX2^*$). De plus, ces souches présentent un effet maternel car elle ne régule un transgène cible que lorsque les clusters $BX2^{\theta}$ et $T-I^{\theta}$ sont transmis maternellement (**Figure 28**).

	$BX2^{\theta}-1$ ($BX2, P-1039$)					$BX2^{\theta}-2$ ($BX2; BQ16$)				
	$BX2^{\theta}-1A$	$BX2^{\theta}-1B$	$BX2^{\theta}-1C$	$BX2^{\theta}-1D$	$BX2^{\theta}-1E$	$BX2^{\theta}-2F$	$BX2^{\theta}-2G$	$BX2^{\theta}-2H$	$BX2^{\theta}-2I$	$BX2^{\theta}-2J$
G_1	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
G_2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
G_3	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
G_4	100%	100%	87,5%	100%	100%	100%	100%	100%	100%	100%
G_5	100%	100%	66.66%	100%	100%	100%	100%	100%	100%	100%
G_6	100%	100%	83.33%	100%	100%	100%	100%	100%	100%	100%
G_{10}	100%	100%	83.33%	100%	100%	100%	100%	100%	100%	100%
G_{15}	100%	100%	71.42%	100%	100%	100%	100%	100%	100%	100%

Figure 26 : Capacités répressives du cluster $BX2^{\theta}$ avec deux transgènes cibles différents : $P-1039$ et $BQ16$ au cours du temps (en pourcentage de répression).

5 sous-lignées $BX2^{\theta}$ portant un transgène cible $P-1039$ ($BX2^{\theta}-1A-E$) ont été testées ainsi que 5 sous-lignées $BX2^{\theta}$ portant un transgène cible $BQ16$ ($BX2^{\theta}-2F-J$). Le pourcentage de répression est calculé sur un effectif minimum $n=300$. 9 sous-lignées sur 10 présentent des capacités répressives fortes et stables. Une seule sous-lignée présente des capacités répressives instables.

	$T-1^{\theta}-1$ ($T-1, P-1039$)				
	$T-1^{\theta}-1E$	$T-1^{\theta}-1F$	$T-1^{\theta}-1H$	$T-1^{\theta}-1I$	$T-1^{\theta}-1J$
G_1	100%	100%	100%	100%	100%
G_2	100%	100%	100%	100%	100%
G_3	100%	100%	100%	100%	100%
G_4	100%	100%	100%	100%	100%
G_5	100%	100%	100%	100%	100%
G_6	100%	100%	100%	100%	100%
G_{10}	100%	100%	100%	100%	100%
G_{15}	100%	100%	100%	100%	100%

Figure 27 : Capacités répressives du cluster $T-1^{\theta}$ avec un transgène cible : $P-1039$ au cours du temps (en pourcentage de répression).

5 sous-lignées $T-1^{\theta}$ portant un transgène cible $P-1039$ ($T-1^{\theta}-1E$, F , H , I et J) ont été testées. Le pourcentage de répression est calculé sur un effectif minimum $n=300$. Toutes ces sous-lignées présentent des capacités répressives fortes et stables.

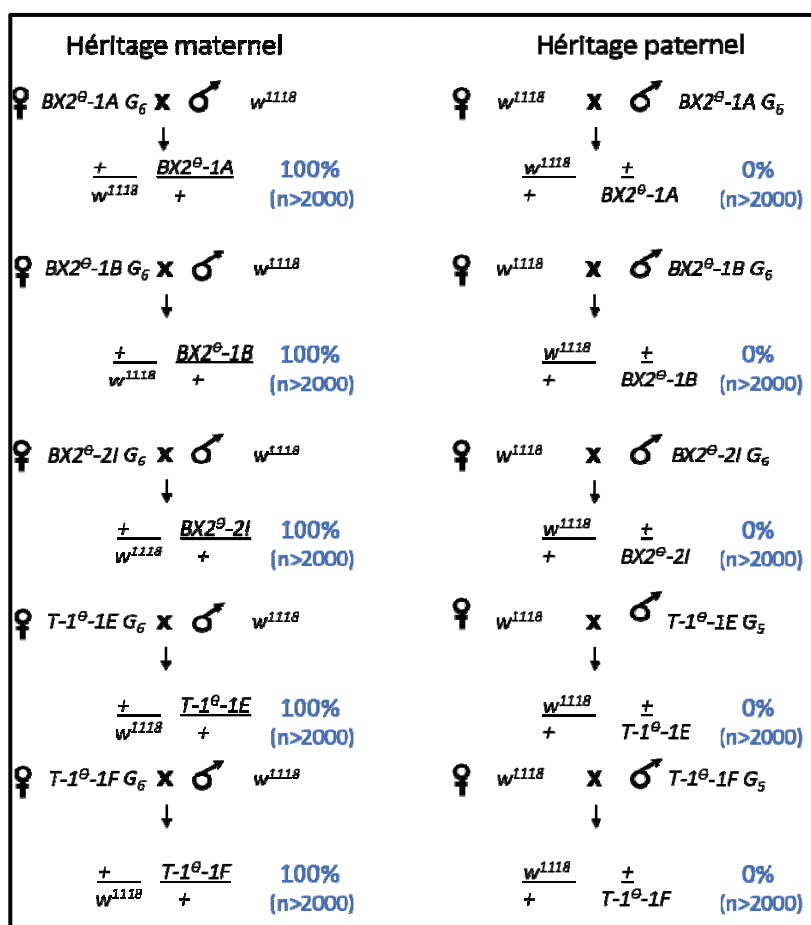


Figure 28 : Effet maternel de la répression des $BX2^{\theta}$ et $T-1^{\theta}$.

Croisement des souches portant les clusters $BX2^{\theta}$ ou $T-1^{\theta}$ et un transgène cible ($P-1039$ pour les sous-lignées $BX2^{\theta}-1A$ et $1B$ et $T-1^{\theta}-1E$ et $1F$ ou $BQ16$ pour la sous-lignée $BX2^{\theta}-2I$) avec une souche $M(w^{1118})$ dépourvue de transgène dans les deux sens de croisements. (A droite) Héritage maternel. (A gauche) Héritage paternel. En dessous de chaque croisement, génotype testé et pourcentage de répression du transgène cible $P-1039$ ou $BQ16$. Pour tous portant les clusters $BX2^{\theta}$ et $T-1^{\theta}$ testés un effet maternel est observé. En effet, lorsque des femelles portant les clusters $BX2^{\theta}$ ou $T-1^{\theta}$ et un transgène cible sont croisées par mâles w^{1118} , il y a

une forte répression du transgène cible dans la descendance dû à l'héritage maternel des clusters $BX2^{\theta}$ et $T-1^{\theta}$. Par contre, lorsque des femelles w^{1118} sont croisées par mâles portant le cluster $BX2^{\theta}$ ou $T-1^{\theta}$ et un transgène cible, il n'y a pas de répression du transgène cible dans la descendance du fait de l'absence d'héritage maternel de cytoplasme contenant des piRNAs.

Ensuite, j'ai analysé les petits ARNs ovariens d'une lignée $BX2^{\theta}$ et $T-1^{\theta}$ (**Figure 29 et 30**). Nous pouvons constater que les clusters $BX2^{\theta}$ et $T-1^{\theta}$ produisent de piRNAs et siRNAs de façon comparable aux clusters $BX2^*$ et $T-1$, respectivement. A noter que $T-1$ et $T-1\theta$ produisent plus de piRNAs homologues au transgène $P\{lacW\}$ sens et antisens que $BX2^*$ et $BX2^{\theta}$ (**Figure 29**), contrairement à la production de piRNAs issus de 42AB (**Figure 30**). Le fait de ne pas observer d'augmentation de la quantité de piRNAs produits par le 42AB est surprenant. En effet, nous aurions pu avoir une augmentation de la quantité de piRNAs issus du locus 42AB en réponse à l'effet de la chaleur ce qui n'est pas le cas. Il est possible que ce soit le démarrage qui soit sensible à la chaleur et non le niveau d'équilibre de production.

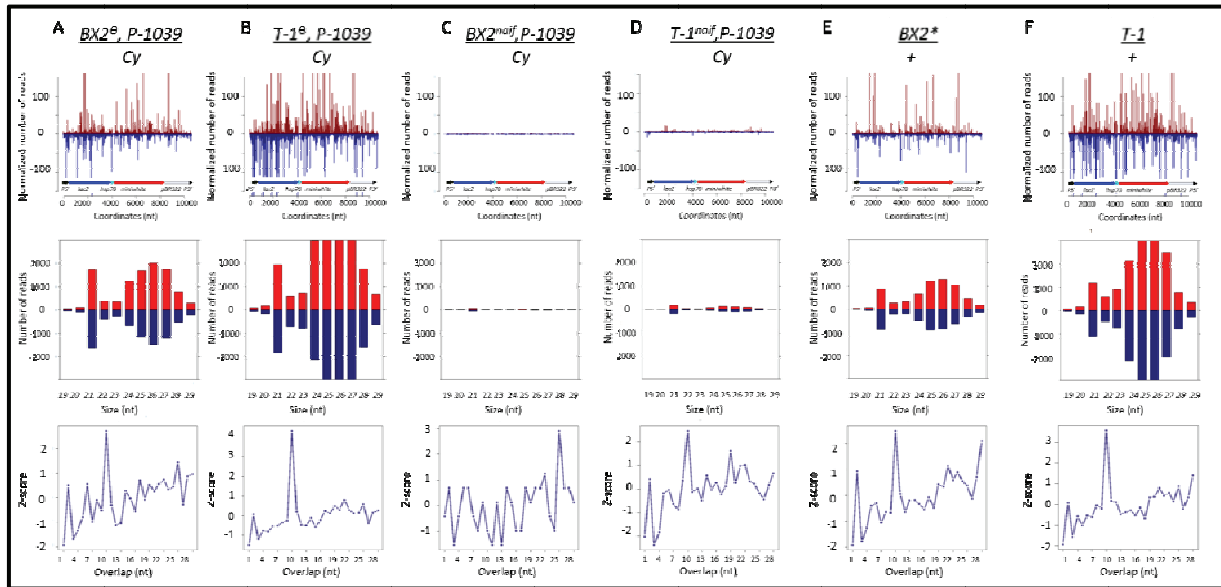


Figure 29 : Analyse des petits ARNs ovariens homologues au transgène $P\{lacW\}$ des souches $BX2^{\theta}$, $P-1039$ ($BX2^{\theta}-1B$ en G_5).

(A), $T-1^{\theta}$, $P-1039$ ($T-1^{\theta}-1F$ en G_5) (B) comparés aux souches témoins : $BX2^{naif}$, $P-1039$ (C), $T-1^{naif}$, $P-1039$ (D) $T-1$ $BX2^*$ en G_{83} (E) et $T-1$ (F). Les petits ARNs ovariens sens sont en rouge et les petits ARNs ovariens antisens sont en bleu. (**Graphique du haut**) Les petits ARNs ovariens de 19 à 29nt sont alignés en fonction de leurs coordonnées (en nucléotides) sur la séquence $P\{lacW\}$. (**Graphique du milieu**) Les petits ARNs ovariens de 19 à 29 nt sont analysés en fonction de leur taille (en nucléotides). (**Graphique du bas**) Signature Ping-Pong des ARNs de 23 à 28nt. Les analyses sont normalisées sur la même banque de référence que les analyses de l'article Hermant *et al.* en préparation (**Figure Table S1 et Tableau 2**) (A) Le cluster $BX2^{\theta}$ activé *de novo* par la chaleur montre un profil des petits ARNs ovariens homologue au transgène $P\{lacW\}$ comparable au cluster $BX2^*$ activé *de novo* par les piRNAs du cluster $T-1$ en (E). (B) le cluster $T-1\theta$ activé *de novo* par la chaleur montre en profil de petits ARNs ovariens homologue au transgène $P\{lacW\}$ comparable au cluster $T-1$ en (F). Le

cluster *BX2^{naif}*, *P-1039* (C) ne produit presque pas piRNAs alors que le cluster *T-1^{naif}*, *P-1039* (D) en produit un peu.

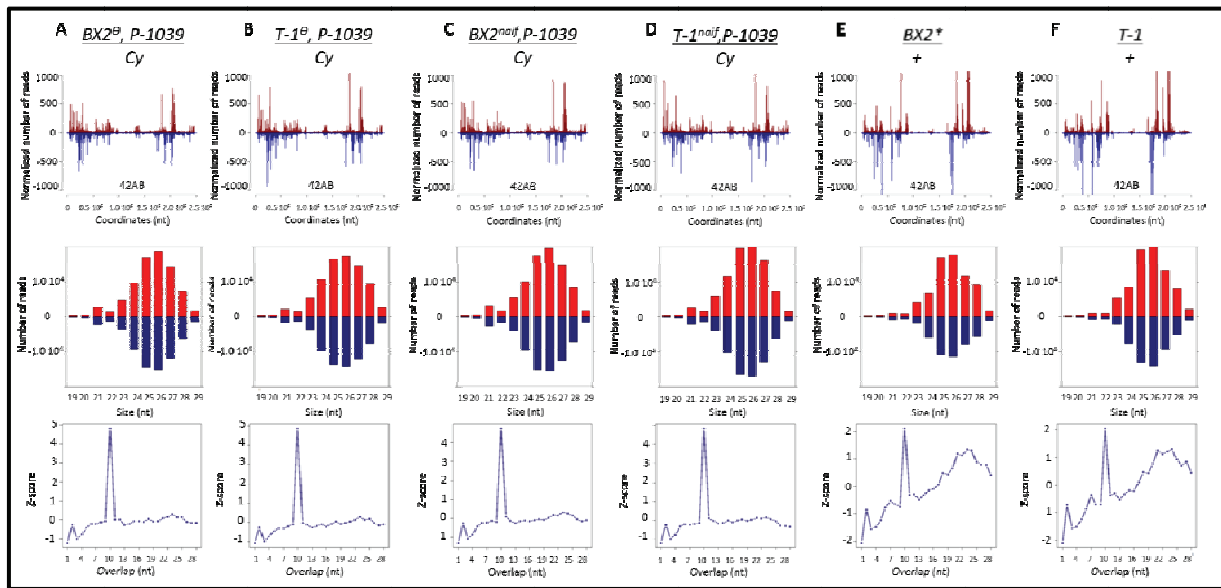


Figure 30 : Analyse des petits ARNs ovariens homologues au locus 42AB des souches *BX2^θ*, *P-1039* (*BX2^θ-1B* en G₅).

(A), *T-1^θ*, *P-1039* (*T-1^θ-1F* en G₅) (B) comparés aux souches témoins : *BX2^{naif}*, *P-1039* (C), *T-1^{naif}*, *P-1039* (D), *T-1^{naif}BX2** en G₈₃ (E) et *T-1* (F). Les petits ARNs ovariens sens sont en rouge et les petits ARNs ovariens antisens sont en bleu. (**Graphique du haut**) Les petits ARNs ovariens de 19 à 29nt sont alignés en fonction de leurs coordonnées (en nucléotides) sur la séquence 42AB. (**Graphique du milieu**) Les petits ARNs ovariens homologues au locus 42AB de 19 à 29 nt sont analysés en fonction de leur taille (en nucléotide). (**Graphique du bas**) Signature Ping-Pong des ARNs de 23 à 28nt. Les analyses sont normalisées sur la même banque de référence que les analyses de l'article Hermant *et al.* en préparation (**Figure Table S1 et Tableau 2**) (A) Le cluster *BX2^θ* activé *de novo* par la chaleur montre un profil des petits ARNs ovariens homologue au locus 42AB comparable au cluster *BX2** activé *de novo* par les piRNAs du cluster *T-1* en (E). (B) Le cluster *T-1^θ* activé *de novo* par la chaleur montre en profil de petits ARNs ovariens homologue au locus 42AB comparable au cluster *T-1* en (F).

Library ID	29	30	31	32	33
Parental origin	<i>BX2^θ, P-1039/Cy</i>	<i>T-1^θ, P-1039/Cy</i>	<i>BX2^{maff}, P-1039/Cy</i>	<i>T-1^{maff}, P-1039/Cy</i>	<i>T-1/Cy x w¹¹¹⁸</i>
female genotype	<i>BX2^θ, P-1039/Cy</i>	<i>T-1^θ, P-1039/Cy</i>	<i>BX2^{maff}, P-1039/Cy</i>	<i>T-1^{maff}, P-1039/Cy</i>	<i>T-1/+</i>
Dmel matched	18 552 704	21 717 559	16 928 361	29 642 142	13 778 234
miRNAs	21.30 %	19.23 %	21.39 %	21.01 %	21.30 %
tRNAs	2.01 %	1.28 %	2.67 %	2.46 %	0.74 %
ncRNAs	1.25 %	1.03 %	1.01 %	1.07 %	0.51 %
mlscRNAs	14.37 %	12.60 %	11.28 %	8.52 %	5.10 %
Transposons	43.91 %	48.32 %	45.59 %	46.80 %	55.43 %
Introns	5.61 %	5.70 %	5.66 %	6.79 %	4.84 %
Transcripts	3.12 %	2.79 %	3.45 %	2.96 %	2.31 %
Intergenic	7.47 %	8.07 %	8.04 %	9.45 %	8.86 %
Remaining	0.96 %	0.97 %	0.90 %	0.94 %	0.90 %
Normalisation factor	0.2807205	0.2328538	0.2989408	0.1650400	0.3357009

Tableau 2 : Annotation et normalisation complémentaire à la Table S1 Hermant *et al.* en préparation.

Les petits ARNs sont préparés à partir d'ovaires de femelles dont le génotype est indiqué « *female genotype* ». L'origine parentale de ces femelles est indiquée « *Parental origin* ». La souche *w¹¹¹⁸* est dépourvue de transgène. Les valeurs de différentes catégories de séquences sont exprimées en pourcentage de nombre total de séquences qui s'alignent sur le génome de *D. melanogaster* 5.49 pour chaque librairie (*Dm Matched*). Les facteurs de normalisation utilisés pour comparer les librairies sont indiqués (*Normalisation factor*). Les librairies sont préparées en utilisant la technologie de séquençage Illumina. La librairie de référence est la librairie n°5 de la Table S1 Hermant *et al.* en préparation.

Une autre différence observée entre *T-1^θ* versus *BX2^θ*, est que *T-1^θ* a des capacités répressives dans le cas de la répression de l'élément *P* (m=67.0% de stérilité dysgénésique ; s=20.4% ; n=15) tout comme le locus *T-1* (m=35,7% de stérilité dysgénésique ; s=6.1% ; n=3) ce qui n'est pas le cas de *BX2** (m=100% de stérilité dysgénésique ; s=0% ; n=3) et *BX2^θ* (m=98.6% de stérilité dysgénésique ; s=2.5% ; n=25). La raison de cette différence n'est pas claire. L'une des hypothèses est que la quantité supérieure de piRNAs produit par *T-1* permettrait la régulation de nombreuses copies actives de l'élément *P*. Cela signifierait qu'il puisse exister un seuil de quantité de piRNAs nécessaire pour réguler les cibles qui serait dépendants du nombre de cibles à réprimer. En effet, dans le cas de la régulation des transgènes *P-1039* et *BQ16*, il n'y a qu'une seule copie du transgène cible à réprimer, alors que dans le cas de la régulation de l'élément, il y a 60 à 80 copies d'éléments *P* dont environ 1/3 correspond à des éléments complets.

L'effet d'activation par la chaleur de la production de piRNAs par un locus « dormant » a été obtenu au cours de ce travail en présence d'un transgène cible. Ce phénomène va être testé maintenant en absence de transgène cible par un nouvel étudiant de Master 2, Georges Ibrahim. Il s'agira de voir si le cluster a une capacité autonome à démarrer ou s'il requiert la présence d'une cible qui pourrait contribuer à la production de piRNAs.

Discussion

Discussion

Au cours de ma thèse, j'ai travaillé sur deux phénomènes d'activation *de novo* d'un locus producteur de piRNAs : activation par les piRNAs transmis maternellement et par la chaleur. Dans le premier cas, nous avons vu qu'il est possible d'activer de façon stable *de novo*, via une transmission maternelle de piRNAs, la production de piRNAs par un locus *BX2* qui est « dormant » pour cette production. Il s'agit du premier cas de paramutation stable chez l'animal. En effet, ce cluster *BX2* activé *de novo* par des piRNAs homologues présente toutes les propriétés d'un locus paramuté : 1/ il a acquis, de façon autonome, la capacité à produire des piRNAs et ce de façon stable, 2/ il est devenu paramutagénique. Le locus *BX2** est capable de produire des piRNAs primaires et secondaires, sens et antisens comme le locus qui l'a paramuté ou le locus producteur de piRNAs de référence 42AB. Il produit également des siRNAs qui, de façon surprenante, sont dépendants de la voie des piRNAs. Nous avons de plus montré que le locus *BX2* peut être paramuté, de façon relativement stable, lors d'une paramutation partiellement homologue, même si toutefois, cette dernière montre une certaine instabilité que ne présente pas le locus *BX2** lorsque la paramutation est complètement homologue. Enfin, nous avons observé un effet somatique dans l'œil des piRNAs hérités maternellement. Dans un second temps, nous avons montré qu'il est possible d'activer *de novo* par la chaleur un locus producteur de piRNAs qui présentent l'ensemble des propriétés fonctionnelles et moléculaires comme un locus activé *de novo* par les piRNAs transmis maternellement.

Le locus *BX2, un cluster bidirectionnel de piRNAs :**

L'analyse des petits ARNs ovariens montre que ce locus produit des piRNAs sens et antisens tout comme le locus 42AB. La localisation et l'analyse transcriptionnelle de *BX2** par RT Q-PCR montre que ce cluster est un cluster bidirectionnel. En effet, il est inséré dans le premier intron du gène *AGO1* en antisens. Nous avons vérifié, par RT Q-PCR, que cette insertion n'altère pas la transcription de ce gène impliqué dans la voie des miRNAs. De façon surprenante, il n'y a pas de différence transcriptionnelle entre *BX2^{naïf}* et *BX2** qui sont tous deux transcrits faiblement. Il est cependant possible que la différence entre les deux loci, activé ou non, soit chromatienne. L'une des hypothèses est qu'il y aurait un recrutement du complexe RDC au niveau du locus activé par les piRNAs transmis maternellement. Ce recrutement permettrait l'envoi des transcrits de *BX2** vers la voie des piRNAs. De plus, il a

été montré que les clusters bidirectionnels pouvaient ou non être dépendant de Rhino et donc du complexe RDC (Mohn *et al.* 2014). Le fait que le locus *BX2* soit sensible à Rhino et Cutoff tend à montrer qu'il est sensible au complexe RDC. Il nous faut montrer maintenant la présence par ChIP Q-PCR des protéines Rhino et Cutoff sur le locus *BX2** et tester si elles sont absentes du locus *BX2^{naïf}*. J'ai de plus essayé de le tester par ChIP Q-PCR mais la mise au point de l'expérience n'a pas abouti pour le moment. J'ai essayé de tester la colocalisation de *BX2* et Rhino par DNA Fish. Il est possible de visualiser le cluster *BX2** dans les noyaux de cellules sous forme de foci similaires à ce qui est observé pour le locus 42AB (Mohn et al 2014). Rechercher une colocalisation avec Rhino et Cutoff constitue un de nos projets. Une expérience indirecte, menée par Adrien Le Thomas, montre la présence de Rhino et Cutoff, par ChIP Q-PCR, sur le locus *T-1* lorsque celui-ci est hérité maternellement (Le Thomas *et al.* 2014). La détection de la présence de Rhino et Cutoff doit être testée maintenant sur le locus *BX2** pour comprendre le mécanisme même de la paramutation. De plus, il a été montré qu'il existe une différence chromatienne entre les loci *BX2^{naïf}* et *BX2**, puisqu'il y a plus de marques H3K9me3 lorsque le locus *BX2* est activé (Le Thomas *et al.* 2014). Toutefois, l'expérience essentielle reste à être réalisée sur un locus *BX2** après au moins deux générations après paramutation pour tester quel est l'état chromatinien d'un locus paramuté à l'équilibre. Il a été montré que le locus *BX2* forme un îlot hétérochromatique en contexte euchromatique, d'après des expériences d'hybridation sur chromosomes polytènes avec la protéine HP1 (Dorer and Henikoff 1994, 1997). Il est possible que le locus non activé soit enrichi en HP1 mais avec peu de Rhino ne permettant pas la production de piRNAs. Par contre, le locus activé serait enrichi en Rhino, mais avec peu de HP1, permettant la production de piRNAs. Il y aurait donc une balance entre Rhino et HP1 pour favoriser la production ou non de piRNAs. Néanmoins, cela ne peut être la seule explication car il existe des clusters de piRNAs germinaux indépendant de Rhino et riche en HP1 qui sont capable de produire des piRNAs, qui représentent 15% des loci sources de piRNAs (Mohn *et al.* 2014). Ces clusters indépendant de Rhino (appelé RI-SL pour « Rhino-independent Source loci ») peuvent être unidirectionnels et bidirectionnels (Mohn *et al.* 2014). Il existe, donc, un paramètre qui permet la production de piRNAs. La grande question est quelle est la nature de ce paramètre ?

Il a été observé que les protéines Rhino et Cutoff n'avaient pas le même effet sur la biogenèse des piRNAs selon les clusters. En effet, il est observé une baisse des piRNAs produits par l'ensemble des clusters de piRNAs en contexte mutant pour *cutoff* alors que certains clusters ne sont pas dépendant de Rhino (Klatenhof *et al.* 2009 ; Pane *et al.* 2011 ;

Mohn *et al.* 2014). L'une des hypothèses concernant la fonction de Cutoff est que Cutoff pourrait se fixer à l'extrémité 5' des transcrits à la place du complexe de coiffe (complexe CBC « *cap-binding complexe* »). Cutoff permettrait ainsi d'identifier ces transcrits en tant que précurseurs de piRNAs. Etant donné que la protéine Cutoff présente un effet général sur tous les clusters, il est possible que ce soit cette protéine qui rende un cluster producteur de piRNAs et soit donc le paramètre recherché qui spécifie un locus producteur de piRNAs. Afin de répondre à cette question, je propose de tester par CHIP seq la présence de Cutoff sur tous les clusters de piRNAs. En effet, si c'est Rhino qui permet le recrutement de Cutoff sur les clusters de piRNAs, il est possible que la protéine HP1 fasse de même pour les clusters indépendants de Rhino.

Le fait que ce locus *BX2* soit bidirectionnel et riche en H3K9me3, et donc hétérochromatique, peut être à l'origine du fait que l'activation *de novo* par les piRNAs ait pu se faire. En effet, les clusters producteurs de piRNAs sont localisés dans des régions hétérochromatiques favorables à la production des piRNAs. De plus, mis à part le cluster 2, tous les clusters germinaux sont bidirectionnels. De plus, ce locus contient des séquences répétées d'ETs tout comme les clusters de piRNAs. Ce locus *BX2* aurait donc toutes les potentialités transcriptionnelles et génétiques requises pour une production de piRNAs pourvue que celle-ci soit initiée. Il serait possible de tester si le fait d'être transcrit dans les deux sens et riche en H3K9me3 est suffisant pour permettre d'activer un cluster producteur de piRNAs. Il a été montré par l'utilisation de transgène *lacI ::Rhino* et *lacO :: GFP*, qu'il est possible de forcer le recrutement de Rhino sur le transgènes *lacO ::GFP* qui produit alors des piRNAs uniquement lorsque ce transgène est transcrit dans les deux sens (Zhang *et al.* 2014). En utilisant ce même système, il serait possible d'hypermétyler un transgène *lacO ::GFP* par un transgène *lacI : histone méthyle transférase (Su(var)3-9, G9a)*, afin de voir s'il est capable de recruter la protéine Rhino et de produire des piRNAs lorsqu'il est transcrit dans les deux sens.

La production de piRNAs dépend d'un phénomène *trans*-générationnel maternel:

Afin d'expliquer la paramutation par transmission maternelle de piRNAs, nous avons proposé un modèle dans lequel le dépôt maternel de piRNAs dans l'embryon pourrait agir sur les cellules de la future lignée germinale (de Vanssay *et al.* 2013). En effet, il a été montré qu'il y a un dépôt de piRNAs maternels dans les embryons, ainsi qu'une forte accumulation de protéines Piwi et Aub au pôle postérieur des embryons où se trouvent ces futures cellules

germinales (Harris and Madocnald, 2001 ; Megosh *et al.* 2006 et Brennecke *et al.* 2007 et 2008). Nous proposons que les piRNAs chargés par Piwi puissent aller agir au niveau des noyaux de ces cellules afin de recruter les protéines du complexe RDC, entre autres, ce qui permettrait d'identifier les loci homologues aux piRNAs comme clusters de piRNAs. Parallèlement, les piRNAs chargés par Aub, et AGO3 dans une moindre mesure, pourraient aller agir dans le nuage afin d'initier la production de piRNAs secondaires *via* le Ping-Pong. Le fait d'avoir activé *de novo* un locus pour la production de piRNAs *via* une transmission maternelle de piRNAs permet de proposer l'hypothèse que l'activation de tous les clusters de piRNAs pourrait se faire à chaque génération de cette façon. En effet, lors du maintien de la paramutation de *BX2**, à chaque génération, le locus apporté paternellement est probablement « mis à nu » d'un point de vue chromatinien et en quelque sorte reparamuté par les piRNAs homologues transmis maternellement. De plus, il a été montré un rôle *trans*-générationnel de piRNAs au locus *T-1* (Le Thomas *et al.* 2014). Il est ainsi possible d'envisager que, non seulement les piRNAs permettraient le maintien des clusters dans les cellules germinales adultes, mais également qu'ils seraient à chaque génération à l'origine de l'activation des clusters dans les cellules germinales embryonnaires.

Le locus *BX2, un cluster de piRNAs qui produit des siRNAs :**

De nombreux papiers relatent la présence de siRNAs présent dans les ovaires (Klenov *et al.* 2007, Malone *et al.* 2009, Olivieri *et al.* 2012, Donertäs *et al.* 2013), et plus précisément issus des clusters de piRNAs (Czech *et al.* 2008, Pane *et al.* 2011, Preal *et al.* 2012, Muerdter *et al.* 2013, Le Thomas *et al.* 2014). Nos analyses montrent que le locus *BX2** produit des siRNAs dépendants de Dicer-2 qui sont sensibles également au mutant de gènes impliqués dans la voie des piRNAs (Hermant *et al.* en préparation). L'une des hypothèses est que ces siRNAs seraient issus de précurseurs de piRNAs. En effet, le fait d'affecter la production des précurseurs de piRNAs, en contexte mutant pour *rhino* et *cutoff*, affecte la production des siRNAs du locus *BX2*. De plus, le fait d'affecter la production des piRNAs primaires, en contexte mutant pour *zucchini*, affecte également la production de siRNAs du locus *BX2*. L'une des données étonnantes est que ces siRNAs sont dépendants d'Aub qui intervient dans le Ping-Pong. Néanmoins, il a été montré que la mutation d'*aub* affecte la biogenèse générale des piRNAs car sa mutation affecte la localisation des principales protéines intervenant dans la voie des piRNAs, à savoir Piwi, Vasa et AGO3 (Li *et al.* 2009). En particulier, Piwi n'est plus nucléaire ce qui aurait pour

conséquence d'affecter toute la biogenèse des piRNAs puisque Piwi a été montré impliqué être dans le maintien des clusters de piRNAs (Mohn *et al.* 2014). L'une des hypothèses pour expliquer le fait que les siRNAs dépendent de la voie des piRNAs est que lors de la biogenèse des piRNAs, les précurseurs de piRNAs entrent dans le nuage où ils sont clampés par Vasa (Xhiol *et al.* 2014). Ceci permettrait un rapprochement des précurseurs de piRNAs sens et antisens afin de favoriser le Ping-Pong. Ce rapprochement de deux molécules complémentaires générerait par le fait du hasard des précurseurs d'ARNs double brins reconnus par Dicer-2 favorisant ainsi la production de siRNAs. A l'inverse, dans la **Figure S1** (J) Hermant *et al.* en préparation, l'absence de Dicer-2 serait favorable à un meilleur fonctionnement du Ping-Pong car tous les précurseurs de piRNAs s'engageraient dans la voie des piRNAs. Une autre question est : est-ce que les siRNAs sont fonctionnels dans les ovaires ? Nous avons montré qu'ils ne sont pas nécessaires à l'activation du locus *BX2*, ni à la capacité répressive de *BX2*. Néanmoins, nous essayons de tester, par génétique à l'aide d'un transgène que générerait un transcrit *lacZ-hairpin*, s'il est possible d'activer le locus *BX2* que par des siRNAs produit à partir de la séquence *lacZ* et d'étudier les capacités répressives et paramutagéniques d'un tel locus.

Un autre résultat intéressant est qu'en contexte mutant pour *Dicer-2*, il y a encore production d'un peu de siRNAs qui présentent une signature siRNAs-piRNAs aussi bien pour le locus *BX2* que pour le locus 42AB (Hermant *et al.* en préparation). Cela signifie qu'il pourrait y avoir une population d'ARNs de 21nt qui ne serait pas dépendante de Dicer-2 mais issue d'un Ping-Pong atypique. Lors de la biogenèse des piRNAs secondaires, il y aurait production d'ARNs de 21nt. Est-ce que ces ARNs de 21nt sont des produits de dégradation ?

Le locus *BX2* activé *de novo* par des piRNAs présentant une couverture partielle du locus :

Nous avons testé un cas d'activation du locus *BX2* par des piRNAs ne couvrant pas toute la longueur du locus. Une première expérience avait été menée dans laquelle cette activation était possible mais instable (de Vanssay *et al.* 2012). Je me suis alors demandé quelle en est raison. Est-ce dû au fait que le cluster donneur de piRNAs est que partiellement homologue ou est-ce dû au fait que le cluster donneur de piRNAs est localisé à un autre site génomique que le locus *BX2* ? Etonnement, mes résultats ont montré qu'une activation *de novo* du locus *BX2* par des piRNAs présentant une couverture partielle du locus ciblé qui est situé sur un autre chromosome est possible de façon stable (Hermant *et al.* en préparation).

Cela montre qu'une paramutation peut être générée : 1- même si le locus paramutagénique est sur un chromosome différent du locus paramuté, 2- les piRNAs produits par le locus paramutagénique ne couvrent pas la totalité du locus paramuté. Cela confirme que le *pairing* entre les deux épiallèles n'est pas nécessaire contrairement à ce qui est suggère dans certains cas de paramutation chez les plantes (de Vanssay *et al.* 2012). En effet, lors de la paramutation du locus *BX2*, les deux épiallèles ne sont jamais présents dans le même génome. Ce sont les piRNAs qui induisent la paramutation. Il s'agit donc une paramutation par un produit diffusible. Je me suis alors demandé quel pouvait être le profil de production de petits ARNs ovariens de ces locus *BX2** paramutés par des piRNAs présentant une couverture partielle du locus. L'analyse de séquençage a montré que le locus *BX2* activé *de novo* par des piRNAs présentant une couverture partielle du locus est capable de produire de piRNAs tout le long du locus (Hermant *et al.* en préparation). Il y a donc, propagation en *cis* de capacité à produire des piRNAs. L'une des hypothèses est que lors de la transmission maternelle des piRNAs, ces derniers vont activer le locus et permettre que le transcrit dans son ensemble devienne précurseur de piRNAs entrant dans la biogenèse de piRNAs une fois sorti du noyau. Ces transcrits pourraient donc générer des piRNAs sur toute sa longueur démarrant ainsi la production de piRNAs primaires et secondaires. De plus, ces piRNAs issus des séquences du transcrits, qui n'ont pas reçu de piRNAs maternels, sont fonctionnels car ils ont des capacités répressives (Hermant *et al.* en préparation). J'ai également testé une paramutation du locus de *BX2* avec une souche portant un élément *P* complet (*Lk-P(1A)*). L'homologie entre l'élément *P* et le locus *BX2* est limité aux pieds de l'élément *P* qui bornent chaque transgène *P{lacW}*, soit 820pb (sur 10691pb). Un locus *BX2* ne présente pas de capacités de répression et ni de production de piRNAs lorsqu'il est paramuté par *Lk-P(1A)*. Cela signifierait que pour activer un locus, il faudrait qu'il y ait une transmission de piRNAs maternels sur une longueur minimale de la séquence.

Néanmoins, les trois expériences d'activation partielle testées ($P^{-1152}BX2^*$, $RS3BX2^*$, $P^{-1152+RS3}BX2^*$) peuvent être instables contrairement à une activation « totale » toujours stable (Hermant *et al.* en préparation **Figure 4**). Ces résultats sont surprenants maintenant qu'il a été montré qu'il y a propagation en *cis* de la production de piRNAs. En effet, l'hypothèse, qui expliquerait l'activation *de novo* par les piRNAs, est que les piRNAs transmis maternellement pourraient aller dans le noyau des cellules de la future lignée germinale afin de promouvoir le recrutement du complexe RDC au locus (de Vanssay *et al.* 2013). Cette hypothèse est supportée par le fait que le recrutement du complexe RDC aux clusters de piRNAs serait Piwi-dépendant, autrement dit dépendant de la production de piRNAs (Mohn *et al.* 2014). Il

est difficile de comprendre cette instabilité puisque le locus « partiellement » activé produit des piRNAs tout le long du transcrit et donc devrait permettre le recrutement du complexe RDC tout le long du locus à la génération suivante, ce qui devrait conduire à la stabilité de la production des piRNAs. Autrement dit, dès la 3^e génération, les piRNAs produits par $P^{1152}BX2^*$ sont apparemment les mêmes que ceux de $T-1BX2^*$, et pourtant leur devenir dans les générations suivantes est différent (Figure 31). Quel est le support moléculaire de cette différence ? Afin de savoir s'il y a un dépôt du complexe RDC tout le long du locus, il est prévu de tester par CHIP la présence de Rhino et Cutoff sur le locus *BX2*, également, dans ces conditions « d'activation partielle ». Il est possible d'envisager également l'implication d'un paramètre protéique inconnu qui pour des raisons, que je ne comprends pas encore, différencierait des loci activés partiellement ou non. Il pourrait être envisagé que le dépôt maternel de piRNAs permette le dépôt, précocement, d'une marque chromatienne permettant d'activer la production de piRNAs. Mais, dès la 3^e génération, il y a production de piRNAs tout le long du locus, cette activation devrait devenir stable au cours du temps. Or, ce n'est pas le cas. L'autre possibilité est que le dépôt de ce paramètre protéique inconnu tout le long du locus soit lent et responsable de l'instabilité. En effet, la propagation de cette protéine en cours du temps pourrait être dépendante des marques euchromatiques déjà présentes. Le ratio de cette protéine inconnue et des protéines euchromatiques déjà présentes pourrait déterminer cette instabilité. Une autre hypothèse est que la quantité de piRNAs transmise maternellement doive dépasser un seuil pour obtenir une activation stable. Néanmoins, dès la 3^e génération, il y a production de piRNAs de façon comparable à un locus activé *de novo* par le locus *T-1*. De la même façon, dès la 3^e génération, cette activation devrait devenir stable. De plus, le locus *BX2^** est capable d'activer *de novo* un locus *BX2^{naïf}* également de façon stable, et ce jusqu'au 5^e ordre de paramutation (de Vanssay *et al.* 2012). Une analyse moléculaire de ces loci doit être réalisée afin de comprendre cette instabilité et identifier ce paramètre caché. On pourrait imaginer faire une analyse en spectrométrie de masse du cluster $P^{1152}BX2^*$ sur les régions qui ont reçu les piRNAs (*lacZ*) versus les régions qui n'ont pas reçu de piRNAs (*white*) afin de voir si le profil de protéines identifiées est le même sur les deux séquences. La difficulté serait d'isoler les séquences d'intérêt sans perdre les protéines qui y sont associées.

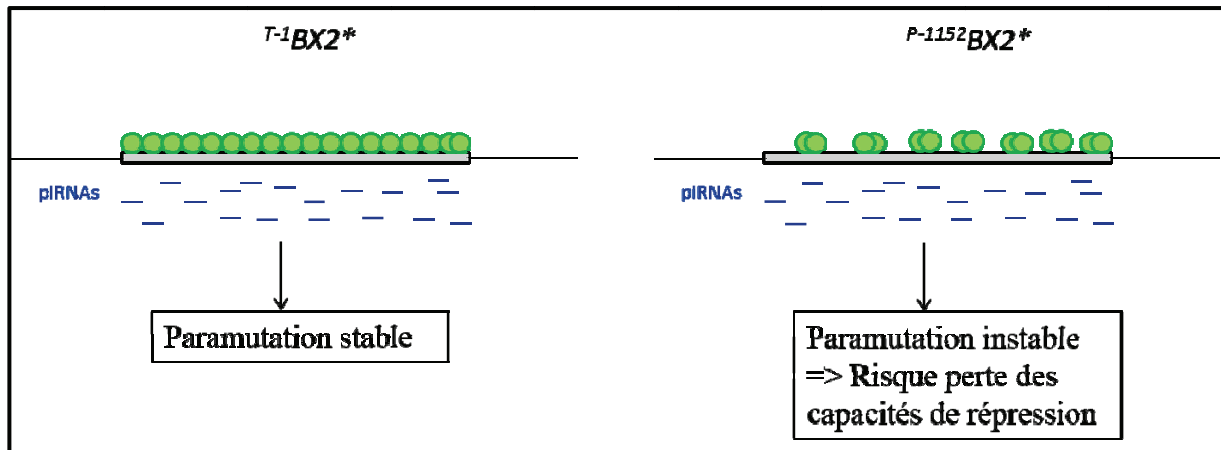


Figure 31 : Modèle de la différence de stabilité entre $T-1 BX2^*$ et $P-1152 BX2^*$.

Les locus $T-1 BX2^*$ et $P-1152 BX2^*$ sont représentés par un rectangle gris encadré en noir. Les piRNAs produits par les deux locus sont représentés en bleu sans distinction de sens. Le paramètre moléculaire inconnu (marques ou protéines chromatinienne) pouvant être responsable de l'instabilité est symbolisé par des ronds verts. Le locus $T-1 BX2^*$ produit des piRNAs tout le long de sa séquence de façon stable au cours du temps. L'hypothèse est que lors de la paramutation par les piRNAs de $T-1$ transmis maternellement, il y ait eu le dépôt d'une ou plusieurs marques et/ou protéines chromatiennes, symbolisées en vert, tout le long de la séquence. Ce dépôt serait maintenu au cours des générations et serait à l'origine de la stabilité de la paramutation. Le locus $P-1152 BX2^*$ produit des piRNAs tout le long de sa séquence de façon instable au cours du temps. L'hypothèse est que lors de la paramutation du locus $P-1152 BX2^*$ par des piRNAs de $P-1152$ transmis maternellement (uniquement sur certaines régions du locus), il y ait eu le dépôt de ces mêmes marques ou protéines chromatiennes mais uniquement sur les régions qui ont reçu des piRNAs de $P-1152$. Ce dépôt partiel serait responsable de l'instabilité de la paramutation avec des piRNAs partiellement homologues.

Néanmoins, ce modèle de travail (Figure 31) n'est pas pleinement satisfaisant et n'arrive pas à expliquer clairement la différence entre $T-1 BX2^*$, $P-1152 BX2^*$ et $BX2^O$. Il est possible que la différence entre ces loci soit due à la qualité de certains piRNAs qu'ils ont reçus maternellement. En effet, il est possible d'envisager que le locus $T-1$ produise un ou plusieurs piRNAs particuliers qui seraient responsables de la stabilité de la paramutation. En effet, la souche $T-1$ a été obtenue après traitement aux rayons X de la souche $BX2$. Une analyse par Southern blot a montré que le locus $T-1$ contenait toujours 7 copies du transgène répétés en tandem tout comme le locus $BX2$. Néanmoins, il est possible que le locus ait subi de petits réarrangements qui pourraient être à l'origine de la production de piRNAs particuliers. Il serait utile de séquencer la souche $T-1$ afin d'identifier d'une façon générale tous les changements induits par les rayons X. Le locus $BX2$ qui reçoit ces piRNAs particuliers, serait alors capable de les reproduire pour les donner à sa descendance. Dans le cas du locus $BX2$ paramuté pas $P-1152$, le locus $P-1152$ ne lui donnerait pas ces piRNAs particuliers. Par contre, dans le cas de l'activation du locus $BX2$ par la chaleur, qui conduit à

des lignées plus stables, il est possible que le locus *BX2* produise « naturellement » ces piRNAs particuliers car ici il ne reproduirait pas des piRNAs transmis.

Activation *de novo* d'un locus producteur de piRNAs par la chaleur :

L'activation *de novo* d'un cluster par les piRNAs est due à un phénomène d'héritage maternel. Ceci pourrait être à l'origine de l'activation et donc au maintien des clusters de piRNAs au cours des générations. Mais comment a pu démarrer le premier cluster ? Cette interrogation est supportée par le fait que nous ne savons pas pourquoi le locus *T-1* est devenu producteur de piRNAs puisque que celui-ci est issu de *BX2^{naïf}*. L'une des hypothèses est que la souche a pu être oubliée dans une pièce à température non contrôlée, supérieure à 25°C, qui serait à l'origine du démarrage du locus *T-1*. En effet, il a été montré dans le cas de la répression de *P*, que le locus télomérique répresseur producteur de piRNAs a de meilleures capacités répressives à haute température (29°C) (Ronsseray *et al.* 1984). Nous avons montré qu'il est possible d'activer *de novo* la production de piRNAs des loci *BX2^{naïf}* et *T-1^{naïf}* par la chaleur (29°C). Ces loci présentent toutes les propriétés fonctionnelles et moléculaires d'un locus activé *de novo* par la transmission de piRNAs maternels. Il est possible que l'augmentation de température agisse au niveau de l'expression du locus producteurs de piRNAs en affectant sa transcription. La biogenèse de piRNAs requière des protéines co-chaperones telles que Hsp83 (une protéine *heat-shock*), qui permettrait le chargement de la protéine Piwi sur les piRNAs (Specchia *et al.* 2010). Une autre hypothèse est qu'une augmentation de la température pourrait stimuler Hsp83 qui stimulerait, à son tour, la production de piRNAs.

Implications évolutives des phénomènes d'activations *de novo* des clusters producteurs de piRNAs :

D'un point de vue évolutif, l'activation *de novo* d'un locus producteur de piRNAs par certains facteurs environnementaux permettrait de comprendre l'émergence des premiers clusters de piRNAs. Ensuite, l'augmentation du nombre de clusters de piRNAs aurait pu se faire par transmission maternelle de piRNAs homologues et donc par paramutation. Les clusters de piRNAs sont composés de nombreuses séquences dégénérées de différents ETs pouvant être communs à plusieurs clusters. L'activation de clusters a pu se faire comme le principe des chaines de dominos, où un ou plusieurs clusters, ayant été activés par la chaleur, ont pu activer à leur tour d'autres clusters par la transmission de leurs piRNAs et ainsi de

suite, permettant ainsi la propagation à travers le génome de la capacité à produire des piRNAs (Figure 32). Cette dernière propagation du nombre de clusters de piRNAs au sein du génome n'aurait plus besoin de stress environnemental pour être maintenue. Néanmoins, l'instabilité observée lors de l'activation par des piRNAs partiellement homologues, révèle que cette propagation entre loci pas complètement homologues n'est pas toujours totalement efficace. Il est possible que ces activations permettent la propagation des clusters de piRNAs mais de façon modérée, afin de ne pas envahir tout le génome. En effet, il n'y a pas que les séquences répétées qui produisent des piRNAs. Il a été identifié des piRNAs issus de régions codantes qui produisent des piRNAs et peuvent être considérés comme des clusters unidirectionnels (Brennecke *et al.* 2007 ; Robine *et al.* 2009 ; Saito *et al.* 2009). Ainsi une efficacité modérée de l'activation des loci producteurs de piRNAs permettrait de maintenir une répression efficace d'ETs sans pour cela être néfaste pour l'expression du génome. Le paramètre moléculaire inconnu en **Figure 31** pourrait permettre cette efficacité modérée. Néanmoins, il est possible que dans la nature la paramutation partiellement homologue soit plus efficace que celle testée avec le locus *P-1152* car les clusters paramutagéniques restent présents dans le génome. Ce n'est pas le cas de la paramutation partiellement homologue où seuls les piRNAs du locus paramutagénique sont transmis.

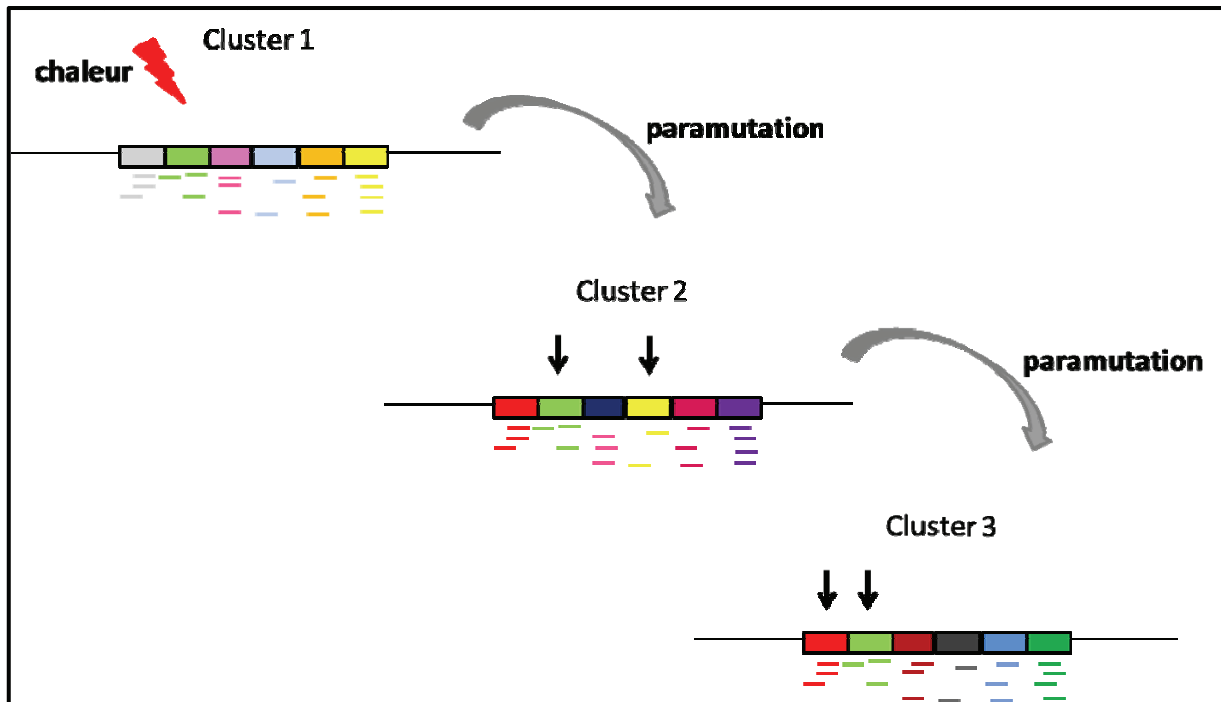


Figure 32 : Effet domino de la propagation en *cis* de l'activation *de novo* des clusters de piRNAs par chaleur et la paramutation.

Trois clusters de séquences répétées d'ETs différents sont représentés sous forme de rectangles de couleurs différentes pour chaque ET. L'éclair rouge symbolise le stress environnemental que produit la chaleur. Les piRNAs sont représentés sous forme de petits bâtonnets de couleur en fonction des ETs dont ils sont issus. Les flèches noires symbolisent les séquences homologues au piRNAs qui viennent paramuter. Dans ce modèle, nous proposons que le 1^{er} cluster de piRNAs ait pu être activé par la chaleur. Le cluster 1 pourrait donc produire de piRNAs. Le cluster 1 a des séquences partiellement homologues au cluster 2 (rectangle vert et jaune), et transmettrait des piRNAs de ces séquences ce qui permettrait de paramuter le cluster 2. Le cluster 2 produirait, à son tour, des piRNAs. Le cluster 2 a des séquences partiellement homologues au cluster 3 (rectangle rouge et vert), et transmettrait des piRNAs de ces séquences ce qui permettrait de paramuter le cluster 3. Le cluster 3 produirait à son tour des piRNAs.

Conclusion

Nous avons découvert qu'il est possible d'activer *de novo* la production de piRNAs par deux processus que sont la transmission maternelle de piRNAs et un facteur environnemental. Ces loci de piRNAs activés *de novo* présentent les propriétés fonctionnelles et moléculaires d'un cluster de piRNAs « classique ». En effet, le locus *BX2* activé *de novo* produit des piRNAs et des siRNAs qui sont dépendants des protéines de la biogenèse de piRNAs. L'étude d'une activation par des piRNAs ne couvrant qu'une partie du locus activé a montré que ce système d'activation est un phénomène général mais montre également que cette activation peut présenter parfois des effets modérés. L'activation *de novo* des loci producteurs de piRNAs par la chaleur peut permettre de comprendre comment s'est faite l'émergence des premiers clusters de piRNAs. L'ensemble des données de cette thèse pourraient permettre de comprendre comment, d'un point de vue évolutif et mécanistique, s'est faite à la fois l'émergence des premiers clusters de piRNAs ainsi que leur propagation dans le génome, une étape importante du façonnage de l'épigénome.

Références bibliographiques

Références Bibliographiques

- Alleman**, Mary, Lyudmila Sidorenko, Karen McGinnis, Vishwas Seshadri, Jane E. Dorweiler, Joshua White, Kristin Sikkink, et Vicki L. Chandler. **2006**. « An RNA-Dependent RNA Polymerase Is Required for Paramutation in Maize ». *Nature* 442 (7100): 295- 98. doi:10.1038/nature04884.
- Ambros**, Victor, Rosalind C. Lee, Ann Lavanway, Peter T. Williams, et David Jewell. **2003**. « MicroRNAs and Other Tiny Endogenous RNAs in *C. Elegans* ». *Current Biology: CB* 13 (10): 807- 18.
- Anand**, Amit, et Toshie Kai. **2012**. « The Tudor Domain Protein Kumo Is Required to Assemble the Nuage and to Generate Germline piRNAs in *Drosophila* ». *The EMBO Journal* 31 (4): 870- 82. doi:10.1038/emboj.2011.449.
- Arabidopsis Genome Initiative**. **2000**. « Analysis of the Genome Sequence of the Flowering Plant *Arabidopsis Thaliana* ». *Nature* 408 (6814): 796- 815. doi:10.1038/35048692.
- Aravin**, Alexei A., Gregory J. Hannon, et Julius Brennecke. **2007**. « The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race ». *Science (New York, N.Y.)* 318 (5851): 761- 64. doi:10.1126/science.1146484.
- Arteaga-Vazquez**, Mario, Lyudmila Sidorenko, Fernando A. Rabanal, Roli Shrivistava, Kan Nobuta, Pamela J. Green, Blake C. Meyers, et Vicki L. Chandler. **2010**. « RNA-Mediated Trans-Communication Can Establish Paramutation at the b1 Locus in Maize ». *Proceedings of the National Academy of Sciences of the United States of America* 107 (29): 12986- 91. doi:10.1073/pnas.1007972107.
- Bellen**, H. J., C. J. O’Kane, C. Wilson, U. Grossniklaus, R. K. Pearson, et W. J. Gehring. **1989**. « P-Element-Mediated Enhancer Detection: A Versatile Method to Study Development in *Drosophila* ». *Genes & Development* 3 (9): 1288- 1300.
- Bergman**, Casey M., Hadi Quesneville, Dominique Anxolabéhère, et Michael Ashburner. **2006**. « Recurrent Insertion and Duplication Generate Networks of Transposable Element Sequences in the *Drosophila Melanogaster* Genome ». *Genome Biology* 7 (11): R112. doi:10.1186/gb-2006-7-11-r112.
- Bernstein**, E., A. A. Caudy, S. M. Hammond, et G. J. Hannon. **2001**. « Role for a Bidentate Ribonuclease in the Initiation Step of RNA Interference ». *Nature* 409 (6818): 363- 66. doi:10.1038/35053110.
- Bingham**, P. M., M. G. Kidwell, et G. M. Rubin. **1982**. « The Molecular Basis of P-M Hybrid Dysgenesis: The Role of the P Element, a P-Strain-Specific Transposon Family ». *Cell* 29 (3): 995- 1004.
- Brennecke**, Julius, Alexei A. Aravin, Alexander Stark, Monica Dus, Manolis Kellis, Ravi Sachidanandam, et Gregory J. Hannon. **2007**. « Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in *Drosophila* ». *Cell* 128 (6): 1089- 1103. doi:10.1016/j.cell.2007.01.043.

- Brennecke, Julius, Colin D. Malone, Alexei A. Aravin, Ravi Sachidanandam, Alexander Stark, et Gregory J. Hannon. 2008.** « An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing ». *Science (New York, N.Y.)* 322 (5906): 1387- 92. doi:10.1126/science.1165171.
- Brink, R. A. 1956.** « A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible ». *Genetics* 41 (6): 872- 89.
- Brower-Toland, Brent, Seth D. Findley, Ling Jiang, Li Liu, Hang Yin, Monica Dus, Pei Zhou, Sarah C. R. Elgin, et Haifan Lin. 2007.** « Drosophila PIWI Associates with Chromatin and Interacts Directly with HP1a ». *Genes & Development* 21 (18): 2300- 2311. doi:10.1101/gad.1564307.
- Bühler, Marc, et Danesh Moazed. 2007.** « Transcription and RNAi in heterochromatic gene silencing ». *Nature Structural & Molecular Biology* 14 (11): 1041- 48. doi:10.1038/nsmb1315.
- C. elegans Sequencing Consortium. 1998.** « Genome Sequence of the Nematode C. Elegans: A Platform for Investigating Biology ». *Science (New York, N.Y.)* 282 (5396): 2012- 18.
- Celniker, Susan E, David A Wheeler, Brent Kronmiller, Joseph W Carlson, Aaron Halpern, Sandeep Patel, Mark Adams, et al. 2002.** « Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence ». *Genome Biology* 3 (12): research0079.1- 79.14. doi:10.1186/gb-2002-3-12-research0079.
- Chandler, Vicki L. 2007.** « Paramutation: From Maize to Mice ». *Cell* 128 (4): 641- 45. doi:10.1016/j.cell.2007.02.007.
- Chen, Yu, Attilio Pane, et Trudi Schüpbach. 2007.** « Cutoff and Aubergine Mutations Result in Retrotransposon Upregulation and Checkpoint Activation in Drosophila ». *Current Biology: CB* 17 (7): 637- 42. doi:10.1016/j.cub.2007.02.027.
- Chung, Wei-Jen, Katsutomo Okamura, Raquel Martin, et Eric C. Lai. 2008.** « Endogenous RNA Interference Provides a Somatic Defense against Drosophila Transposons ». *Current Biology: CB* 18 (11): 795- 802. doi:10.1016/j.cub.2008.05.006.
- Coe, E. H. 1966.** « The Properties, Origin, and Mechanism of Conversion-Type Inheritance at the B Locus in Maize ». *Genetics* 53 (6): 1035- 63.
- Creamer, Kevin M., et Janet F. Partridge. 2011.** « RITS-Connecting Transcription, RNA Interference, and Heterochromatin Assembly in Fission Yeast ». *Wiley Interdisciplinary Reviews. RNA* 2 (5): 632- 46. doi:10.1002/wrna.80.
- Czech, Benjamin, Colin D. Malone, Rui Zhou, Alexander Stark, Catherine Schlingehayde, Monica Dus, Norbert Perrimon, et al. 2008.** « An Endogenous Small Interfering RNA Pathway in Drosophila ». *Nature* 453 (7196): 798- 802. doi:10.1038/nature07007.
- Czech, Benjamin, Jonathan B. Preall, Jon McGinn, et Gregory J. Hannon. 2013.** « A Transcriptome-Wide RNAi Screen in the Drosophila Ovary Reveals Factors of the Germline piRNA Pathway ». *Molecular Cell* 50 (5): 749- 61. doi:10.1016/j.molcel.2013.04.007.

- Deng, W., et H. Lin. 2001.** « Asymmetric Germ Cell Division and Oocyte Determination during *Drosophila* Oogenesis ». *International Review of Cytology* 203: 93- 138.
- Dennis, Cynthia, Vanessa Zanni, Emilie Brasset, Angeline Eymery, Liang Zhang, Rana Mteirek, Silke Jensen, Yikang S. Rong, et Chantal Vaury. 2013.** « “Dot COM”, a Nuclear Transit Center for the Primary piRNA Pathway in *Drosophila* ». *PloS One* 8 (9): e72752. doi:10.1371/journal.pone.0072752.
- Desset, Sophie, Carine Meignin, Bernard Dastugue, et Chantal Vaury. 2003.** « COM, a Heterochromatic Locus Governing the Control of Independent Endogenous Retroviruses from *Drosophila Melanogaster* ». *Genetics* 164 (2): 501- 9.
- de Vanssay, Augustin, Anne-Laure Bougé, Antoine Boivin, Catherine Hermant, Laure Teyssset, Valérie Delmarre, Christophe Antoniewski, et Stéphane Ronsseray. 2012.** « Paramutation in *Drosophila* Linked to Emergence of a piRNA-Producing Locus ». *Nature* 490 (7418): 112- 15. doi:10.1038/nature11416.
- de Vanssay, Augustin, Anne-Laure Bougé, Antoine Boivin, Catherine Hermant, Laure Teyssset, Valérie Delmarre, Christophe Antoniewski, et Stéphane Ronsseray. 2013.** « piRNAs and Epigenetic Conversion in *Drosophila* ». *Fly* 7 (4): 237- 41. doi:10.4161/fly.26522.
- « Diapositive08.JPG (Image JPEG, 780 × 540 pixels) - Redimensionnée (0%) ». 2014. <http://sfa.univ-poitiers.fr/databio/DATABIO/GENETIQUE/Droso2001/Diapositive08.JPG>.
- Dönertas, Derya, Grzegorz Sienski, et Julius Brennecke. 2013.** « *Drosophila* Gtsfl Is an Essential Component of the Piwi-Mediated Transcriptional Silencing Complex ». *Genes & Development* 27 (15): 1693- 1705. doi:10.1101/gad.221150.113.
- Dorer, D. R., et S. Henikoff. 1994.** « Expansions of Transgene Repeats Cause Heterochromatin Formation and Gene Silencing in *Drosophila* ». *Cell* 77 (7): 993- 1002.
- Dorer, D. R., et S. Henikoff. 1997.** « Transgene Repeat Arrays Interact with Distant Heterochromatin and Cause Silencing in Cis and Trans ». *Genetics* 147 (3): 1181- 90.
- Dorweiler, J. E., C. C. Carey, K. M. Kubo, J. B. Hollick, J. L. Kermicle, et V. L. Chandler. 2000.** « Mediator of paramutation1 Is Required for Establishment and Maintenance of Paramutation at Multiple Maize Loci ». *The Plant Cell* 12 (11): 2101- 18.
- Drosophila 12 Genomes Consortium, Andrew G. Clark, Michael B. Eisen, Douglas R. Smith, Casey M. Bergman, Brian Oliver, Therese A. Markow, et al. 2007.** « Evolution of Genes and Genomes on the *Drosophila* Phylogeny ». *Nature* 450 (7167): 203- 18. doi:10.1038/nature06341.
- Dufourt, J., E. Brasset, S. Desset, P. Pouchin, et C. Vaury. 2011.** « Polycomb Group-Dependent, Heterochromatin Protein 1-Independent, Chromatin Structures Silence Retrotransposons in Somatic Tissues Outside Ovaries ». *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes* 18 (6): 451- 61. doi:10.1093/dnares/dsr031.

- Eberl, D. F., L. J. Lorenz, M. B. Melnick, V. Sood, P. Lasko, et N. Perrimon. 1997.** « A New Enhancer of Position-Effect Variegation in *Drosophila Melanogaster* Encodes a Putative RNA Helicase That Binds Chromosomes and Is Regulated by the Cell Cycle ». *Genetics* 146 (3): 951- 63.
- Eddy, E. M. 1974.** « Fine Structural Observations on the Form and Distribution of Nuage in Germ Cells of the Rat ». *The Anatomical Record* 178 (4): 731- 57.
doi:10.1002/ar.1091780406.
- Eddy, E. M. 1975.** « Germ Plasm and the Differentiation of the Germ Cell Line ». *International Review of Cytology* 43: 229- 80.
- Engels, W. R. 1979.** « Extrachromosomal Control of Mutability in *Drosophila Melanogaster* ». *Proceedings of the National Academy of Sciences of the United States of America* 76 (8): 4011- 15.
- Engels, W.R. 1989.** « P elements in *Drosophila* » (ED.Berg, D.E., Howe, M. M.) American Society for Microbiology, Washington,DC.
- Ephrussi, A., et R. Lehmann. 1992.** « Induction of Germ Cell Formation by Oskar ». *Nature* 358 (6385): 387- 92. doi:10.1038/358387a0.
- Fanti, L., D. R. Dorer, M. Berloco, S. Henikoff, et S. Pimpinelli. 1998.** « Heterochromatin Protein 1 Binds Transgene Arrays ». *Chromosoma* 107 (5): 286- 92.
- Feschotte, Cédric, et Ellen J. Pritham. 2007.** « DNA Transposons and the Evolution of Eukaryotic Genomes ». *Annual Review of Genetics* 41: 331- 68.
doi:10.1146/annurev.genet.40.110405.090448.
- Findley, Seth D., Mio Tamanaha, Nigel J. Clegg, et Hannele Ruohola-Baker. 2003.** « Maelstrom, a *Drosophila* Spindle-Class Gene, Encodes a Protein That Colocalizes with Vasa and RDE1/AGO1 Homolog, Aubergine, in Nuage ». *Development (Cambridge, England)* 130 (5): 859- 71.
- Finnegan, D. J. 1989.** « Eukaryotic Transposable Elements and Genome Evolution ». *Trends in Genetics: TIG* 5 (4): 103- 7.
- Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, et C. C. Mello. 1998.** « Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis Elegans* ». *Nature* 391 (6669): 806- 11. doi:10.1038/35888.
- Frank, Filipp, Jesse Hauver, Nahum Sonenberg, et Bhushan Nagar. 2012.** « Arabidopsis Argonaute MID Domains Use Their Nucleotide Specificity Loop to Sort Small RNAs ». *The EMBO Journal* 31 (17): 3588- 95. doi:10.1038/emboj.2012.204.
- Frank, Filipp, Nahum Sonenberg, et Bhushan Nagar. 2010.** « Structural Basis for 5'-Nucleotide Base-Specific Recognition of Guide RNA by Human AGO2 ». *Nature* 465 (7299): 818- 22. doi:10.1038/nature09039.
- Gatfield, D., H. Le Hir, C. Schmitt, I. C. Braun, T. Köcher, M. Wilm, et E. Izaurralde. 2001.** « The DEXH/D Box Protein HEL/UAP56 Is Essential for mRNA Nuclear Export in *Drosophila* ». *Current Biology: CB* 11 (21): 1716- 21.

- Ghildiyal**, Megha, Hervé Seitz, Michael D. Horwich, Chengjian Li, Tingting Du, Soohyun Lee, Jia Xu, et al. **2008**. « Endogenous siRNAs Derived from Transposons and mRNAs in *Drosophila* Somatic Cells ». *Science (New York, N.Y.)* 320 (5879): 1077- 81. doi:10.1126/science.1157396.
- Ghildiyal**, Megha, et Phillip D. Zamore. **2009**. « Small Silencing RNAs: An Expanding Universe ». *Nature Reviews. Genetics* 10 (2): 94- 108. doi:10.1038/nrg2504.
- González-Reyes**, Acaimo. **2003**. « Stem Cells, Niches and Cadherins: A View from *Drosophila* ». *Journal of Cell Science* 116 (Pt 6): 949- 54.
- Goriaux**, Coline, Sophie Desset, Yoan Renaud, Chantal Vauray, et Emilie Brasset. **2014**. « Transcriptional Properties and Splicing of the Flamenco piRNA Cluster ». *EMBO Reports* 15 (4): 411- 18. doi:10.1002/embr.201337898.
- Gunawardane**, Lalith S., Kuniaki Saito, Kazumichi M. Nishida, Keita Miyoshi, Yoshinori Kawamura, Tomoko Nagami, Haruhiko Siomi, et Mikiko C. Siomi. **2007**. « A Slicer-Mediated Mechanism for Repeat-Associated siRNA 5' End Formation in *Drosophila* ». *Science (New York, N.Y.)* 315 (5818): 1587- 90. doi:10.1126/science.1140494.
- Hamilton**, A. J., et D. C. Baulcombe. **1999**. « A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants ». *Science (New York, N.Y.)* 286 (5441): 950 - 52.
- Hamilton**, Andrew, Olivier Voinnet, Louise Chappell, et David Baulcombe. **2002**. « Two Classes of Short Interfering RNA in RNA Silencing ». *The EMBO Journal* 21 (17): 4671 - 79.
- Hammond**, S. M., E. Bernstein, D. Beach, et G. J. Hannon. **2000**. « An RNA-Directed Nuclease Mediates Post-Transcriptional Gene Silencing in *Drosophila* Cells ». *Nature* 404 (6775): 293- 96. doi:10.1038/35005107.
- Handler**, Dominik, Katharina Meixner, Manfred Pizka, Kathrin Lauss, Christopher Schmied, Franz Sebastian Gruber, et Julius Brennecke. **2013**. « The Genetic Makeup of the *Drosophila* piRNA Pathway ». *Molecular Cell* 50 (5): 762- 77. doi:10.1016/j.molcel.2013.04.031.
- Handler**, Dominik, Daniel Olivieri, Maria Novatchkova, Franz Sebastian Gruber, Katharina Meixner, Karl Mechtler, Alexander Stark, Ravi Sachidanandam, et Julius Brennecke. **2011**. « A Systematic Analysis of *Drosophila* TUDOR Domain-Containing Proteins Identifies Vreteno and the Tdrd12 Family as Essential Primary piRNA Pathway Factors ». *The EMBO Journal* 30 (19): 3977- 93. doi:10.1038/emboj.2011.308.
- Harris**, A. N., et P. M. Macdonald. **2001**. « Aubergine Encodes a *Drosophila* Polar Granule Component Required for Pole Cell Formation and Related to eIF2C ». *Development (Cambridge, England)* 128 (14): 2823- 32.
- Hay**, B., L. Y. Jan, et Y. N. Jan. **1990**. « Localization of Vasa, a Component of *Drosophila* Polar Granules, in Maternal-Effect Mutants That Alter Embryonic Anteroposterior Polarity ». *Development (Cambridge, England)* 109 (2): 425- 33.

- Horwich**, Michael D., Chengjian Li, Christian Matranga, Vasily Vagin, Gwen Farley, Peng Wang, et Phillip D. Zamore. **2007**. « The Drosophila RNA Methyltransferase, DmHen1, Modifies Germline piRNAs and Single-Stranded siRNAs in RISC ». *Current Biology: CB* 17 (14): 1265- 72. doi:10.1016/j.cub.2007.06.030.
- Huang**, Huiyan, Qun Gao, Xiaoxue Peng, Seok-Yong Choi, Krishna Sarma, Hongmei Ren, Andrew J. Morris, et Michael A. Frohman. **2011**. « piRNA-Associated Germline Nuage Formation and Spermatogenesis Require MitoPLD Profusogenic Mitochondrial-Surface Lipid Signaling ». *Developmental Cell* 20 (3): 376- 87. doi:10.1016/j.devcel.2011.01.004.
- Ipsaro**, Jonathan J., Astrid D. Haase, Simon R. Knott, Leemor Joshua-Tor, et Gregory J. Hannon. **2012**. « The Structural Biochemistry of Zucchini Implicates It as a Nuclease in piRNA Biogenesis ». *Nature* 491 (7423): 279- 83. doi:10.1038/nature11502.
- Ishizu**, Hirotugu, Haruhiko Siomi, et Mikiko C. Siomi. **2012**. « Biology of PIWI-Interacting RNAs: New Insights into Biogenesis and Function inside and outside of Germlines ». *Genes & Development* 26 (21): 2361- 73. doi:10.1101/gad.203786.112.
- Jiao**, Xinfu, Jeong Ho Chang, Turgay Kilic, Liang Tong, et Megerditch Kiledjian. **2013**. « A Mammalian Pre-mRNA 5' End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing ». *Molecular Cell* 50 (1): 104- 15. doi:10.1016/j.molcel.2013.02.017.
- Jiao**, Xinfu, Song Xiang, Chanseok Oh, Charles E. Martin, Liang Tong, et Megerditch Kiledjian. **2010**. « Identification of a Quality-Control Mechanism for mRNA 5'-End Capping ». *Nature* 467 (7315): 608- 11. doi:10.1038/nature09338.
- Josse**, Thibaut, Corinne Maurel-Zaffran, Augustin de Vanssay, Laure Teyssset, Anne-Laure Todeschini, Valerie Delmarre, Nicole Chaminade, Dominique Anxolabéhère, et Stéphane Ronsseray. **2008**. « Telomeric Trans-Silencing in Drosophila Melanogaster: Tissue Specificity, Development and Functional Interactions between Non-Homologous Telomeres ». *PloS One* 3 (9): e3249. doi:10.1371/journal.pone.0003249.
- Josse**, Thibaut, Laure Teyssset, Anne-Laure Todeschini, Clara M. Sidor, Dominique Anxolabéhère, et Stéphane Ronsseray. **2007**. « Telomeric Trans-Silencing: An Epigenetic Repression Combining RNA Silencing and Heterochromatin Formation ». *PLoS Genetics* 3 (9): 1633- 43. doi:10.1371/journal.pgen.0030158.
- Jurka**, J., V. V. Kapitonov, A. Pavlicek, P. Klonowski, O. Kohany, et J. Walichiewicz. **2005**. « Repbase Update, a Database of Eukaryotic Repetitive Elements ». *Cytogenetic and Genome Research* 110 (1-4): 462- 67. doi:10.1159/000084979.
- Katzenberger**, Rebeccah J., Matthew S. Marengo, et David A. Wassarman. **2009**. « Control of Alternative Splicing by Signal-Dependent Degradation of Splicing-Regulatory Proteins ». *The Journal of Biological Chemistry* 284 (16): 10737- 46. doi:10.1074/jbc.M809506200.
- Khurana**, Jaspreet S., Jie Wang, Jia Xu, Birgit S. Koppetsch, Travis C. Thomson, Anetta Nowosielska, Chengjian Li, Phillip D. Zamore, Zhiping Weng, et William E. Theurkauf. **2011**. « Adaptation to P Element Transposon Invasion in Drosophila Melanogaster ». *Cell* 147 (7): 1551- 63. doi:10.1016/j.cell.2011.11.042.

- Kidwell, M. G., J. F. Kidwell, et J. A. Sved. 1977.** « Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination ». *Genetics* 86 (4): 813- 33.
- Kim, Minkyu, Nevan J. Krogan, Lidia Vasiljeva, Oliver J. Rando, Eduard Nedeia, Jack F. Greenblatt, et Stephen Buratowski. 2004.** « The Yeast Rat1 Exonuclease Promotes Transcription Termination by RNA Polymerase II ». *Nature* 432 (7016): 517- 22. doi:10.1038/nature03041.
- Kim, V. Narry, Jinju Han, et Mikiko C. Siomi. 2009.** « Biogenesis of Small RNAs in Animals ». *Nature Reviews. Molecular Cell Biology* 10 (2): 126- 39. doi:10.1038/nrm2632.
- King, Robert C. 1970.** *Ovarian Development in Drosophila Melanogaster*. Academic Press.
- Klattenhoff, Carla, Hualin Xi, Chengjian Li, Soohyun Lee, Jia Xu, Jaspreet S. Khurana, Fan Zhang, et al. 2009.** « The Drosophila HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters ». *Cell* 138 (6): 1137- 49. doi:10.1016/j.cell.2009.07.014.
- Klenov, Mikhail S., Sergey A. Lavrov, Alina P. Korbut, Anastasia D. Stolyarenko, Evgeny Y. Yakushev, Michael Reuter, Ramesh S. Pillai, et Vladimir A. Gvozdev. 2014.** « Impact of Nuclear Piwi Elimination on Chromatin State in Drosophila Melanogaster Ovaries ». *Nucleic Acids Research* 42 (10): 6208- 18. doi:10.1093/nar/gku268.
- Klenov, Mikhail S., Sergey A. Lavrov, Anastasia D. Stolyarenko, Sergey S. Ryazansky, Alexei A. Aravin, Thomas Tuschl, et Vladimir A. Gvozdev. 2007.** « Repeat-Associated siRNAs Cause Chromatin Silencing of Retrotransposons in the Drosophila Melanogaster Germline ». *Nucleic Acids Research* 35 (16): 5430- 38. doi:10.1093/nar/gkm576.
- Klenov, Mikhail S., Olesya A. Sokolova, Evgeny Y. Yakushev, Anastasia D. Stolyarenko, Elena A. Mikhaleva, Sergey A. Lavrov, et Vladimir A. Gvozdev. 2011.** « Separation of Stem Cell Maintenance and Transposon Silencing Functions of Piwi Protein ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (46): 18760- 65. doi:10.1073/pnas.1106676108.
- Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001.** « Initial Sequencing and Analysis of the Human Genome ». *Nature* 409 (6822): 860 - 921. doi:10.1038/35057062.
- Lasko, P. F., et M. Ashburner. 1990.** « Posterior Localization of Vasa Protein Correlates With, but Is Not Sufficient For, Pole Cell Development ». *Genes & Development* 4 (6): 905- 21.
- Lécher, P., A. Bucheton, et A. Péliisson. 1997.** « Expression of the Drosophila Retrovirus Gypsy as Ultrastructurally Detectable Particles in the Ovaries of Flies Carrying a Permissive Flamenco Allele ». *The Journal of General Virology* 78 (Pt 9) (septembre): 2379- 88.
- Lee, Young Sik, Kenji Nakahara, John W. Pham, Kevin Kim, Zhengying He, Erik J. Sontheimer, et Richard W. Carthew. 2004.** « Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways ». *Cell* 117 (1): 69- 81.

- Lehmann, R.**, et A. Ephrussi. **1994**. « Germ Plasm Formation and Germ Cell Determination in *Drosophila* ». *Ciba Foundation Symposium* 182: 282- 96; discussion 296- 300.
- Le Thomas, Adrien**, Alicia K. Rogers, Alexandre Webster, Georgi K. Marinov, Susan E. Liao, Edward M. Perkins, Junho K. Hur, Alexei A. Aravin, et Katalin Fejes Tóth. **2013**. « Piwi Induces piRNA-Guided Transcriptional Silencing and Establishment of a Repressive Chromatin State ». *Genes & Development* 27 (4): 390- 99. doi:10.1101/gad.209841.112.
- Le Thomas, Adrien**, Evelyn Stuwe, Sisi Li, Jiamu Du, Georgi Marinov, Nikolay Rozhkov, Yung-Chia Ariel Chen, et al. **2014**. « Transgenerationally Inherited piRNAs Trigger piRNA Biogenesis by Changing the Chromatin of piRNA Clusters and Inducing Precursor Processing ». *Genes & Development* 28 (15): 1667- 80. doi:10.1101/gad.245514.114.
- Liang, L.**, W. Diehl-Jones, et P. Lasko. **1994**. « Localization of Vasa Protein to the *Drosophila* Pole Plasm Is Independent of Its RNA-Binding and Helicase Activities ». *Development (Cambridge, England)* 120 (5): 1201- 11.
- Li, Chengjian**, Vasily V. Vagin, Soohyun Lee, Jia Xu, Shengmei Ma, Hualin Xi, Hervé Seitz, et al. **2009**. « Collapse of Germline piRNAs in the Absence of Argonaute3 Reveals Somatic piRNAs in Flies ». *Cell* 137 (3): 509- 21. doi:10.1016/j.cell.2009.04.027.
- Lim, Ai Khim**, et Toshie Kai. **2007**. « Unique Germ-Line Organelle, Nuage, Functions to Repress Selfish Genetic Elements in *Drosophila Melanogaster* ». *Proceedings of the National Academy of Sciences of the United States of America* 104 (16): 6714- 19. doi:10.1073/pnas.0701920104.
- Linder, Patrick**, et Eckhard Jankowsky. **2011**. « From Unwinding to Clamping - the DEAD Box RNA Helicase Family ». *Nature Reviews. Molecular Cell Biology* 12 (8): 505- 16. doi:10.1038/nrm3154.
- Li, Xin Zhiguo**, Christian K. Roy, Xianjun Dong, Ewelina Bolcun-Filas, Jie Wang, Bo W. Han, Jia Xu, et al. **2013**. « An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes ». *Molecular Cell* 50 (1): 67- 81. doi:10.1016/j.molcel.2013.02.016.
- Malone, Colin D.**, Julius Brennecke, Monica Dus, Alexander Stark, W. Richard McCombie, Ravi Sachidanandam, et Gregory J. Hannon. **2009**. « Specialized piRNA Pathways Act in Germline and Somatic Tissues of the *Drosophila* Ovary ». *Cell* 137 (3): 522- 35. doi:10.1016/j.cell.2009.03.040.
- Marin, L.**, M. Lehmann, D. Nouaud, H. Izaabel, D. Anxolabéhère, et S. Ronsseray. **2000**. « P-Element Repression in *Drosophila Melanogaster* by a Naturally Occurring Defective Telomeric P Copy ». *Genetics* 155 (4): 1841- 54.
- Mason, James M.**, Radmila Capkova Frydrychova, et Harald Biessmann. **2008**. « *Drosophila* telomeres: an exception providing new insights ». *BioEssays : news and reviews in molecular, cellular and developmental biology* 30 (1): 25- 37. doi:10.1002/bies.20688.

- McClintock, B. 1950.** « The Origin and Behavior of Mutable Loci in Maize ». *Proceedings of the National Academy of Sciences of the United States of America* 36 (6): 344- 55.
- Megosh, Heather B., Daniel N. Cox, Chris Campbell, et Haifan Lin. 2006.** « The Role of PIWI and the miRNA Machinery in Drosophila Germline Determination ». *Current Biology: CB* 16 (19): 1884- 94. doi:10.1016/j.cub.2006.08.051.
- Meignin, Carine, et Ilan Davis. 2008.** « UAP56 RNA Helicase Is Required for Axis Specification and Cytoplasmic mRNA Localization in Drosophila ». *Developmental Biology* 315 (1): 89- 98. doi:10.1016/j.ydbio.2007.12.004.
- Mével-Ninio, Maryvonne, Alain Pelisson, Jennifer Kinder, Ana Regina Campos, et Alain Bucheton. 2007.** « The Flamenco Locus Controls the Gypsy and ZAM Retroviruses and Is Required for Drosophila Oogenesis ». *Genetics* 175 (4): 1615- 24. doi:10.1534/genetics.106.068106.
- Miyoshi, Keita, Hiroko Tsukumo, Tomoko Nagami, Haruhiko Siomi, et Mikiko C. Siomi. 2005.** « Slicer function of Drosophila Argonautes and its involvement in RISC formation ». *Genes & Development* 19 (23): 2837- 48. doi:10.1101/gad.1370605.
- Mohn, Fabio, Grzegorz Sienski, Dominik Handler, et Julius Brennecke. 2014.** « The Rhino-Deadlock-Cutoff Complex Licenses Noncanonical Transcription of Dual-Strand piRNA Clusters in Drosophila ». *Cell* 157 (6): 1364- 79. doi:10.1016/j.cell.2014.04.031.
- Moshkovich, Nellie, et Elissa P. Lei. 2010.** « HP1 Recruitment in the Absence of Argonaute Proteins in Drosophila ». *PLoS Genetics* 6 (3): e1000880. doi:10.1371/journal.pgen.1000880.
- Mouse Genome Sequencing Consortium, Robert H. Waterston, Kerstin Lindblad-Toh, Ewan Birney, Jane Rogers, Josep F. Abril, Pankaj Agarwal, et al. 2002.** « Initial Sequencing and Comparative Analysis of the Mouse Genome ». *Nature* 420 (6915): 520- 62. doi:10.1038/nature01262.
- Muerdter, Felix, Paloma M. Guzzardo, Jesse Gillis, Yicheng Luo, Yang Yu, Caifu Chen, Richard Fekete, et Gregory J. Hannon. 2013.** « A Genome-Wide RNAi Screen Draws a Genetic Framework for Transposon Control and Primary piRNA Biogenesis in Drosophila ». *Molecular Cell* 50 (5): 736- 48. doi:10.1016/j.molcel.2013.04.006.
- Murota, Yukiko, Hirotugu Ishizu, Shinichi Nakagawa, Yuka W. Iwasaki, Shinsuke Shibata, Miha K. Kamatani, Kuniaki Saito, Hideyuki Okano, Haruhiko Siomi, et Mikiko C. Siomi. 2014.** « Yb Integrates piRNA Intermediates and Processing Factors into Perinuclear Bodies to Enhance piRISC Assembly ». *Cell Reports* 8 (1): 103- 13. doi:10.1016/j.celrep.2014.05.043.
- Nishida, Kazumichi M., Tomoko N. Okada, Takeshi Kawamura, Toutai Mituyama, Yoshinori Kawamura, Sachi Inagaki, Haidong Huang, et al. 2009.** « Functional Involvement of Tudor and dPRMT5 in the piRNA Processing Pathway in Drosophila Germlines ». *The EMBO Journal* 28 (24): 3820- 31. doi:10.1038/emboj.2009.365.

- Nishimasu**, Hiroshi, Hirotsugu Ishizu, Kuniaki Saito, Satoshi Fukuhara, Miharu K. Kamatani, Luc Bonnefond, Naoki Matsumoto, et al. **2012**. « Structure and Function of Zucchini Endoribonuclease in piRNA Biogenesis ». *Nature* 491 (7423): 284- 87. doi:10.1038/nature11509.
- O'Hare**, K., et G. M. Rubin. **1983**. « Structures of P Transposable Elements and Their Sites of Insertion and Excision in the *Drosophila Melanogaster* Genome ». *Cell* 34 (1): 25- 35.
- Okamura**, Katsutomo, Sudha Balla, Raquel Martin, Na Liu, et Eric C Lai. **2008**. « Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in *Drosophila melanogaster* ». *Nature structural & molecular biology* 15 (6): 581- 90. doi:10.1038/nsmb.1438.
- Okamura**, Katsutomo, Wei-Jen Chung, J. Graham Ruby, Huili Guo, David P. Bartel, et Eric C. Lai. **2008**. « The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs ». *Nature* 453 (7196): 803- 6. doi:10.1038/nature07015.
- Olivieri**, Daniel, Kirsten-André Senti, Sailakshmi Subramanian, Ravi Sachidanandam, et Julius Brennecke. **2012**. « The Cochaperone Shutdown Defines a Group of Biogenesis Factors Essential for All piRNA Populations in *Drosophila* ». *Molecular Cell* 47 (6): 954- 69. doi:10.1016/j.molcel.2012.07.021.
- Olivieri**, Daniel, Martina M. Sykora, Ravi Sachidanandam, Karl Mechtler, et Julius Brennecke. **2010**. « An in Vivo RNAi Assay Identifies Major Genetic and Cellular Requirements for Primary piRNA Biogenesis in *Drosophila* ». *The EMBO Journal* 29 (19): 3301- 17. doi:10.1038/emboj.2010.212.
- Pane**, Attilio, Peng Jiang, Dorothy Yanling Zhao, Mona Singh, et Trudi Schüpbach. **2011**. « The Cutoff Protein Regulates piRNA Cluster Expression and piRNA Production in the *Drosophila* Germline ». *The EMBO Journal* 30 (22): 4601- 15. doi:10.1038/emboj.2011.334.
- Pane**, Attilio, Kristina Wehr, et Trudi Schüpbach. **2007**. « Zucchini and Squash Encode Two Putative Nucleases Required for rasiRNA Production in the *Drosophila* Germline ». *Developmental Cell* 12 (6): 851- 62. doi:10.1016/j.devcel.2007.03.022.
- Park**, Jung W., Katherine Parisky, Alicia M. Celotto, Robert A. Reenan, et Brenton R. Graveley. **2004**. « Identification of alternative splicing regulators by RNA interference in *Drosophila* ». *Proceedings of the National Academy of Sciences of the United States of America* 101 (45): 15974- 79. doi:10.1073/pnas.0407004101.
- Perrat**, Paola N., Shamik DasGupta, Jie Wang, William Theurkauf, Zhiping Weng, Michael Rosbash, et Scott Waddell. **2013**. « Transposition-Driven Genomic Heterogeneity in the *Drosophila* Brain ». *Science (New York, N.Y.)* 340 (6128): 91- 95. doi:10.1126/science.1231965.
- Picard**, G. **1976**. « Non-Mendelian Female Sterility in *Drosophila Melanogaster*: Hereditary Transmission of I Factor ». *Genetics* 83 (1): 107- 23.

- Preall**, Jonathan B., Benjamin Czech, Paloma M. Guzzardo, Felix Muerdter, et Gregory J. Hannon. **2012**. « Shutdown Is a Component of the *Drosophila* piRNA Biogenesis Machinery ». *RNA (New York, N.Y.)* 18 (8): 1446- 57. doi:10.1261/rna.034405.112.
- Pritham**, Ellen J. **2009**. « Transposable Elements and Factors Influencing Their Success in Eukaryotes ». *The Journal of Heredity* 100 (5): 648- 55. doi:10.1093/jhered/esp065.
- Pryde**, Fiona E, Hazel C Gorham, et Edward J Louis. **1997**. « Chromosome ends: all the same under their caps ». *Current Opinion in Genetics & Development* 7 (6): 822- 28. doi:10.1016/S0959-437X(97)80046-9.
- Rand**, Tim A., Krzysztof Ginalski, Nick V. Grishin, et Xiaodong Wang. **2004**. « Biochemical Identification of Argonaute 2 as the Sole Protein Required for RNA-Induced Silencing Complex Activity ». *Proceedings of the National Academy of Sciences of the United States of America* 101 (40): 14385- 89. doi:10.1073/pnas.0405913101.
- Rangan**, Prashanth, Colin D. Malone, Caryn Navarro, Sam P. Newbold, Patrick S. Hayes, Ravi Sachidanandam, Gregory J. Hannon, et Ruth Lehmann. **2011**. « piRNA Production Requires Heterochromatin Formation in *Drosophila* ». *Current Biology: CB* 21 (16): 1373 - 79. doi:10.1016/j.cub.2011.06.057.
- Rassoulzadegan**, Minoo, Valérie Grandjean, Pierre Gounon, Stéphane Vincent, Isabelle Gillot, et François Cuzin. **2006**. « RNA-Mediated Non-Mendelian Inheritance of an Epigenetic Change in the Mouse ». *Nature* 441 (7092): 469- 74. doi:10.1038/nature04674.
- Reinhart**, Brenda J., et David P. Bartel. **2002**. « Small RNAs Correspond to Centromere Heterochromatic Repeats ». *Science (New York, N.Y.)* 297 (5588): 1831. doi:10.1126/science.1077183.
- Rio**, D. C. **1991**. « Regulation of *Drosophila* P Element Transposition ». *Trends in Genetics: TIG* 7 (9): 282- 87. doi:10.1016/0168-9525(91)90309-E.
- Robert**, V., N. Prud'homme, A. Kim, A. Bucheton, et A. Péliisson. **2001**. « Characterization of the Flamenco Region of the *Drosophila Melanogaster* Genome ». *Genetics* 158 (2): 701- 13.
- Robine**, Nicolas, Nelson C. Lau, Sudha Balla, Zhigang Jin, Katsutomo Okamura, Satomi Kuramochi-Miyagawa, Michael D. Blower, et Eric C. Lai. **2009**. « A Broadly Conserved Pathway Generates 3'UTR-Directed Primary piRNAs ». *Current Biology: CB* 19 (24): 2066- 76. doi:10.1016/j.cub.2009.11.064.
- Roche**, S. E., et D. C. **Rio**. **1998**. « Trans-Silencing by P Elements Inserted in Subtelomeric Heterochromatin Involves the *Drosophila* Polycomb Group Gene, Enhancer of Zeste ». *Genetics* 149 (4): 1839- 55.
- Ronsseray**, S., D. Anxolabéhère, et G. Périquet. **1984**. « Hybrid Dysgenesis in *Drosophila Melanogaster*: Influence of Temperature on Cytotype Determination in the P-M System ». *Molecular & General Genetics: MGG* 196 (1): 17- 23.
- Ronsseray**, S., A. Boivin, et D. Anxolabéhère. **2001**. « P-Element Repression in *Drosophila Melanogaster* by Variegating Clusters of P-lacZ-White Transgenes ». *Genetics* 159 (4): 1631- 42.

- Ronsseray, S., M. Lehmann, et D. Anxolabéhère. 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 (2): 501- 12.
- Ronsseray, S., M. Lehmann, D. Nouaud, et D. Anxolabéhère. 1996.** « The Regulatory Properties of Autonomous Subtelomeric P Elements Are Sensitive to a Suppressor of Variegation in *Drosophila Melanogaster* ». *Genetics* 143 (4): 1663- 74.
- Ronsseray, S., B. Lemaitre, et D. Coen. 1993.** « Maternal Inheritance of P Cytotype in *Drosophila Melanogaster*: A “Pre-P Cytotype” Is Strictly Extra-Chromosomally Transmitted ». *Molecular & General Genetics: MGG* 241 (1-2): 115- 23.
- Saito, Kuniaki, Sachi Inagaki, Toutai Mituyama, Yoshinori Kawamura, Yukiteru Ono, Eri Sakota, Hazuki Kotani, Kiyoshi Asai, Haruhiko Siomi, et Mikiko C. Siomi. 2009.** « A Regulatory Circuit for Piwi by the Large Maf Gene Traffic Jam in *Drosophila* ». *Nature* 461 (7268): 1296- 99. doi:10.1038/nature08501.
- Saito, Kuniaki, Hirotugu Ishizu, Mihar Komai, Hazuki Kotani, Yoshinori Kawamura, Kazumichi M. Nishida, Haruhiko Siomi, et Mikiko C. Siomi. 2010.** « Roles for the Yb Body Components Armitage and Yb in Primary piRNA Biogenesis in *Drosophila* ». *Genes & Development* 24 (22): 2493- 98. doi:10.1101/gad.1989510.
- Saito, Kuniaki, Kazumichi M. Nishida, Tomoko Mori, Yoshinori Kawamura, Keita Miyoshi, Tomoko Nagami, Haruhiko Siomi, et Mikiko C. Siomi. 2006.** « Specific Association of Piwi with rasiRNAs Derived from Retrotransposon and Heterochromatic Regions in the *Drosophila* Genome ». *Genes & Development* 20 (16): 2214- 22. doi:10.1101/gad.1454806.
- Saito, Kuniaki, Yuriko Sakaguchi, Takeo Suzuki, Tsutomu Suzuki, Haruhiko Siomi, et Mikiko C. Siomi. 2007.** « Pimet, the *Drosophila* Homolog of HEN1, Mediates 2'-O-Methylation of Piwi- Interacting RNAs at Their 3' Ends ». *Genes & Development* 21 (13): 1603- 8. doi:10.1101/gad.1563607.
- Schalch, Thomas, Godwin Job, Sreenath Shanker, Janet F. Partridge, et Leemor Joshua-Tor. 2011.** « The Chp1-Tas3 Core Is a Multifunctional Platform Critical for Gene Silencing by RITS ». *Nature Structural & Molecular Biology* 18 (12): 1351- 57. doi:10.1038/nsmb.2151.
- Schnable, Patrick S., Doreen Ware, Robert S. Fulton, Joshua C. Stein, Fusheng Wei, Shiran Pasternak, Chengzhi Liang, et al. 2009.** « The B73 Maize Genome: Complexity, Diversity, and Dynamics ». *Science (New York, N.Y.)* 326 (5956): 1112- 15. doi:10.1126/science.1178534.
- Schüpbach, T., et E. Wieschaus. 1991.** « Female Sterile Mutations on the Second Chromosome of *Drosophila Melanogaster*. II. Mutations Blocking Oogenesis or Altering Egg Morphology ». *Genetics* 129 (4): 1119- 36.
- Senti, Kirsten-André, et Julius Brennecke. 2010.** « The piRNA Pathway: A Fly's Perspective on the Guardian of the Genome ». *Trends in Genetics: TIG* 26 (12): 499- 509. doi:10.1016/j.tig.2010.08.007.

- Sentmanat**, Monica F., et Sarah C. R. Elgin. **2012**. « Ectopic Assembly of Heterochromatin in *Drosophila Melanogaster* Triggered by Transposable Elements ». *Proceedings of the National Academy of Sciences of the United States of America* 109 (35): 14104- 9. doi:10.1073/pnas.1207036109.
- Sidorenko**, Lyudmila, Jane E. Dorweiler, A. Mark Cigan, Mario Arteaga-Vazquez, Meenal Vyas, Jerry Kermicle, Diane Jurcin, Jan Brzeski, Yu Cai, et Vicki L. Chandler. **2009**. « A Dominant Mutation in Mediator of paramutation2, One of Three Second-Largest Subunits of a Plant-Specific RNA Polymerase, Disrupts Multiple siRNA Silencing Processes ». *PLoS Genetics* 5 (11): e1000725. doi:10.1371/journal.pgen.1000725.
- Sienski**, Grzegorz, Derya Dönertas, et Julius Brennecke. **2012**. « Transcriptional Silencing of Transposons by Piwi and Maelstrom and Its Impact on Chromatin State and Gene Expression ». *Cell* 151 (5): 964- 80. doi:10.1016/j.cell.2012.10.040.
- Simmons**, Michael J., John D. Raymond, Jarad B. Niemi, Jeremy R. Stuart, et Peter J. Merriman. **2004**. « The P Cytotype in *Drosophila Melanogaster*: A Maternally Transmitted Regulatory State of the Germ Line Associated with Telomeric P Elements ». *Genetics* 166 (1): 243- 54.
- Siomi**, Mikiko C., Taro Mannen, et Haruhiko Siomi. **2010**. « How Does the Royal Family of Tudor Rule the PIWI-Interacting RNA Pathway? ». *Genes & Development* 24 (7): 636- 46. doi:10.1101/gad.1899210.
- Siomi**, Mikiko C., Kaoru Sato, Dubravka Pezic, et Alexei A. Aravin. **2011**. « PIWI-Interacting Small RNAs: The Vanguard of Genome Defence ». *Nature Reviews. Molecular Cell Biology* 12 (4): 246- 58. doi:10.1038/nrm3089.
- Specchia**, Valeria, Lucia Piacentini, Patrizia Tritto, Laura Fanti, Rosalba D'Alessandro, Gioacchino Palumbo, Sergio Pimpinelli, et Maria P. Bozzetti. **2010**. « Hsp90 Prevents Phenotypic Variation by Suppressing the Mutagenic Activity of Transposons ». *Nature* 463 (7281): 662- 65. doi:10.1038/nature08739.
- Spradling**, A., D. Drummond-Barbosa, et T. Kai. **2001**. « Stem Cells Find Their Niche ». *Nature* 414 (6859): 98- 104. doi:10.1038/35102160.
- Stam**, Maïke, Christiane Belele, Jane E. Dorweiler, et Vicki L. Chandler. **2002**. « Differential Chromatin Structure within a Tandem Array 100 Kb Upstream of the Maize b1 Locus Is Associated with Paramutation ». *Genes & Development* 16 (15): 1906- 18. doi:10.1101/gad.1006702.
- Tcheressiz**, S., V. Calco, F. Arnaud, L. Arthaud, B. Dastugue, et C. Vaury. **2002**. « Expression of the Idefix Retrotransposon in Early Follicle Cells in the Germarium of *Drosophila Melanogaster* Is Determined by Its LTR Sequences and a Specific Genomic Context ». *Molecular Genetics and Genomics: MGG* 267 (2): 133- 41. doi:10.1007/s00438-002-0641-1.
- « **The Yeast Genome Directory** » **1997**. *Nature*. 387 (6632 Suppl): 5.

- Todeschini**, Anne-Laure, Laure Teyssset, Valérie Delmarre, et Stéphane Ronsseray. **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 (6): e11032. doi:10.1371/journal.pone.0011032.
- Tomari**, Yukihide, Tingting Du, Benjamin Haley, Dianne S. Schwarz, Ryan Bennett, Heather A. Cook, Birgit S. Koppetsch, William E. Theurkauf, et Phillip D. Zamore. **2004**. « RISC Assembly Defects in the *Drosophila* RNAi Mutant Armitage ». *Cell* 116 (6): 831 - 41.
- Vagin**, Vasily V., Mikhail S. Klenov, Alla I. Kalmykova, Anastasia D. Stolyarenko, Roman N. Kotelnikov, et Vladimir A. Gvozdev. **2004**. « The RNA Interference Proteins and Vasa Locus Are Involved in the Silencing of Retrotransposons in the Female Germline of *Drosophila Melanogaster* ». *RNA Biology* 1 (1): 54 - 58.
- Verdel**, André, Songtao Jia, Scott Gerber, Tomoyasu Sugiyama, Steven Gygi, Shiv I. S. Grewal, et Danesh Moazed. **2004**. « RNAi-Mediated Targeting of Heterochromatin by the RITS Complex ». *Science (New York, N.Y.)* 303 (5658): 672 - 76. doi:10.1126/science.1093686.
- Vermaak**, Danielle, Steven Henikoff, et Harmit S. Malik. **2005**. « Positive Selection Drives the Evolution of Rhino, a Member of the Heterochromatin Protein 1 Family in *Drosophila* ». *PLoS Genetics* 1 (1): 96 - 108. doi:10.1371/journal.pgen.0010009.
- Volpe**, Thomas A., Catherine Kidner, Ira M. Hall, Grace Teng, Shiv I. S. Grewal, et Robert A. Martienssen. **2002**. « Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi ». *Science (New York, N.Y.)* 297 (5588): 1833 - 37. doi:10.1126/science.1074973.
- Wang**, Sidney H., et Sarah C. R. Elgin. **2011**. « *Drosophila* Piwi Functions Downstream of piRNA Production Mediating a Chromatin-Based Transposon Silencing Mechanism in Female Germ Line ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (52): 21164 - 69. doi:10.1073/pnas.1107892109.
- Xiol**, Jordi, Pietro Spinelli, Maike A. Laussmann, David Homolka, Zhaolin Yang, Elisa Cora, Yohann Couté, et al. **2014**. « RNA Clamping by Vasa Assembles a piRNA Amplifier Complex on Transposon Transcripts ». *Cell* 157 (7): 1698 - 1711. doi:10.1016/j.cell.2014.05.018.
- Yan**, Zheng, Hai Yang Hu, Xi Jiang, Vera Maierhofer, Elena Neb, Liu He, Yuhui Hu, et al. **2011**. « Widespread Expression of piRNA-like Molecules in Somatic Tissues ». *Nucleic Acids Research* 39 (15): 6596 - 6607. doi:10.1093/nar/gkr298.
- Yin**, Hang, et Haifan Lin. **2007**. « An Epigenetic Activation Role of Piwi and a Piwi-Associated piRNA in *Drosophila Melanogaster* ». *Nature* 450 (7167): 304 - 8. doi:10.1038/nature06263.
- Yoon**, Jeongheon, Kyu-Sun Lee, Jung Sun Park, Kweon Yu, Sang-Gi Paik, et Yong-Kook Kang. **2008**. « dSETDB1 and SU(VAR)3-9 Sequentially Function during Germline-Stem Cell Differentiation in *Drosophila Melanogaster* ». *PloS One* 3 (5): e2234. doi:10.1371/journal.pone.0002234.

- Zamore, P. D., T. Tuschl, P. A. Sharp, et D. P. Bartel. 2000.** « RNAi: Double-Stranded RNA Directs the ATP-Dependent Cleavage of mRNA at 21 to 23 Nucleotide Intervals ». *Cell* 101 (1): 25- 33. doi:10.1016/S0092-8674(00)80620-0.
- Zamparini, Andrea L., Marie Y. Davis, Colin D. Malone, Eric Vieira, Jiri Zavadil, Ravi Sachidanandam, Gregory J. Hannon, et Ruth Lehmann. 2011.** « Vreteno, a Gonad-Specific Protein, Is Essential for Germline Development and Primary piRNA Biogenesis in *Drosophila* ». *Development (Cambridge, England)* 138 (18): 4039- 50. doi:10.1242/dev.069187.
- Zhang, Fan, Jie Wang, Jia Xu, Zhao Zhang, Birgit S. Koppetsch, Nadine Schultz, Thom Vreven, et al. 2012.** « UAP56 Couples piRNA Clusters to the Perinuclear Transposon Silencing Machinery ». *Cell* 151 (4): 871- 84. doi:10.1016/j.cell.2012.09.040.
- Zhang, Y., et D. Kalderon. 2001.** « Hedgehog Acts as a Somatic Stem Cell Factor in the *Drosophila* Ovary ». *Nature* 410 (6828): 599- 604. doi:10.1038/35069099.
- Zhang, Zhao, Birgit S. Koppetsch, Jie Wang, Cindy Tipping, Zhiping Weng, William E. Theurkauf, et Phillip D. Zamore. 2014.** « Antisense piRNA Amplification, but Not piRNA Production or Nuage Assembly, Requires the Tudor-Domain Protein Qin ». *The EMBO Journal* 33 (6): 536- 39. doi:10.1002/embj.201384895.
- Zhang, Zhao, Jie Wang, Nadine Schultz, Fan Zhang, Swapnil S. Parhad, Shikui Tu, Thom Vreven, Phillip D. Zamore, Zhiping Weng, et William E. Theurkauf. 2014.** « The HP1 Homolog Rhino Anchors a Nuclear Complex That Suppresses piRNA Precursor Splicing ». *Cell* 157 (6): 1353- 63. doi:10.1016/j.cell.2014.04.030.
- Zilberman, Daniel, Xiaofeng Cao, et Steven E. Jacobsen. 2003.** « ARGONAUTE4 Control of Locus-Specific siRNA Accumulation and DNA and Histone Methylation ». *Science (New York, N.Y.)* 299 (5607): 716- 19. doi:10.1126/science.1079695.

