



**HAL**  
open science

# De l'œuf à l'adulte : étude moléculaire et fonctionnelle de la répression des éléments transposables par les piARN au cours du développement chez drosophila melanogaster

Pauline Marie

► **To cite this version:**

Pauline Marie. De l'œuf à l'adulte : étude moléculaire et fonctionnelle de la répression des éléments transposables par les piARN au cours du développement chez drosophila melanogaster. Génétique. Université Pierre et Marie Curie - Paris VI, 2016. Français. NNT : 2016PA066426 . tel-01879095

**HAL Id: tel-01879095**

**<https://theses.hal.science/tel-01879095>**

Submitted on 22 Sep 2018

**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

École doctorale Complexité du vivant

*Laboratoire de biologie du développement / Épigénétique et ADN mobile*

## **De l'œuf à l'adulte : étude moléculaire et fonctionnelle de la répression des éléments transposables par les piARN au cours du développement chez *Drosophila melanogaster***

Par Pauline Marie

Thèse de doctorat de génétique

Dirigée par Stéphane Ronsseray et Antoine Boivin

Présentée et soutenue publiquement le 20 septembre 2016

Devant un jury composé de :

Pret Anne Marie	Professeur	Rapporteur
Simonelig Martine	Directrice de recherche	Rapporteur
Higuet Dominique	Professeur	Président du jury
Huynh Jean-René	Directeur de recherche	Examineur
Brasset Emilie	Maître de Conférences	Examineur
Boivin Antoine	Maître de Conférences	Co-encadrant
Ronsseray Stéphane	Directeur de Recherche	Directeur de thèse



## Remerciements

Tout d'abord, je souhaite remercier les membres du jury pour avoir accepté d'évaluer mon travail de thèse, en particulier les Dr. Anne-Marie Pret et Martine Simonelig qui ont bien voulu être rapportrices. Ces trois années ont été très enrichissantes et je me sens très chanceuse d'avoir pu les passer au sein de la communauté des Drosophilistes, la plus chaleureuse de toutes, de leur propre avis que je partage évidemment!

Merci Stéphane pour m'avoir accueillie dans ton équipe, initiée aux secrets de la Droso, de la génétique et bien sûr de l'élément *P* qui, comme chacun sait, a envahi le génome de *Drosophila melanogaster* au siècle dernier ! Merci pour les métaphores automobiles en réunion de labo qui m'ont toujours fait beaucoup rire, pour tes idées de manip toujours ingénieuses et pour ton soutien tout au long de cette aventure.

Je te remercie particulièrement Antoine pour m'avoir encadré durant trois ans sur ce projet un peu nouveau pour l'équipe qui a nécessité pas mal de mise au point et généré des gigas d'images en couleurs. J'ai appris plein de choses grâce à toi et notre rencontre en TP de M1 me semble déjà très loin! J'ai beaucoup apprécié de travailler avec toi dans une ambiance détendue et agréable, j'espère que la suite de ton histoire va bien se dérouler et que les reviewers seront enthousiastes! Un grand merci.

Il me semble que notre bureau a toujours été particulièrement bien (sur)peuplé et tous ses membres resteront d'excellents souvenirs.

Laure, merci de m'avoir réconcilié avec les petits ARN en master, je ne serais jamais arrivée là sans toi ! Merci pour ta bonne humeur et tes conseils et enfin pour l'exemple que tu donnes quant au métier de MCU que je compte bien viser d'ici... une dizaine d'années.

Val, j'ai adoré passer ces trois ans et demi en ta compagnie, j'ai l'impression que nous nous sommes trouvés des passions communes assez facilement : je n'avais jamais rencontré d'Austinienne aussi mordue que moi et ce fut un plaisir immense. Le fait que tu aies fait une petite incursion dans le monde zombie grâce à Jane est une petite fierté personnelle, j'avoue ! J'espère que nos échanges de livres et de films ne s'arrêteront pas là et que j'aurai toujours de tes nouvelles.

Anna, compagne de route, aussi bien spirituellement qu'au sens propre, merci d'avoir accompagné mes démarches hésitantes dans l'administration, mes délires dans la pièce à mouche et mes aventures dans le monde des congrès ! J'ai plein de souvenirs précieux grâce à toi et il faut bien avouer que l'ambiance de la pièce à mouche est bien moins survoltée depuis que tu es partie ! Un grand merci pour tout.

Karine, tu prends la relève. Je te souhaite bonne chance pour cette aventure qu'est la thèse. J'espère qu'elle se déroulera sans trop d'accros!

Je n'oublie pas nos voisins drosophilistes de la pièce à mouches : Frédérique, Delphine, Jérôme, Hélène, Emmanuèle, Jean-Michel, Sandra, Neel, Valérie et Immane.

En particulier Jérôme, bien sûr, mon autre compagnon d'aventure toujours là pour discuter, écouter mes doléances, me donner de bons conseils et m'accompagner en congrès au bout de la France ou au bout du monde. Évidemment, ce qui se passe à Key West... Delphine merci pour ta bonne humeur et pour l'animation des repas dans cette pièce qui est ton bureau! Je garderai un très bon souvenir de ces midis passés à manger des plats Picard, à papoter en partageant des expériences de vie avec Val, Valérie, Amna, Karine, Hélène et Sandra (et en choquant Jérôme au passage, le pauvre). Les mots croisés instaurés par Karine. Sandra, qui se révèle un peu plus cette année et qui me fait bien rire.

Merci Emmanuèle pour tes conseils, surtout vers la fin pour les écueils administratifs. Je te remercie particulièrement Neel pour ta disponibilité et tes conseils inappréciables quant à notre manuscrit! Mention spéciale pour Camille qui nous a abandonné mais qui reste dans nos cœurs... et qui m'a donné une excellente excuse pour découvrir Londres.

Merci aux anciens membres de l'équipe, Catherine, Georges d'avoir fait partie de mon expérience. Merci aux autres personnes de l'unité qui m'ont accompagnée durant ces trois ans, Margarita, Affaf, Sara, Camille C, Mafalda, les autres représentants des doctorants, Merci à tous les membres de l'UMR7622 et à ceux de l'IFR83 que j'ai beaucoup vu également, Susanne, Jean-François et Richard.

Je remercie mon compagnon Jonathan pour son soutien, sa compréhension et le partage des tâches ménagères ^^ . Merci pour les relectures et les conseils, les discussions science et l'appui que je ressens quand nous sommes confrontés à des non-scientifiques. On ne s'est pas beaucoup vu ces derniers mois, emplois du temps de thésard oblige, mais j'espère que nous allons rattraper ça au plus vite. Je te souhaite une dernière année de thèse mémorable et pleine de publications!

Merci aussi aux amis pour le bon temps passé entre les manip!

Et je remercie également toute ma famille pour son soutien, en particulier ma mère et ma sœur pour leur confiance inconditionnelle en moi, tous les autres pour leurs encouragements et je pense fort à celui qui n'est plus là car je sais qu'il aurait été très fier.

## Sommaire

Remerciements .....	3
Sommaire .....	5
Préambule.....	7
<b>I - Introduction</b> .....	9
Chapitre 1 : Les Éléments Transposables .....	9
A. Classification des éléments transposables .....	10
A.1 Familles d'éléments transposables .....	10
A.2 Stratégies de conquête du génome .....	12
B. Défendre le génome.....	15
B.1 Historique de la découverte des petits ARN interférents .....	15
B.2 La répression des éléments transposables .....	16
B.3 Différents petits ARN, même combat .....	17
B.4 Un nouveau système de régulation, les piARN.....	22
Chapitre 2 : Les Clusters Producteurs de piARN.....	29
A. Clusters endogènes producteurs de piARN.....	30
A.1 Uni-strand clusters.....	31
A.2 Transcription des dual-strand clusters .....	33
B. Le <i>Trans</i> -Silencing Effect .....	37
B.1 Les clusters de piARN transgéniques <i>T-1</i> et <i>BX2</i> .....	37
B.2 Transgènes « silencer » et transgènes « cible » .....	38
B.3 Propriétés et analyse du TSE .....	40
Chapitre 3 : La Biogenèse des piARN .....	45
A. Les clusters de piARN.....	46
A.1 Transcription des clusters de piARN.....	46
A.2 Le nuage .....	52
A.3 Retour au noyau : le choix de Piwi.....	59
A.4 Les piARN somatiques.....	61
B. Autres sources de piARN dans le génome .....	67
B.1 Les piARN géniques.....	67
B.2 Connexions entre les différentes voies de synthèse des petits ARN interférents ..	67
B.3 Impact de l'Exon Junction Complex sur la voie des piARN.....	70
Chapitre 4 : La Lignée Germinale, un Bastion à Protéger .....	73
A. De l'œuf à la mouche. ....	73
A.1 Mise en place des territoires embryonnaires .....	73
A.2 Développement des gonades .....	76
A.3 Stade adulte : l'ovogenèse .....	80
B. Maintien de la répression des ET au cours de la vie des cellules germinales .....	85
B.1 La protection du génome embryonnaire, une responsabilité maternelle. ....	85
B.2 Activation des clusters de piARN .....	86

<b>II – Résultats</b> .....	89
<b>Répression d'un transgène rapporteur</b> dans les cellules germinales en développement ....	
1. Résultats du crible .....	92
1.1 Lignées enhancer-trap <i>lacZ</i> .....	92
1.2 Lignées enhancer-trap <i>gal4</i> .....	93
2. Test des constructions pilotes <i>gal4</i> spécifiques de la lignée germinale adulte .....	94
<b>Article 1</b> : « From egg to adult: piRNA-mediated silencing throughout germline development in <i>Drosophila</i> » .....	93
Présentation de l'article .....	99
Résultats additionnels.....	133
Effet de la mutation de gènes impliqués dans la voie des piARN .....	133
Augmentation de la transcription de différents éléments transposables .....	139
Impact de la mutation germinale de Rhino et HP1a.....	140
Conclusions .....	143
<b>Persistance des piARN maternels</b> .....	145
1. Stabilité des piARN hérités maternellement .....	145
2. Effet et demi-vie des piARN maternels .....	145
2.1 Effet des piARN maternels sur une cible germinale.....	146
2.2 Effet des piARN maternels sur une cible somatique .....	149
<b>Article 2</b> «Knock down of <i>half pint</i> gives birth to a meiotically stable piRNA-producing locus in <i>Drosophila</i> » .....	147
Présentation de l'article .....	153
Résultats additionnels.....	187
Conclusions .....	189
<b>III – Discussion et Perspectives</b> .....	185
A. Les piARN larvaires différents des piARN adultes ?.....	191
A.1 Relâcher la pression ?.....	191
A.2 Utilisation du crible shARN .....	191
A.3 Quelle différence entre les transgènes silenciers <i>RS3</i> et <i>P-1152</i> ?.....	193
B. Le rôle des protéines Rhino et HP1a dans la production de piARN .....	195
B.1 Rôle de Rhino et de HP1a dans la transcription des clusters de piARN .....	195
B.2 Modèle proposé pour la variéation .....	196
B.3 Éviter les piARN .....	198
C. Perspectives : observer le mécanisme moléculaire de la Variéation .....	201
<b>IV – Matériel et Méthodes</b> .....	203
A. Souches utilisées.....	203
B. Détection d'une cible <i>lacZ</i> : coloration Xgal.....	204
C. Immunohistochimie .....	205
D. Tri cellulaire .....	207
E. Extraction d'ARN et Deep-Sequencing des petits ARN .....	208
F. RTqPCR.....	210
<b>V – Références Bibliographiques</b> .....	213
<b>VI – Annexe</b> .....	231
<b>Résumé</b> .....	233

## Préambule

Pour des raisons pratiques (et éthiques), la drosophile est un organisme modèle très utilisé dans les laboratoires de génétique. Sa relative simplicité, 15000 gènes répartis sur quatre paires de chromosomes séquencés et annotés (Adams et al. 2000) permettent l'étude de nombreux processus. Sa petite taille, son cycle de vie court et le grand nombre de ses descendants en font un puissant outil pour la recherche en génétique. Une vaste communauté de chercheurs l'a adoptée : il existe une base de données en ligne contenant de nombreuses informations telles que les profils de transcription des gènes, les phénotypes mutants, la description de différents allèles... De façon encore plus tangible, des banques de mouches auprès desquelles il est possible de se procurer les souches transgéniques créées dans les laboratoires du monde entier existent sur tous les continents.

Notre équipe s'intéresse à la répression des éléments transposables dans la lignée germinale femelle chez *Drosophila melanogaster*. Ces séquences d'ADN parasites, aussi appelées transposons, ont la capacité de s'intégrer dans le génome d'organismes hôtes. Elles peuvent représenter une importante proportion de l'ADN d'une espèce : 45% de l'ADN humain est composé de séquences de transposons. Une fois en place, ils génèrent des copies supplémentaires de façon anarchique, causant une instabilité génomique : leur mobilisation peut entraîner des cassures chromosomiques et leur intégration peut perturber des gènes essentiels. Les mutations ainsi causées peuvent entraîner la mort de la cellule concernée, voire même la stérilité de l'organisme hôte. En effet, le maintien de l'intégrité du génome est particulièrement important dans la lignée germinale qui transmet l'information génétique à travers les générations. Les organismes ont développé une sorte de système immunitaire basé sur de petits ARN non-codants pour limiter l'activité de ces « parasites génomiques » (Brennecke et al. 2007). La répression des éléments transposables est très étudiée au cours de la gamétogenèse chez l'adulte, où elle est particulièrement efficace, mais très peu d'informations sont connues à propos de cette régulation au cours du développement. Or c'est tout au long de la vie des cellules germinales qu'elle doit être maintenue pour que les générations futures héritent d'un génome en bon état. Mon projet de thèse est donc d'analyser les propriétés fonctionnelles et moléculaires de cette régulation à chaque étape du développement de la lignée germinale chez la drosophile. Dans ce manuscrit, je me concentrerai sur la description des interactions entre la lignée germinale femelle en formation et les éléments transposables.





# I - Introduction

## Chapitre 1 : Les Éléments Transposables

Les éléments transposables (ET) sont présents chez toutes les espèces en plus ou moins grandes quantités. Ainsi, ils représentent moins de 5% de l'ADN des nématodes (Kidwell et Lisch 2000), mais peuvent former jusqu'à 70% du génome de certaines espèces végétales comme le maïs (Biémont 2010). Il existe même des amphibiens tels que *Salamander hydromantes* dont le génome est composé à 90% de séquences d'éléments transposables (Marracci et al. 1996). Dans de telles proportions, il paraît évident qu'une trop forte mobilisation de ces éléments peut avoir des effets délétères sur l'organisme hôte : la nouvelle insertion pourrait interrompre un gène important (Finnegan 1992) ou interférer avec sa régulation (McDonald et al. 1997). La présence de plusieurs copies d'un transposon peut même induire, par recombinaison homologue, des réarrangements chromosomiques importants tels que des délétions et des duplications (Goldberg et al, 1983). L'effet délétère des transposons constitue un dogme qui semble valider les découvertes récentes sur les nombreuses stratégies de défense du génome mises en place par les différents organismes pour lutter contre les transposons. Néanmoins, de nombreuses exceptions ont été notées, comme des cas de « domestication » de protéines codées par un ET comme les transposases PiggyBac-like de *Paramecium tetraurelia* et *Tetrahymena thermophila* (Baudry et al. 2009; Cheng et al. 2010). En outre, des exemples d'autorégulation, et plus largement de coopération transposon/hôte ont été mis en évidence. Ainsi, les éléments transposables *HeT-A*, *TART* et *TAHRE* jouent le rôle de télomères chez *Drosophila melanogaster*. De fait, les ET ne sont plus seulement considérés comme de l'ADN « égoïste » mais aussi comme une force motrice pour l'évolution des génomes. Cette idée a été proposée pour la première fois en 1984 par Barbara McClintock lors de la remise de son prix Nobel. Un exemple connu de création d'un allèle est l'insertion d'un transposon dans le gène codant SBEI, l'enzyme responsable de la transformation de l'amylose en amylopectine chez le pois (Bhattacharyya et al, 1993). Cette insertion a donné lieu au phénotype de « graine ridée » observé par Gregor Mendel en 1865.

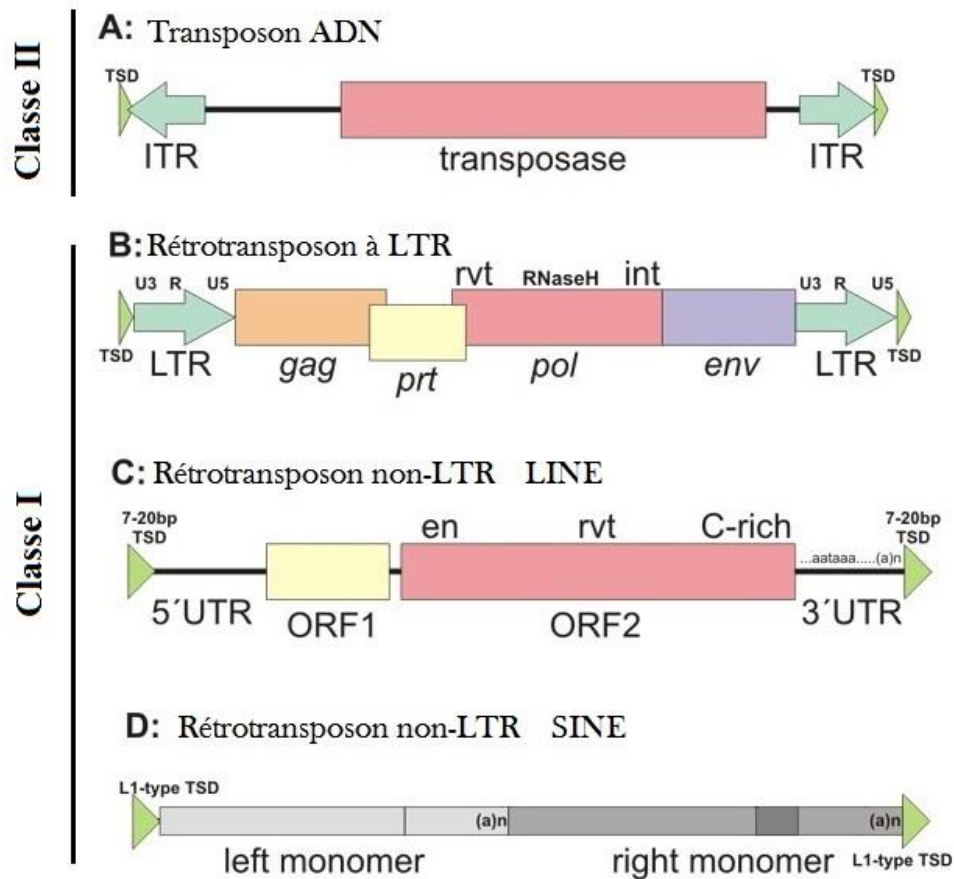
## **A. Classification des éléments transposables**

Il existe de nombreuses familles d'éléments transposables insérés dans les génomes, depuis les organismes unicellulaires jusqu'aux plantes et animaux. Depuis 1989, ces séquences mobiles ont été classifiées en deux grandes catégories en fonction de leur structure et de leur molécule intermédiaire de transposition (Finnegan 1992). Ainsi, les rétrotransposons (classe I) produisent un intermédiaire ARN qui peut être recopié à de nombreuses reprises. Au contraire, les transposons (classe II) changent de localisation *via* une molécule d'ADN excisée du génome puis déplacée.

### **A.1 Familles d'éléments transposables**

Les éléments transposables de classe II fonctionnent grâce à un mécanisme de « couper-coller ». Ces transposons ont généralement une structure de base constante : des séquences « Inverted Terminal Repeat » (ITR) situées de part et d'autre d'une séquence codant une transposase (Fig. 1A). Lorsque l'élément est transcrit, l'ARNm de la transposase est traduit dans le cytoplasme puis l'enzyme retourne dans le noyau et coupe l'ADN au niveau des ITR. Après excision, la molécule d'ADN est réintégrée ailleurs dans le génome.

Les rétrotransposons sont divisés en deux grandes familles : ceux qui possèdent des « Long Terminal Repeat » (LTR), avec une structure comparable à celle des rétrovirus (Fig. 1B) et ceux qui n'en possèdent pas, les rétrotransposons non-LTR comme *HeT-A* et *TART* chez *D. melanogaster*. Ces deux familles d'éléments se déplacent grâce à un mécanisme de « copier-coller » leur permettant d'augmenter leur nombre de copie mais leurs mécanismes de fonctionnement sont différents.



**Figure 1 - Familles de transposons.** A : Transposon à ADN, les séquences inversées répétées (ITR en vert jade) flanquent la séquence d'une transposase, elles-mêmes encadrées par les séquences de «Target site duplication» (TSD en vert pomme) qui sont laissées dans le génome lorsque le transposon est mobilisé. B : Rétrotransposon à LTR, constitués généralement de séquences ressemblant à celles des rétrovirus, *gag*, *pol* et *env*. Les longues répétitions terminales qui entourent ces séquences codantes sont les LTR (en vert jade) constitués d'U3 (3' untranslated region), R (recombination region) et U5 (5' untranslated region). C : Rétrotransposon non-LTR type LINE (Long Interspersed Element) constitué de 2 cadres ouverts de lecture qui codent une endonucléase (*en*) et une transcriptase inverse (*rvt*). Le 5' UTR possède un promoteur interne ARN pol II tandis qu'un signal de polyadénylation est présent dans le 3'UTR, suivi d'une queue polyA. D : Rétrotransposon non-LTR type « Short Interspersed Nuclear Element » (SINE), un promoteur interne pour l'ARN pol III permet l'expression des régions internes L et R. (Adapté de « Current genetics - Repetitive DNA » 2016).

Les rétrotransposons produisent un ARN qui sera reverse transcrit en ADN complémentaire (ADNc) pour s'insérer à un nouveau site. Ces insertions possèdent une signature particulière : on retrouve à leurs extrémités un «Target site duplication» (TSD) de 4 à 8 paires de bases. La comparaison entre les éléments à LTR et les rétrovirus ne s'arrête pas aux LTR. En effet, les ET comme *gypsy* chez *D. melanogaster* codent des protéines comparables aux « Group specific AntiGen » (*gag*, Fig 1.B en orange), et la protéase (Prt) est

nécessaire à l'assemblage des rétrovirus. *gag* et *pol* peuvent être traduits en une seule grande protéine, *pol* qui contient un domaine reverse transcriptase (rvt), un domaine RNaseH et un domaine intégrase (int). (Fig. 1B). Les protéines GAG peuvent former une particule virale, sorte de compartiment où a lieu la reverse transcription (comme le rétrovirus de levure Ty (AL-Khayat et al. 1999). L'intermédiaire ADNc est transporté jusqu'au noyau où il est réintégré dans le génome grâce à l'intégrase codée par *pol*. Parmi les rétroéléments à LTR on distingue les Pseudoviridae (*copia*), les Metaviridae (*gypsy*) et les rétrovirus endogènes (Benachenhou et al. 2013).

Les rétrotransposons non LTR sont eux aussi divisés en deux familles, les LINE (Fig. 1C) et les SINE (Fig. 1D). Les LINE possèdent deux cadres ouverts de lecture (« Open Reading Frame », ORF) (1 et 2) dont au moins un code une transcriptase inverse permettant la transposition. Ces protéines forment un complexe ribonucleo-protéique (RNP) avec l'ARN messager du LINE qui retourne dans le noyau où la reverse transcription a lieu. L'ADNc est alors intégré dans le génome. Les LINE sont également caractérisés par la présence d'une région riche en A à leur extrémité 3'. Les SINE possèdent quant à eux un promoteur ARN pol III permettant l'expression de leurs régions internes L et R. Ils utilisent la transcriptase inverse codée par des éléments LINE pour leur rétro-transcription.

## **A.2 Stratégies de conquête du génome**

Les transposons de classe I et II diffèrent par leur structure, leur origine évolutive et leurs modalités d'expression à la fois temporelle et spatiale. Ils peuvent être transcrits à certains stades et/ou uniquement dans certains tissus spécifiques. La diversité des stratégies de transposition et la relative similitude entre l'expression de ces « parasites » et celle des gènes codants pour des protéines est un obstacle important pour les génomes. Une régulation de la transcription de ces éléments est nécessaire pour protéger l'information génétique, en particulier au niveau de la lignée germinale. La comparaison de données de séquençages haut débit de petits ARN provenant d'ovaires adulte et de jeunes embryons a permis de distinguer les transposons exprimés spécifiquement dans la lignée germinale de ceux transcrits dans les cellules somatiques de l'ovaire (Malone et al. 2009). Trois catégories d'éléments sont décrites : les transposons dits « somatiques ovariens » comme *gypsy*, *ZAM* et *tabor* qui ne s'expriment que dans les cellules somatiques folliculaires de l'ovaire de drosophile qui entourent la lignée germinale. Les transposons « germinaux » comme l'élément *P* et *copia* s'expriment uniquement dans les cellules germinales. Enfin, certains dont *tirant*, *blood* et *412*

semblent s'exprimer dans ces deux tissus. D'autres études ont mis en évidence que des éléments transposables peuvent s'exprimer dans d'autres tissus somatiques, comme *LINE-1* dans cerveau (Upton et al. 2015). Durant ma thèse je me suis particulièrement intéressée à la régulation des ET dans la lignée germinale.

### ***A.2.1 Éléments transposables germinaux***

Le tissu d'expression de nombreux ET est la lignée germinale, où les effets de leur mobilisation et l'accroissement du nombre de copies seront transmises à la descendance. Par exemple chez la drosophile, l'élément *P*, un transposon à ADN, est actif uniquement en tissu germinale. L'expression de ce dernier est régulée génétiquement et de façon tissu-spécifique. Les éléments *P* autonomes portent un gène codant une transposase nécessaire à leur mobilisation (W. R. Engels 1984). La production de celle-ci est uniquement induite dans la lignée germinale grâce à l'épissage spécifique de l'intron 2-3 de *P* dans ce tissu (Laski et al, 1986). Un autre exemple, l'élément *I*, un rétroélément non-LTR, est transcrit sous le contrôle d'un promoteur interne activé spécifiquement dans les cellules de la lignée germinale femelle (Chambeyron et Bucheton 2005). Enfin, les rétroéléments non-LTR, *HetT-A* et *TART* sont également exprimés dans la lignée germinale. Ils sont retrouvés aux télomères dans la famille *Drosophila* et jouent le rôle de la télomérase. La symbiose établie entre ces rétrotransposons et le génome de *Drosophila* rappelle l'idée de McClintock que les éléments transposables ont des rôles régulateurs dans l'organisme. La mobilisation de ces éléments transposables est finement régulée grâce au mécanisme d'ARN interférence que j'introduirai dans le chapitre suivant. Les cellules germinales présentent le mécanisme de régulation le plus restrictif.

### ***A.2.2 Éléments transposables somatiques folliculaires***

Chez les mouches, les cellules folliculaires somatiques entourent chaque ovocyte en développement, fournissant des signaux moléculaires participant à l'établissement des axes antéro-postérieur et dorso-ventral du futur embryon. Elles protègent ainsi l'œuf durant sa maturation. Bien que la plupart des transposons agissent dans la lignée germinale, *gypsy*, un rétroélément à LTR, échappe à la régulation germinale en produisant des particules de type viral dans les cellules folliculaires. Il infecte ensuite l'ovocyte adjacent grâce à l'export de son ARN à travers les membranes qui séparent les cellules germinales des cellules folliculaires (A. Kim et al. 1994; S. U. Song et al. 1994) probablement grâce à la présence de *env* (Teyssset et al. 1998). D'autres éléments transposables à LTR comme *Zam* et *Gypsy* sont contrôlés dans les cellules folliculaires grâce à des séquences régulatrices particulières, le *locus flamenco* par

exemple, dont l'expression est également limitée à ce tissu (Desset et al. 2003, Pélisson et al. 2007).

### ***A.2.3 Éléments transposables somatiques***

D'autres tissus somatiques sont sujets à la mobilisation d'éléments transposables. Le laboratoire de G. Faulkner a effectué des séquençages du génome de neurones de l'hippocampe, du cortex et sur des cellules gliales. Ils trouvent environ 14 nouvelles insertions de *LINE-1* par cellules dans ces tissus. Les neurones de l'hippocampe sont particulièrement ciblés : les enhancers et gènes transcrits dans ce tissu présentent souvent une insertion. Ceci est probablement dû à des raisons mécaniques : les gènes exprimés sont peut-être plus accessibles à cause de l'ouverture de la chromatine pour la transcription. Les insertions introniques dans le sens de la transcription sont majoritaires. G. Faulkner et ses collaborateurs vont même jusqu'à poser l'hypothèse d'une éventuelle fonction des insertions de ce rétroélément dans le cerveau *via* la diversité génétique qu'elles entraînent (Upton et al. 2015).

## **B. Défendre le génome, un système basé sur les petits ARN et les marques chromatiniques**

Les éléments transposables jouent un rôle majeur dans l'évolution tout en représentant une source d'instabilité génétique. Leur transposition peut entraîner, par exemple, la perturbation de gènes fonctionnels *via* l'altération de leur expression ou un épissage aberrant. L'augmentation du nombre de copies d'un ET fournit la matrice pour des recombinaisons pouvant donner lieu à des duplications de gènes, des délétions ou des réarrangements à grande échelle. Les espèces qui assurent la conservation de l'intégrité de leurs génomes sur de larges échelles de temps ont établi des mécanismes de défense moléculaires qui emploient des « RNA-induced silencing complexes » (RISC), pour protéger leurs génomes en limitant la mobilisation et multiplication des ET. Les RISC sont composés de petits ARN non codants complexés à des protéines effectrices Argonautes. Les différentes voies d'interférence ARN (ARNi) peuvent être classifiées en fonction du type de petit ARN non codant et de la protéine Argonaute associée.

### **B.1 Historique de la découverte des petits ARN interférents**

#### ***B.1.1 L'interférence ARN***

En 1998 Fire et Mello découvrent l'efficacité de l'injection de doubles brins d'ARN (dsARN) qui répriment l'expression de transcrits ARN : les ARNm de *mex-3* diminuent en présence du dsARN correspondant (Fire et al. 1998). En revanche, les dsARN correspondant à des promoteurs ou des introns ne conduisent pas à l'extinction du gène en question. L'effet de ces dsARN est capable de traverser les membranes cellulaires et les générations : peu importe l'endroit de l'injection dans le parent, la génération suivante montre une répression des gènes correspondants parfois même plus importante que le phénotype du parent. Ces expériences ont poussé la génétique vers une nouvelle ère, cette technique permettant de mimer des mutants pour n'importe quelle protéine dont la séquence du message est connue et donc de tester les phénotypes de régions codantes dont les produits n'étaient pas connus. Néanmoins, Fire et Mello avaient émis quelques réserves : l'homologie de séquence entre des ARNm proches, des gènes de la même famille par exemple, leur permet d'être ciblés par un même dsARN (off-target effet) sans qu'on puisse les discriminer. Plus généralement, ils prédisaient que certains messages pourraient échapper à l'interférence et qu'un niveau d'expression résiduel perdurerait. Enfin, dans la discussion de leur article de 1998, Fire et



Mello émettent l'hypothèse d'une régulation des gènes par l'organisme en mettant en avant que si les dsARN peuvent voyager aussi facilement chez le nématode, il doit exister un mécanisme actif de transport des ARN interférents, et donc une fonction physiologique de cette répression. Aujourd'hui, l'interférence ARN (ARNi) au sens large, constitue un ensemble de procédés qui utilise des petits ARN (de 20 à 30 nucléotides) capables de reconnaître et d'interférer avec l'expression d'ARN possédant une séquence complémentaire : on distingue trois grandes catégories: les miARN, les siARN et les piARN.

### ***B.1.2 Le complexe RISC***

L'interférence ARN est réalisée par le complexe RISC, constitué d'une protéine Argonaute, responsable du clivage de l'ARNm, associée à un petit ARN. Ce dernier guide le complexe au site de clivage, reconnu par complémentarité de séquence. Les protéines Argonautes possèdent un domaine PAZ (pour Piwi-Argonaute-Zwille), un domaine intermédiaire MID et un domaine PIWI. Le petit ARN va se fixer entre les domaines PAZ et PIWI et la structure formée par le domaine MID est essentiel au clivage auquel participe le domaine PIWI, similaire à un domaine RNase H (J.-J. Song et al. 2004; W. Wang et al. 2015). Chez les eucaryotes, la famille Argonaute est divisée en deux. Les Ago-like apparentées à AGO1 de *A. thaliana* sont exprimées dans tous les tissus et prennent en charge les miARN et les siARN. Les PIWI-like, homologues de la protéine Piwi de *D. melanogaster*, sont plus exprimées dans les tissus germinaux, voire spécifiques de ces tissus, et sont associées aux piARN (Vagin et al. 2006).

## **B.2 La répression des éléments transposables**

Les petits ARN sont impliqués dans deux modes de répression des éléments transposables. D'une part, une régulation transcriptionnelle au niveau de la chromatine par l'établissement de marques chromatiniennes répressives permet d'empêcher ceux-ci d'être transcrits, c'est le « Transcriptional Gene Silencing » (TGS). Des petits ARN chargés par certaines protéines Argonautes sont capables de retourner dans le noyau et ciblent les copies euchromatiques actives des ET : ce mécanisme de répression transcriptionnelle appelé hétérochromatinisation est une façon de réguler les éléments transposables en rendant leur séquence moins accessible à la transcription. D'autre part, le « Post-Transcriptional Gene Silencing » (PTGS) a lieu dans le cytoplasme si les ET sont transcrits. Lorsqu'un ARNm produit par un ET traverse la membrane nucléaire, il est reconnu par un petit ARN complémentaire chargé par une protéine Argonaute qui va le cliver, entraînant sa dégradation.

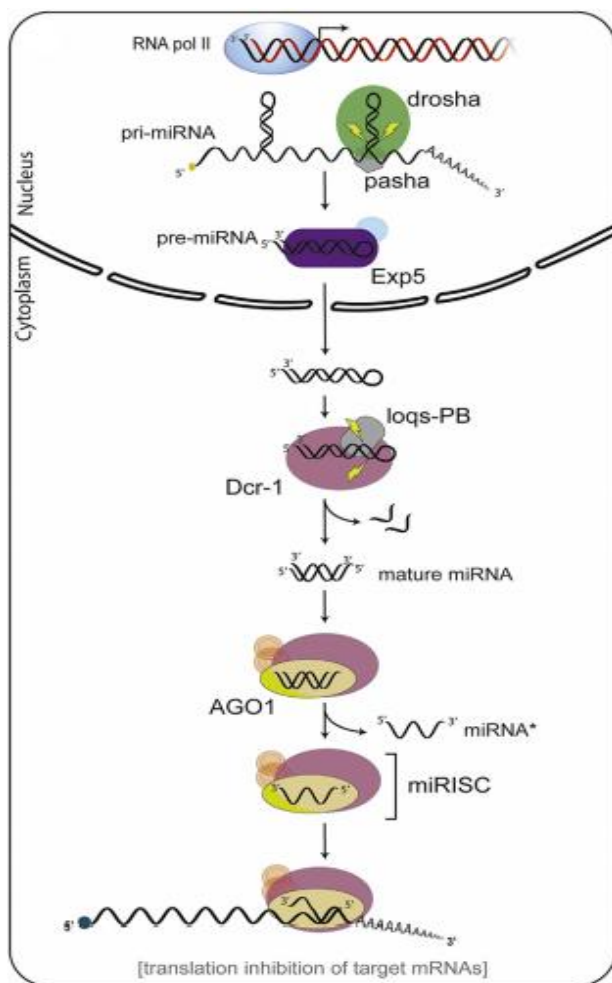
Ces deux mécanismes sont complémentaires quoi qu'on ne connaisse pas l'importance respective de chacun d'entre eux dans la répression. On peut néanmoins postuler que si un ARNm échappe au TGS dans le noyau, sa prise en charge par le PTGS dans le cytoplasme est une deuxième opportunité de répression.

### **B.3 Différents petits ARN, même combat**

Si les éléments transposables sont présents dans presque toutes les espèces vivantes, ils ne sont pas toujours régulés de la même façon. Il a été montré que les trois types de petits ARN interférents peuvent être impliqués dans cette répression.

#### ***B.3.1 Les miARN impliqués dans la régulation des éléments transposables ?***

Les miARN (micro-ARN) font majoritairement 22nt de long, ils sont dérivés de structures tiges boucles qui proviennent de régions génomiques contenant des répétitions inversées ou des paires de transcrits complémentaires qui peuvent donner lieu à des duplexes d'ARN (Fig. 2). On les retrouve généralement des régions inter-géniques, dans les ARN pré-messagers au niveau des introns et au niveau des séquences 3'UTR. Chez les animaux ils sont d'abord pris en charge par Drosha pour former le pré-miARN puis par Dicer-1 et Loquacious (Y. S. Lee et al. 2004). Le miARN mature sera associé majoritairement avec AGO1 chez la Drosophile pour cibler des ARNm *via* une complémentarité de séquence imparfaite avec leur 3'UTR (pour revue Bartel 2004). Une partie des miARN semble dérivée de transposons: Y. Li et collaborateurs montrent que chez les plantes, de nombreux miARN proviendraient en particulier de MITE (Miniature Inverted-repeat Transposable Elements) (Li et al. 2011). En revanche, ils ne seraient pas forcément impliqués dans leur régulation car ils ciblent généralement des gènes indépendants.



**Figure 2 - La voie miARN chez *D. melanogaster*.** Les gènes contenant des miARN sont généralement transcrits par l'ARN pol II, les précurseurs simples brins appelés pri-miARN sont reconnus par le complexe Drosha/Pasha grâce à la structure secondaire en épingle à cheveux du pri-miARN. Cette structure sera prise en charge dans le noyau pour générer le pré-miARN qui sera exporté vers le cytoplasme par l'exportin-5 où ils sont processés par Dcr-1 associé à Loqs-PB pour générer un duplex de miARN matures. Ce duplex va être pris en charge par AGO1 qui libère le brin passager également appelé miARN star (\*) alors que le brin guide reste fixé à AGO1 pour former le miRISC. Ce complexe va rechercher des sites partiellement complémentaires dans le 3'UTR d'ARN messagers et conduire à la déstabilisation de leur traduction (Soares et al. 2014).

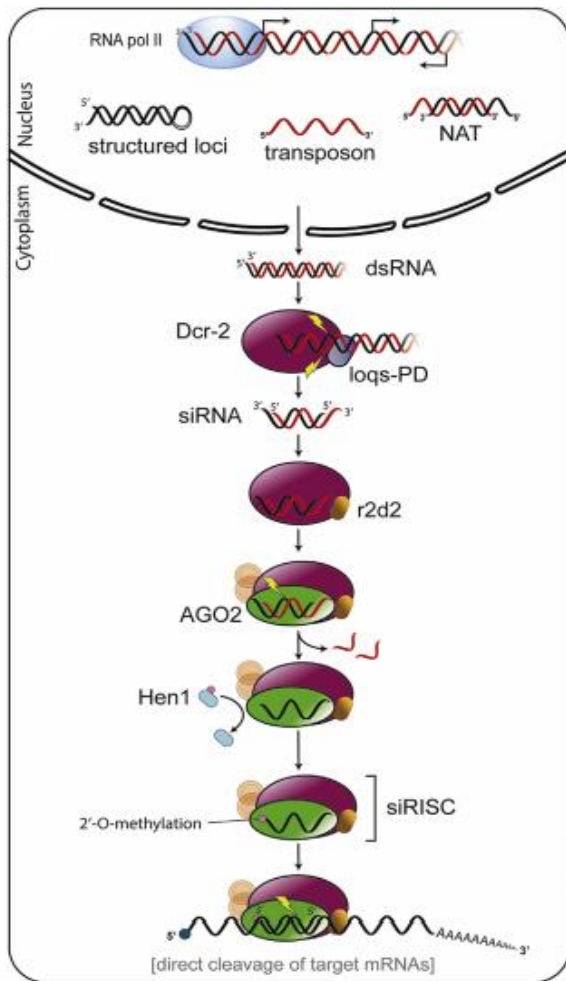
La mutation de Drosha/Dicer-2 dans les cellules folliculaires a montré que l'abolition de la voie miARN dans les cellules folliculaires entraîne l'expression d'un rapporteur *ZAM-lacZ* (fusion transcriptionnelle du promoteur de l'élément transposable ZAM avec le gène *lacZ* d'*E. coli*) ce qui reflète la dérégulation des transposons dans ces cellules chez la drosophile (Mugat et al. 2015). La mutation d'AGO1 a aussi été testée avec le même résultat. Les protéines impliquées dans la fabrication des miARN semblent nécessaires à la répression des transposons en cellules somatiques. Les auteurs ont voulu vérifier si cette perte de répression des ET était due à l'abolition du TGS (généralement majoritaire dans le soma) ou du PTGS. Des expériences d'immunoprécipitation de la chromatine (ChIP-qPCR) ont montré que la perte de Drosha entraîne sur le transgène *ZAM-lacZ* l'accumulation des marques activatrices de la transcription : la diméthylation de l'histone H3 au niveau de la Lysine 4 (H3K4me2) et la perte de la marque répressive H3K9me3 qui correspond à la triméthylation de l'histone H3 au niveau de la Lysine 9. Le TGS est donc impacté par la perte des miARN. D'ailleurs, la perte de deux miARN spécifiques est suffisante pour la réexpression d'un rapporteur *gypsy-lacZ* et de trois ET endogènes somatiques : *ZAM*, *tabor* et *Stalker2*. Cette

perte de Droscha est corrélée avec la perte d'une autre famille de petits ARN interférents, les piARN. En effet, en absence de miR-14 et miR-34 les piARN homologues à *ZAM* et *tabor* disparaissent (Mugat et al. 2015). Enfin, un troisième miARN influençant la production de piARN a été trouvé chez la souris: le miR-17v (Du et al. 2016). Son absence aurait un effet négatif sur la fertilité et sur le développement embryonnaire avec une diminution de la production des piARN et une hausse de la transposition des ET. Ces effets montrent que ces deux voies sont interconnectées, des miARN pouvant réguler la production de certaines protéines de la voie des piARN par exemple.

### ***B.3.2 siARN et marques chromatiniennes répressives***

Les petits ARN interférents, siARN, (Short Interfering RNA) font généralement 21nt de long et proviennent majoritairement d'ARN viraux double brins, pris en charge par Dicer-2, Loquacious (Loqs) et R2D2. Ils s'associent ensuite avec AGO2 pour interférer avec les transcrits viraux au niveau post-transcriptionnel (Okamura et al. 2004; S.-W. Ding et Voinnet 2007). Néanmoins dans le soma, une partie des siARN est principalement destinée à réguler l'activité de transposons. On les appelle les endo-siARN car ils dépendent de Dicer-2 pour leur production, comme les siARN, et d'AGO2 pour leur action (Péllisson et al. 2007; Ghildiyal et al. 2008; Chung et al. 2008) mais ils sont produits de façon endogène dans les cellules somatiques à partir de *loci* pouvant être transcrits dans les deux sens (dsARN NAT Fig. 3). Les endo-siARN peuvent agir post-transcriptionnellement. De plus, des expériences de CHIP ont montré que les séquences répétées des mutants *Dicer-2* présentent moins de marques chromatiniennes répressives de la transcription, en particulier H3K9me3, ceci laisse à penser que les siARN peuvent également induire du TGS.

Chez les plantes il existe de nombreuses protéines des familles Argonaute (AGO) et Dicer : 10 protéines AGO chez *A. thaliana* et 4 Dicer (Baulcombe 2004). Le TGS prédomine : la plupart des ET sont hétérochromatinisés. Ils sont transcrits par l'ARN polymérase IV (Pol IV) spécifique des plantes, leur ARNm est copié par RdR2 (ARN polymérase ARN dépendante) ce qui permet la production d'ARN double brins qui sera coupé par une protéine de la famille Dicer en petits ARN interférents de 24nt. Ce sont les protéines Dicer-Like 3 (Dcl3) qui sont majoritairement impliquées dans la production de ces siARN spécifiques des éléments transposables (Qi et al. 2005), même si les Dicer-Like de *A. thaliana* semblent avoir des fonctions redondantes (Ito et al. 2012).

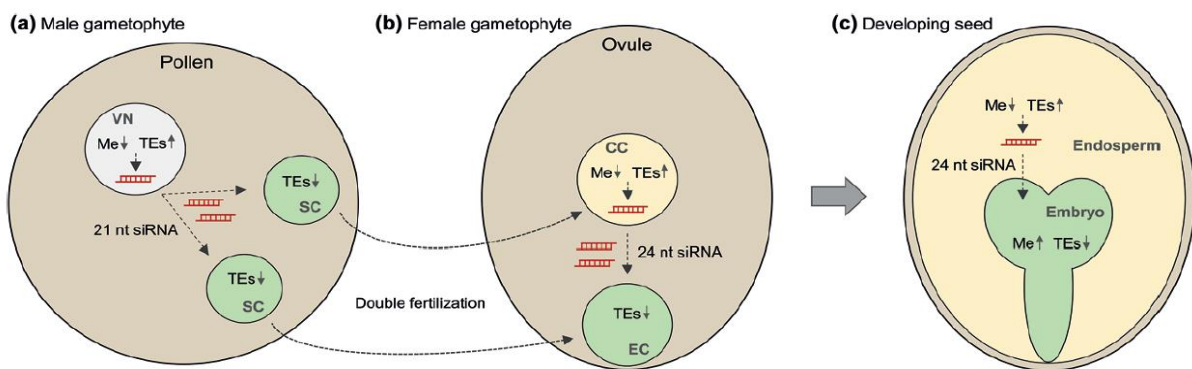


**Figure 3 - La voie siARN chez *D. melanogaster*.** Les ARN double brins provenant de la transcription de *loci* structurés par deux répétitions sens-antisens, de paires naturelles de transcrits complémentaires (NAT) et de séquences répétées comme les éléments transposables sont exportés vers le cytoplasme où ils sont reconnus et transformés en siARN par le complexe Dcr-2 /Loqs-PD. Les longs duplexes de siARN sont reconnus par Dcr-2 et R2D2, maturés puis chargés par AGO2. Le brin passager du duplex siARN est clivé par Ago2 qui va rester associé au brin guide. Dans AGO2, le brin guide est 2'O-méthylé par Hen1 pour donner le siRISC mature. La reconnaissance de séquences parfaitement complémentaires d'ARN messagers par le siRISC entraînera leur dégradation *via* un clivage par AGO2 (Soares et al. 2014).

Bien que la plupart des transposons soient inactifs dans les plantes hôtes, certains chocs génomiques, comme un stress environnemental ou un événement d'hybridation, peuvent déclencher l'activation des transposons. H. Ito a étudié la régulation du rétroélément *copialike* ONSEN en contexte de stress thermique : la mobilisation de cet élément est observée à 37°C et la lignée germinale n'est pas épargnée (Ito 2012). Dans un premier temps il a été observé qu'ONSEN n'est plus régulé dans les mutants pour NRPd2 (ARN polymérase) et RdR2, mais qu'une mutation de Ddl3 seule n'a pas d'effet: seul le triple mutant Dcl2/Dcl3/Dcl4 montre une dérégulation d'ONSEN. Les siARN sont ensuite chargés par AGO4 (Argonaute 4) et guident une méthyltransférase (RdDM) vers les *loci* transcrits par l'ARN polIV et ainsi renforcent la boucle de répression des ET. Des siARN de 21-22 nt sont également produits lors de l'invasion par un nouveau transposon. Celui-ci étant transcrit par l'ARN PolIII, le mécanisme de PTGS est nécessaire dans les premiers temps.

La régulation de la mobilisation des ET dans la lignée germinale des plantes dépend des siARN : les cellules végétatives et centrales qui accompagnent respectivement l'ovule et

le pollen subissent une déméthylation complète de leur ADN *via* DEMETER (Ibarra et al. 2012). Les éléments transposables sont alors réactivés et peuvent se mobiliser dans ces cellules. Ceci induit la production de siARN de 21nt dans le noyau végétatif qui seront transférés à la cellule germinale (SC Fig. 4). AGO9 s'associe à des siARN spécifiques de 24nt produits dans la cellule centrale et réprime les ET dans la cellule germinale avant la fécondation. Enfin, les petits ARN de 21 et 24nt induisent une hyperméthylation des séquences ADN cibles et « verrouillent » les ET de l'embryon.



**Figure 4 - Répression dans la lignée germinale des plantes.** Une déméthylation globale de l'ADN des cellules « compagnes » et de l'endosperme conduit à la réactivation de ET et à la production de siARN de 21 nt. Ces siARN sont transférés dans les gamètes et dans l'embryon pour renforcer le silencing transcriptionnel des éléments transposables. Abréviations: VN, noyau végétatif; SC, cellule de sperme; CC, cellule centrale; Me, méthylation de l'ADN (Gebert et Rosenkranz 2015).

### B.3.3 Des siARN spécifiques de la lignée germinale, les rasiARN

Chez *D. melanogaster* les « Repeat-Associated siARN » (rasiARN) font entre 23 et 30nt, et sont spécialisés dans la répression germinale des transposons. Contrairement aux siARN la majorité de ces petits ARN sont antisens aux transposons endogènes et correspondent à des régions répétées du génome, comme les 50 copies de *Suppressor of Stellate* sur le chromosome X. Leur présence est indépendante des protéines responsables de la production des siARN Dcr-2/R2D2/Ago2, ou des miARN Dcr-1/Ago1. Ce sont les Argonautes de la famille PIWI qui sont en charge de leur action. Aubergine et Piwi ont été précipitées en association avec ces petits ARN et des mutations d'Armitage et Spindle-E (hélicases) entraînent la perte de ces rasiARN ainsi que l'augmentation de la transcription de plusieurs ET (Vagin et al. 2006). La perte des rasiARN pendant la formation de l'ovocyte équivaut à la non-spécification de l'axe antéro-postérieur de l'embryon et à la désorganisation du réseau de microtubules entraînant une perte de localisation d'ARNm (*oskar*) et de protéines

(Gurken). Mais en utilisant des combinaisons de doubles mutants *aubergine* ou *armitage* avec *mnk* et *mei-41* ces défauts sont évités (Klattenhoff et al. 2007). *mnk* et *mei-41* codent pour les homologues de ATM et Chk2 chez la drosophile. Ce qui signifie que l'absence de rasiARN activerait les points de contrôle de dommage à l'ADN entraînant à leur tour des défauts de développement. Chez *D. melanogaster*, la forme phosphorylée du variant de l'histone H2A (y-H2Av) s'accumule aux points de cassures de l'ADN. Des foyers de y-H2Av sont effectivement retrouvés en grande quantité dans les mutants *armi* et *aubergine* et ne sont pas éliminés en présence de mutation *mnk* et *mei-41* (Klattenhoff et al. 2007). En 2007, le laboratoire de T. Schüpbach découvre deux nucléases impliquées dans la production des rasiARN : Zucchini (Zuc) et Squash (Squ) (Pane et al, 2007). Ils font l'hypothèse que les défauts d'embryogenèse détectés dans les mutants *zuc* et *squ* sont dus, au moins en partie, à une activation des points de contrôle de dommage à l'ADN comme pour les mutants de *aub* et *armi*. De plus, ils montrent que comme Aubergine et Armi, les protéines Zuc and Squ sont localisées dans le nuage, une structure électroniquement dense qui entoure le noyau des cellules germinales animales (Mahowald 1971). La co-immunoprécipitation de ces deux nucléases dans des lignées Aub-GFP/HA-Zuc et Aub-GFP/HA-Squ montre une interaction physique entre elles. Enfin, les mutants *Zuc* et *Squ* ne régulent plus les transposons *HeT-A* et *TART* ni le *locus Stellate* : les rasiARN correspondants à ces éléments ne sont plus produits (Pane et al, 2007). En résumé, ces rasiARN proviennent de séquences répétées, sont maturés par Zucchini et Squash aidés d'Armitage et de Spindle-E et associés avec Piwi et Aubergine au niveau du nuage pour réprimer les ARNm d'ET et les *loci* répétés dans la lignée germinale. Quelques années plus tard ils seront requalifiés de « PIWI- interacting RNA », les piARN, dont la biogenèse est détaillée dans le Chapitre 2.

#### **B.4 Un nouveau système de régulation, les piARN**

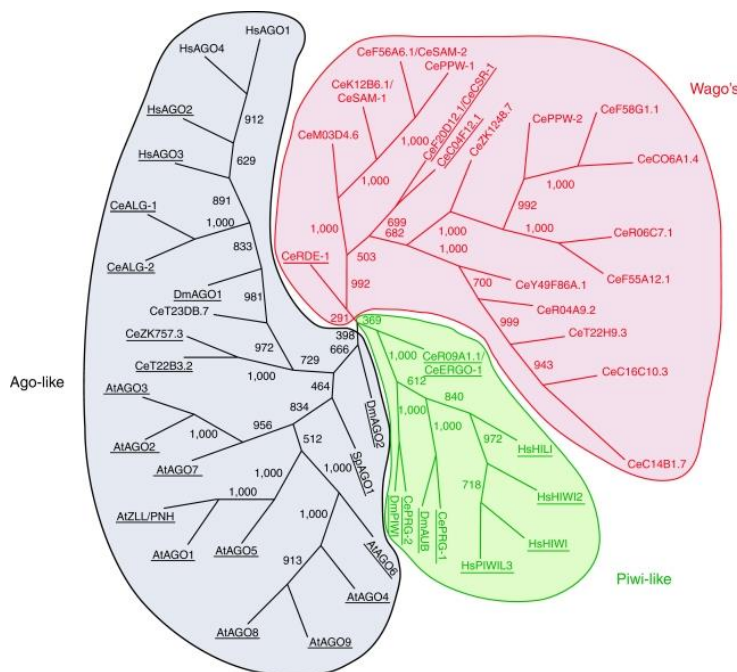
Dans les cellules germinales, de petits ARN sens et antisens homologues à des transposons sont présents en grande quantité. En 2006, les piARN sont découverts simultanément par quatre équipes, dirigées par T. Tuschl, H. Imai, H. Lin et G. Hannon, dans la lignée germinale de la souris en regardant les petits ARN produits par ces cellules. Le laboratoire d'Haifan Lin fait le lien en 2007 entre les rasiARN et les piARN, il caractérise 12903 piARN chez *D. melanogaster* et montre que les rasiARN sont une sous-catégorie de piARN (Yin et Lin 2007). À cette époque, même s'ils ne semblent pas provenir de longs ARN double brins, les chercheurs pensent que les piARN partagent plusieurs caractéristiques avec



les autres petits ARN interférents. L'écart de distribution de leur taille est relativement faible (23-31 nt), ils sont associés avec les protéines Argonautes connues pour être les actrices principales de l'ARN interférence et ils ont une forte tendance à commencer par un U en 5', ce qui est une particularité des petits ARN interférents. Enfin, ils sont 2'O-méthylés par dmHen1/Pimet comme les siARN ce qui les protège de l'ubiquitination et de la dégradation par l'exosome (Kirino et Mourelatos 2007).

#### B.4.1 Apparition évolutive

L'évolution d'organismes capables de se reproduire a été suivie par l'émergence rapide de parasites génomiques capables de profiter de cette capacité. Les hôtes en mesure de contrôler ces séquences parasites grâce au développement d'une panoplie de stratégies « immunitaires » pour discriminer le soi et le non-soi au niveau de l'ADN ont pu survivre jusqu'à nos jours. Les protéines AGO sont présentes dans environ 65% des génomes eucaryotes séquencés, elles sont dispersées dans au moins quatre des cinq super-groupes eucaryotes, dans 32% des archées et dans 46% des phylums bactériens (Swarts et al. 2014). Aujourd'hui, l'analyse phylogénétique des Argonautes de *H. sapiens*, *A. thaliana*, *C. elegans*, *C. briggsae*, *D. melanogaster* et de *S. pombe* a suggéré que ces protéines peuvent être divisées en trois clades (Fig. 5), les Ago-like, Wago et Piwi-like (Tolia et Joshua-Tor 2007).



**Figure 5 – Distribution des protéines argonautes.** À partir de l'analyse phylogénétique des AGO de *H. sapiens*, *A. thaliana*, *C. elegans*, *C. briggsae*, *D.melanogaster*, et *S. pombe*. Le clade Ago-like, basé sur la similitude avec *AtAgo1* est représenté en noir, le clade Piwi-like sur la base de similitude avec *Dm Piwi* en vert et le clade WAGO, spécifique du nématode, est représenté en rouge (Joshua-Tor et Hannon 2011).

Les Ago-like sont distribuées largement : on les retrouve chez les animaux, les plantes et les levures, elles sont présentes dans presque tous les tissus de ces organismes. Les Wago



sont les très nombreuses protéines Argonautes spécifiques des nématodes. Enfin, la sous-famille Piwi est spécifique à l'animal et exprimée presque exclusivement dans les tissus gonadiques, souvent uniquement dans les cellules germinales. Les protéines Piwi ne sont pas exclusives des bilatériens, on en retrouve dans les cellules souches des Cnidaires, *Nematostella vectensis* (Grimson et al. 2008), *Hydra magnipapillata* (Juliano et al. 2014), des Planaires, *Schmidtea mediterranea* (Palakodeti et al. 2008) et des Eponges (*Amphimedon queenslandica* (Grimson et al. 2008). Le dernier ancêtre commun des eucaryotes aurait donc possédé au minimum une protéine de chaque famille de l'ARNi : Argonaute, Piwi, Dicer et une RdRP (Cerutti et Casas-Mollano 2006). Les piARN auraient émergés dès l'apparition des animaux montrant une coévolution entre les ET et leurs hôtes.

#### **B.4.2 Découverte**

Les premiers piARN à avoir été observés provenaient de testicules de *Mus musculus*. Ils sont très abondants : on en trouve environ un million par cellule dans la lignée germinale de la souris (A. Aravin et al. 2006). Ils sont essentiels pour la spermatogenèse qui peut être divisée en trois phases durant lesquelles des piARN associés à différents membres de la famille PIWI (MIWI, MILI, and MIWI2) sont exprimés (V. N. Kim 2006). Les chercheurs ont séparé les piARN en deux catégories suivant la protéine qu'ils guident : les 26-28nt qui s'associent à MILI (A. Aravin et al. 2006) et les 29-31nt qui sont complexés avec MIWI (Girard et al. 2006), ces deux protéines sont homologues à Aubergine chez *D. melanogaster*.

Chez *D. melanogaster*, on trouve trois protéines Argonautes de la sous-famille PIWI: Piwi, Aubergine et Argonaute 3 (Ago3). Elles sont présentes spécifiquement dans la lignée germinale (mâle et femelle) sauf Piwi que l'on peut également observer dans les cellules somatiques de l'ovaire (Malone et al. 2009). En 2006, l'équipe de M. Siomi a mis en avant que c'est la protéine Piwi qui charge les petits ARN venant des régions répétées et des éléments transposables. La production d'un anticorps anti-Piwi leur a permis de montrer que la protéine était très présente chez l'embryon jeune, puis qu'elle se concentrait ensuite dans la lignée germinale mâle et femelle (Saito et al. 2006). On détecte également Piwi dans les cellules somatiques de la pointe des testicules, au niveau de la niche des cellules germinales mâles et dans les noyaux somatiques et germinaux sur toute l'ovariole chez la femelle. Des immunoprécipitations de Piwi à partir de lysat d'ovaires ont permis de montrer que les petits ARN associés à cette protéine sont un peu plus longs que des miARN et l'utilisation de sondes contre différents miARN a définitivement écarté la possibilité que les miARN soient

chargés par Piwi (Saito et al. 2006). 86% des petits ARN trouvés en complexe avec Piwi correspondent à des rasiARN, 40% d'entre eux sont homologues à des séquences d'éléments transposables avec une légère majorité d'antisens à ces transposons. L'autre partie correspond à des séquences hétérochromatiques. Enfin, l'étude *in vitro* d'une protéine de fusion GST-Piwi a prouvé que Piwi est capable de couper des ARNm cibles lorsqu'elle est guidée par un siARN complémentaire à l'ARN de la luciférase. Piwi montre donc autant d'activité endonucléase qu'Ago1.

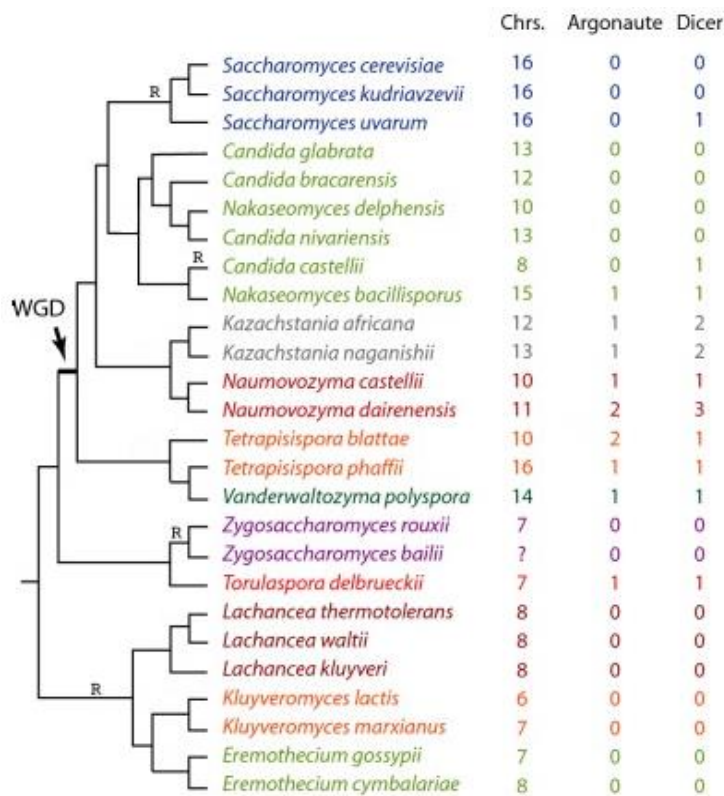
Chez *C. elegans*, des piARN de 21nt nommés 21U-ARN sont produits à partir de deux *loci* sur le chromosome 4. Ils sont transcrits par l'ARN polIII. Les précurseurs sont maturés dans les granules P situées à la périphérie du noyau et pris en charge par la protéine Argonaute Piwi PRG-1 puis maturés en petits ARN de 21nt. Ils ciblent les ARN étrangers par appariement imparfait, ce qui déclenche la production de 22G-ARN qui conduiront à la répression des transcrits envahisseurs. Ces 22G-ARN associés à WAGO peuvent également agir dans le noyau pour diriger la formation de chromatine répressive. Ils sont exprimés dans la lignée germinale à partir du stade L4 et transmis à l'embryon indépendamment par les deux parents (Batista et al. 2008). Ces piARN de nématode semblent pour l'instant avoir comme cibles les transposons de la famille *Tc3* (Batista et al. 2008), et PRG-1 est nécessaire à la spermatogenèse en assurant la bonne expression des gènes exprimés pendant la fabrication des spermatozoïdes (G. Wang et Reinke 2008).

#### **B.4.3 Perte chez certains organismes**

Un mécanisme d'ARN interférence, le Quelling, a été découvert chez les champignons par Romano et Macino en 1992. Il est activé par la présence d'ARN aberrants et requiert les acteurs canoniques de l'interférence ARN : DICER, AGO, RdRP et des siARN de 20–25nt (Romano et Macino 1992). Des mécanismes d'hétérochromatinisation ont été mis en évidence chez *S. pombe* au niveau de l'ADN des rétroéléments (Verdel et al. 2004). RdRP serait attirée par le spliceosome vers des ARNm possédant des introns cryptiques. L'origine de ce mécanisme serait que les ARNm « parasites » ne sont pas aussi efficacement épissés que les gènes endogènes ce qui déclencherait la réponse ARNi.

Si les mécanismes de défense contre les transposons sont nombreux chez les champignons, une grande partie de ce phylum a perdu les composants de l'ARNi au cours de l'évolution. Par exemple, la levure modèle *S. cerevisiae* ne possède ni protéines Dicer, ni protéines Argonautes (Fig. 6). Ces pertes peuvent être expliquées de plusieurs façons. La

réintroduction d'Ago1 et Dcr1 de *S. castellii* permet de réactiver l'ARNi chez *S. saccharomyces* : cette réactivation a peu d'effet sur le phénotype et sur le transcriptome de la levure. La seule différence notée par les chercheurs est la perte d'un virus à ARN double brin: le virus L-A aussi appelé Killer Virus, et son ARN satellite M (Drinnenberg et al, 2011). Or ces deux ARN viraux confèrent une immunité aux cellules qui les portent et tuent les autres cellules proches avec une toxine, augmentant ainsi l'accès aux ressources des levures infectées. Cette symbiose levure/virus est présente chez toutes les espèces ayant perdu l'ARNi. Une autre hypothèse énonce qu'un phénomène de mutation dirigée a pu éliminer la capacité de nuisance des éléments transposables à une époque lointaine, supprimant la nécessité de leur régulation par l'ARNi. Enfin, dans le cas des champignons unicellulaires, l'aneuploïdie peut être un avantage et la perte de l'interférence ARN a été pointée comme facteur de risque pour la mauvaise ségrégation des chromosomes chez *S. pombe* (Billmyre et al. 2013).



**Figure 6 - Phylogénie des différentes espèces de la famille *Saccharomycetaceae*.** Résumé des différences au niveau du nombre de protéines des familles Argonaute et Dicer. La lettre R montre le moment de la perte de tous les gènes Argonaute ou de tous les gènes Dicer dans une espèce. Modifié d'après (Wolfe et al. 2015).

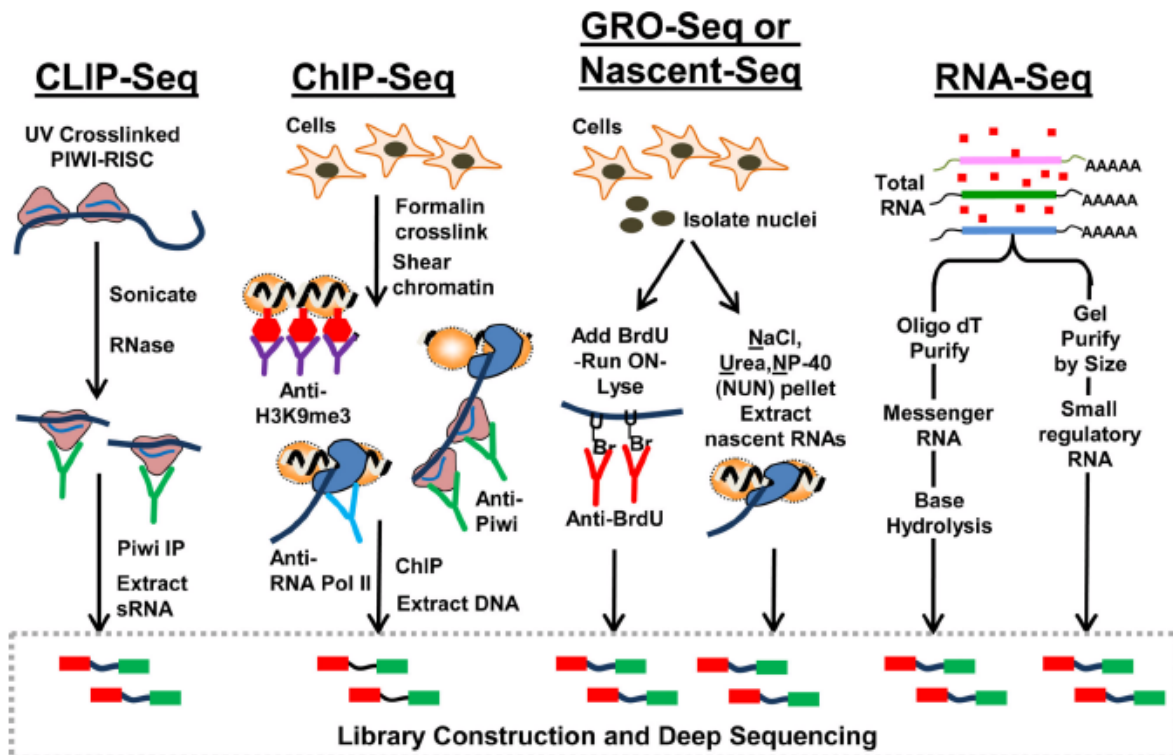
L'instabilité génomique n'est pas forcément une mauvaise chose chez les unicellulaires, elle peut même être un avantage évolutif important. Néanmoins, les éléments transposables *Ty2* et *Ty3p* de *S. cerevisiae* montrent des signes d'activité récente et seraient arrivés par transfert horizontal chez cette levure (Carr et al. 2012) indiquant que la perte de l'ARNi entraîne d'inévitables conséquences. Ceci explique peut-être pourquoi il n'y a pas de

perte des protéines Ago ou Dicer dans un grand groupe d'espèces mais plutôt plusieurs pertes successives au sein de la famille *Saccharomycetaceae* (Fig. 6).



## Chapitre 2 : Les Clusters Producteurs de piARN

De nombreuses méthodes ont été développées pour l'analyse des clusters producteurs de piARN, des ARN, piARN et de leurs protéines partenaires. La Figure 7 présente un schéma récapitulatif des expériences de CLIP-seq, ChIP-seq, Gro-seq et RNA-seq.



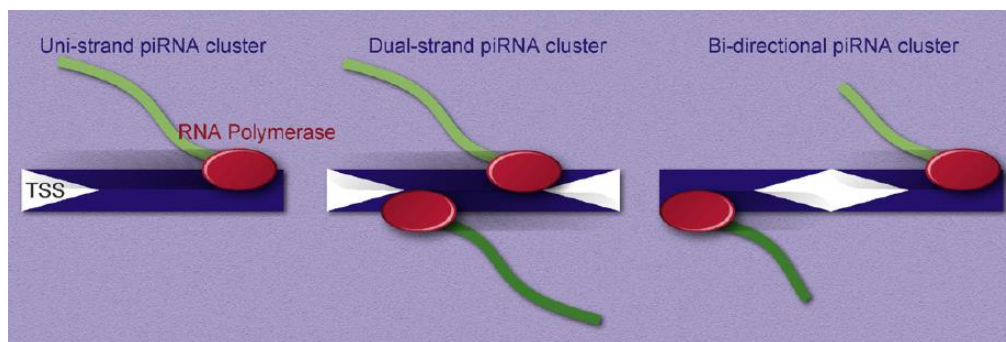
**Figure 7 - Présentation des techniques actuelles les plus courantes pour l'analyse des ARN et de leurs protéines partenaires.** Cross-Linking ImmunoPrécipitation (CLIP), ces expériences sont généralement suivies d'un séquençage des ADNc obtenus par transcription inverse. Chromatin ImmunoPrécipitation (ChIP) qui permet l'analyse de séquences ADN associées à des protéines d'intérêt. Global Run On-sequencing (GRO-seq), et le séquençage de l'ARN (RNA-seq) (Clark et Lau 2014).

Le Cross-Linking Immuno Précipitation (CLIP) permet d'enrichir un extrait cellulaire en transcrits interagissant avec une protéine d'intérêt, par exemple Piwi. La Chromatin ImmunoPrécipitation (ChIP) permet de récupérer des séquences ADN où se fixe une protéine d'intérêt, par exemple l'ARN polymérase II ou l'histone H3 triméthylée sur la Lysine 9. Le GRO-seq sert à observer quels sont les *loci* transcrits dans un certain contexte physiologique. Enfin, le RNA-seq permet de récupérer les ARN d'un extrait cellulaire, on peut par exemple se concentrer sur les ARNm en utilisant leur queue PolyA ou bien faire migrer tous les ARN et récupérer sur gel ceux dont la taille est intéressante. C'est ce que l'on fait dans le cas du

séquençage haut débit des petits ARN non-codants comme les piARN que j'appellerai Deep-seq dans la suite de ce manuscrit.

## A. Clusters endogènes producteurs de piARN

Les piARN sont produits à partir de longs précurseurs simple brins originaires de *loci* allant de quelques milliers de paires de bases à plusieurs centaines de kilobases. Chez la drosophile, 142 clusters producteurs de piARN ont été identifiés et classés selon le nombre de piARN qu'ils produisent. Ces sites représentent une petite portion du génome annoté de *D. melanogaster* (3.5%) mais produisent 92% des piARN (Brennecke et al. 2007). Ces clusters constituent un catalogue de fragments d'éléments transposables qui, après mobilisation, ont fini par s'insérer et se fixer dans ces régions. Les longs transcrits provenant de ces régions sont pris en charge par des protéines spécifiques pour éviter leur traduction et les diriger vers la voie de fabrication des piARN où ils seront découpés et maturés. Il existe deux types majoritaires de clusters de piARN : les « uni-strand clusters », transcrits depuis un seul des deux brins de l'ADN et les « dual-strand clusters », au niveau desquels les deux brins d'ADN sont transcrits de façon convergente (Fig. 8). De plus, des *loci* producteurs de piARN présents chez les mammifères ont été qualifiés de bidirectionnels : ils possèdent un promoteur central à partir duquel chacun des brins est transcrit de façon divergente.



**Figure 8 - Classification des clusters producteurs de piARN.** Ils sont divisés en trois types selon les directions de leur transcription. Les flèches blanches indiquent des sites d'initiation de la transcription (TSS) pointant le sens de transcription. Les « uni-strand clusters » sont transcrits dans une seule direction, un seul brin de l'ADN du cluster est transcrit. Les « dual-strand clusters » contiennent deux promoteurs aux extrémités orientés vers l'intérieur de sorte que les deux brins du cluster sont transcrits. Enfin, les « bi-directional clusters » ont un promoteur bidirectionnel à partir duquel chaque côté est transcrit sur le brin ADN opposé (Hirakata et Siomi 2016).

Tous ces clusters producteurs de piARN sont situés majoritairement dans les régions hétérochromatiques péri-centromériques et télomériques du génome. On note toutefois chez *D.*

*melanogaster* l'existence de 7 clusters localisés dans des régions potentiellement euchromatiques (Brennecke et al. 2007). Chez les mammifères, la distribution des piARN clusters est très conservée : chez la souris les piARN viennent de clusters de 20 à 90kb qui peuvent être uni-strand ou bidirectionnel, une organisation que l'on retrouve chez l'humain et le rat (Girard et al. 2006; Lau et al. 2006).

### A.1 Uni-strand clusters

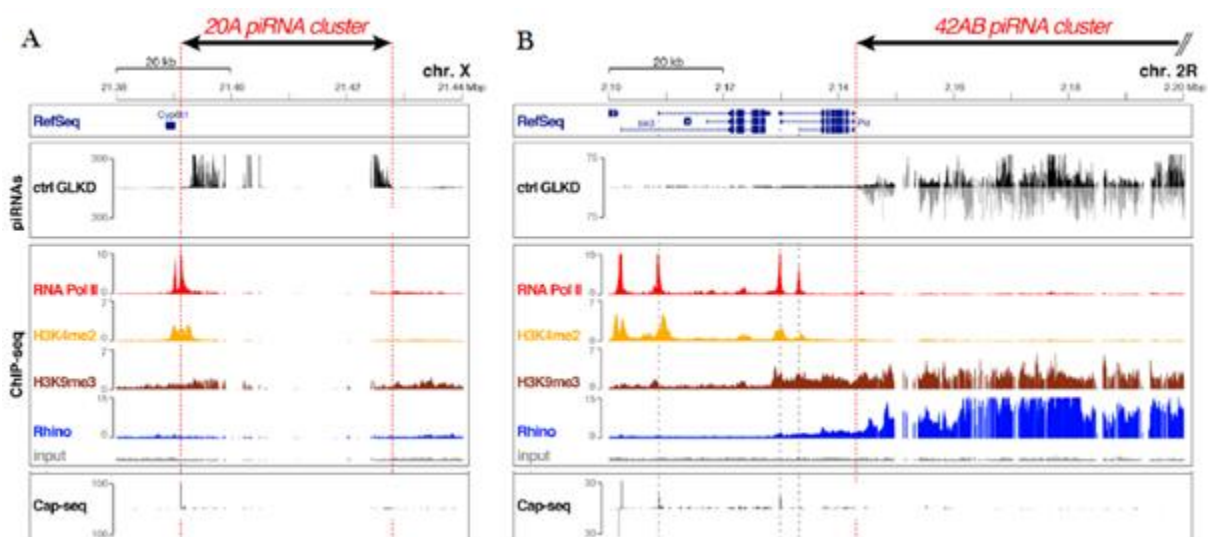
Chez *D. melanogaster*, les clusters dont un seul brin est transcrit sont généralement exprimés dans les cellules somatiques de l'ovaire : les cellules folliculaires, qui entourent la lignée germinale chez la femelle. Ils sont riches en fragments d'éléments transposables dont la majorité est insérée dans le sens contraire à la transcription du cluster lui-même. Ils produisent ainsi des piARN antisens aux ARNm d'ET qui sont donc directement efficaces pour la répression : ils sont complémentaires aux ARN sens produits par les copies euchromatiques d'éléments transposables.

Le cluster le plus exprimé dans ces cellules est *flamenco* (179 kb dont 87% d'ET ou de fragments d'ET), situé près du centromère du chromosome X au site *20A1-3*. 99% des piARN uniques de *flamenco* (qui ne peuvent pas être attribués à une autre séquence dans le génome) viennent du brin sens et plus de 95% d'entre eux sont chargés par Piwi (Brennecke et al. 2007). Ce cluster possède un promoteur enrichi en marques épigénétiques activatrices de la transcription, H3K4me2, il est transcrit par l'ARN Pol II et son expression dépend du facteur de transcription *Cubitus interruptus* (*Ci*) (Goriaux et al. 2014). Comme les transcrits des séquences géniques, les ARN provenant du *locus flamenco* (*flam*) sont coiffés par le Cap Binding Complexe (CBC), épissés et transportés dans le cytoplasme. Ils sont ensuite maturés en piARN dans une structure proche du noyau appelée Yb body qui concentre une grande partie des protéines impliquées dans la biogenèse des piARN. Bien qu'exprimé uniquement dans les cellules somatiques ovariennes, *flamenco* participe à la formation de la lignée germinale: sa transcription est requise dans ce tissu somatique pour le bon développement de l'ovaire. Des mouches « mosaïques » dont la lignée germinale est homozygote mutante pour *flamenco* (délétion du *locus flam*<sup>BG(P)</sup>) et dont le soma est hétérozygote (*flam*<sup>BG(P)/+</sup>) présentent des ovarioles parfaitement formées. En revanche, les mouches entièrement mutantes présentent des problèmes de développement de la lignée germinale : les ovaires sont atrophiés. La différenciation et la division des cellules germinales sont affectées. Les auteurs postulent que ce phénotype est dû à un défaut de communication entre le soma et la lignée



germinale : si les cellules folliculaires n'ont pas le *locus flam*, leur action d'escorte et de soutien à la lignée germinale est perturbée (Mével-Ninio et al. 2007). Les cellules somatiques folliculaires constituent probablement la première barrière de défense contre certains éléments transposables somatiques comme *gypsy*. La réplication de ce rétrotransposon est bloquée par la présence du *locus flam* via des protéines PIWI (Sarot et al. 2004). En effet, la présence de piARN correspondant à la séquence de *gypsy* dans ces cellules semble corrélée avec la répression de ce dernier (Péllisson et al. 2007). De la même façon, la mobilisation des éléments transposables *Idefix*, *gypsy* et *Zam* est empêchée par la présence de cette région (*flamenco* au *locus* 20A1-3) et *COM* au *locus* 20A2-3 (Desset et al. 2003; Mével-Ninio et al. 2007).

Le cluster *20A* (Fig. 9A), proche de *flamenco*, est le second cluster uni-strand le plus productif chez *D. melanogaster* (Brennecke et al. 2007). C'est le seul cluster unidirectionnel à être transcrit dans les cellules germinales : il est actif à la fois dans les cellules germinales, les cellules somatiques ovariennes et même dans le jeune embryon (Lau et al. 2009).



**Figure 9 – (A) Comparaison d'un cluster uni-strand (20A) et (B) dual-strand (42AB).** Deep-seq des piARN uniques correspondants à la séquence de ces deux clusters. Visualisation de la présence de l'ARNpolII sur ces clusters par ChIP-seq, suivie de la localisation des marques épigénétiques H3K4me2 (activatrices de la transcription) et H3k9me3, généralement inhibitrice de la transcription mais également caractéristique des dual-strand clusters, ainsi que de la protéine hétérochromatique germinale Rhino (Mohn et al. 2014)

Sa transcription ressemble à celle des clusters producteurs de piARN chez les mammifères : un enrichissement en H3K4me2 au TSS permet le recrutement de l'ARN polymérase II (Pol II), son transcrit est coiffé (coiffe 5'méthyl-guanosine) et il possède un terminateur de transcription : le motif consensus « AATAAA » de clivage et de

polyadénylation est présent en 3' de ce cluster. On note la présence modérée de la marque répressive H3K9me3 uniquement au début de la séquence du cluster *20A* alors que les dual-strands clusters en sont recouverts (Sienski et al, 2012; Mohn et al. 2014).

## **A.2 Transcription des dual-strand clusters**

Les cellules germinales montrent l'expression quasiment exclusive de dual-strand clusters. Ceux-ci ont une transcription très particulière, les clusters *42AB* (Fig. 9B), *80F* et *38C1*, les plus productifs en piARN (avec *flam* et *20A*) sont totalement recouverts de H3K9me3. (Mohn et al. 2014). Bien que ce soit une marque épigénétique considérée comme répressive, la présence de l'ARN PolIII a été établie par CHIP-seq sur toute la longueur de ces 3 clusters et les expériences de séquençage ARN montrent clairement qu'ils sont transcrits et qu'ils produisent des piARN en absence de TSS identifié (pas d'enrichissement en H3K4me2 au début des séquences). Les transcrits n'ont pas non plus de coiffe, ni de signal de polyadénylation. L'absence de coiffe devrait les adresser à la voie de dégradation des ARN. Pourtant, ils permettent de produire des piARN matures, suggérant qu'ils y échappent. Dans l'étude de Mohn et collaborateurs (2014), la différence majeure entre les uni-strand et dual-strand clusters est le recrutement de la protéine Rhino au niveau de ces derniers, l'implication de cette protéine à chromodomaine dans la transcription des clusters de piARN sera détaillée dans le Chapitre 3.

### ***A.2.1 Un cluster centromérique producteur de piARN, le 42AB***

Le cluster *42AB* est exclusivement transcrit dans les cellules germinales. En effet, dans la lignée cellulaire somatique « Ovary Somatic Sheet » qui provient de cellules entourant la lignée germinale dans les ovaires, moins de 1% de piARN correspondent à ce *locus* (Lau et al. 2009). Il produit en revanche 12% des piARN ovariens uniques ce qui en fait le plus gros producteur de piARN germinaux. Ceux-ci sont transmis à l'ovocyte et protègent l'embryon de la mobilisation des ET au début du développement : 15% des piARN embryonnaires correspondent aux 240kb de cette séquence (Brennecke et al. 2008). Ce cluster est constitué de plusieurs centaines de fragments d'éléments transposables incomplets, insérés dans les deux sens (familles *gypsy*, *copia*, *invader*, ...). Il présente toutes les caractéristiques d'un dual-strand cluster : on ne détecte pas de promoteur identifiable immédiatement au début du cluster, en revanche, ceux des gènes qui l'entourent pourraient servir de départ de transcription. Il n'y a pas de splicing des ARNm produits par le cluster ni de polyadénylation

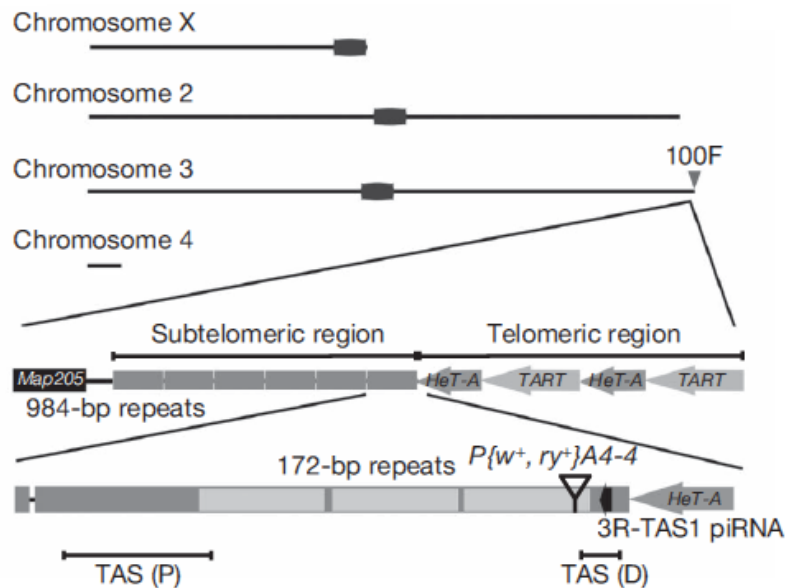
(Fig. 9B). Enfin, on peut observer un enrichissement en marques H3K9me3 sur toute sa longueur (Mohn et al. 2014).

### ***A.2.2 Les Telomeric Associated Sequences, des clusters producteurs de piARN situés aux extrémités des chromosomes***

Les « Telomeric Associated Sequences » (TAS) sont des régions sub-télomériques situées entre les télomères strictement parlant (rétrotransposons *HeT-A*, *TART* et *TAHRE*) et les premiers gènes euchromatiques. La présence de ce genre de séquences télomériques productrices de petits ARN est assez commune au sein de l'arbre du vivant. Elles ont été découvertes chez *D. melanogaster* grâce à leur capacité à réprimer la dysgénésie des hybrides (Kidwell et al. 1977). En effet, l'élément *P*, un transposon, est connu pour causer la stérilité chez les mouches incapables de réprimer sa mobilisation. Des cassures chromosomiques et un taux de mutation élevée engendrent une stérilité thermo-dépendante de la descendance couplée à l'atrophie des gonades à 29°C (Kidwell et al. 1977; Engels et Preston 1979). La dysgénésie des hybrides a lieu quand une femelle ne portant pas d'élément *P* est croisée avec un mâle en possédant. Mais lors du croisement réciproque, c'est à dire une femelle avec des éléments *P* avec un mâle n'en possédant pas, aucun symptôme n'est observé. Plus tard, il a été montré que cette propriété était due à la présence d'un élément *P* inséré dans le TAS du chromosome X (Ronsseray et al. 1998). La localisation de cet élément à ce *locus* précis lui a permis d'acquérir la propriété de produire des piARN dirigés contre sa séquence réprimant ainsi la transposition des copies euchromatiques de *P* et ainsi de protéger le génome (Todeschini et al. 2010; Pöyhönen et al. 2012). Il a été appelé TSE pour *Trans-Silencing Effect* (Roche et Rio 1998) et fait l'objet de nombreuses études au laboratoire. Ce phénomène de répression n'est pas uniquement valable pour cet élément, il s'agit d'un phénomène général basé sur l'homologie de séquence entre un fragment d'ADN inséré dans les TAS (ou plus largement dans un cluster producteur de piARN) et des ARNm (Pöyhönen et al. 2012).

Des piARN sens et antisens correspondant aux TAS ont été retrouvés ce qui laisse penser que ces séquences répétées sont transcrites dans les deux sens (Brennecke et al. 2007). Chez *D. melanogaster*, il existe deux types de TAS : les TAS-R et les TAS-L (Mason et al. 2008). Les TAS-R sont présents sur les bras droits des chromosomes 2 et 3 et sur le bras gauche du chromosome X (Karpen et Spradling 1992). Elles sont composées de répétitions « d'unité TAS » composées de trois répétitions d'une partie du LTR du rétrotransposon *Invader 4* (172-bp repeats, Fig. 10) entourées de séquences qui ne correspondent pas à des

éléments transposables connus. Ces « unités TAS » sont différentes entre elles mais semblent avoir une origine commune (Laidin, en préparation). A. Asif-Laidin a également montré qu'il existe un polymorphisme d'une souche à l'autre, la présence de TAS-R aux télomères du chromosome X, du 2R et du 3R n'est pas donc systématique. En revanche, aucune souche de *D. melanogaster* n'en est totalement dépourvue, ce qui laisse penser que ces séquences ont une fonction essentielle.



**Figure 10 - Organisation du TAS 3R situé au site 100F du chromosome 3.** La bande polytène 100F est une portion chromosomique localisée d'après l'observation de chromosomes polytènes de larves, elle contient le TAS et le télomère du chromosome 3R. La région TAS est composée de plusieurs sous unités placées entre les rétroéléments du télomère et les premiers gènes euchromatiques. Chacune de ces sous unités contient trois répétitions d'une partie du LTR du rétrotransposon *Invader 4* (172-bp repeats) et une séquence ne correspondant pas à des éléments transposables connus. (Yin et Lin 2007).

Les TAS-L se composent d'un mini-satellite de 460pb répété plusieurs fois (Walter et al. 1995). Aucune séquence d'élément transposable n'est reconnaissable mais des petits ARN homologues à ces séquences ont été trouvés dans plusieurs membres de la famille *Drosophilae* (Laidin, en préparation). Pour les TAS-R, des expériences de ChIP-seq ont montré que ces séquences possèdent une chromatine particulière : l'ADN de la séquence télomérique TAS 3R présente à la fois des marques épigénétiques spécifiques de l'euchromatine, H3K4me2, H3K4me3 et H3K9ac, mais aussi de l'hétérochromatine, H3K9me3, H4K12ac et HP1 (Yin et Lin 2007). En condition mutante pour *piwi*, les marques euchromatiques diminuent et les marques répressives augmentent au niveau du TAS 3R. Les chercheurs observent l'interaction de la protéine Piwi avec la région 3R TAS et on observe la

perte d'un piARN abondant spécifique de cette région (le piARN 3R-TAS1) en mutant *piwi*. Le recrutement de Piwi semble donc impliqué dans la définition de l'identité de ces séquences et dans leur transcription (Yin et Lin 2007).

## **B. Utilisation de transgènes « *P-lacZ* » pour l'étude de la régulation de l'élément *P* : Le *Trans-Silencing Effect***

La découverte des TAS en tant que *loci* producteur de piARN a été précédée par leur identification comme *locus* responsable de la répression d'éléments *P* euchromatiques grâce à l'insertion d'une copie de *P* dans le TAS du chromosome X. Ceci ayant pour effet d'éviter la dysgénésie des hybrides et donc protège le génome de la mobilisation de cet élément (Chap. 2 Partie A.2.2). Il a ensuite été montré que ce phénomène de répression n'est pas uniquement efficace pour ce transposon mais qu'il s'agit d'un phénomène général basé sur l'homologie de séquence qui a été appelé TSE (Roche et Rio 1998), entre un fragment d'ADN inséré dans les TAS (ou plus largement dans un cluster producteur de piARN) et des transcrits partiellement homologues (Marin et al. 2000; Josse et al. 2007).

### **B.1 Les clusters de piARN transgéniques *T-1* et *BX2***

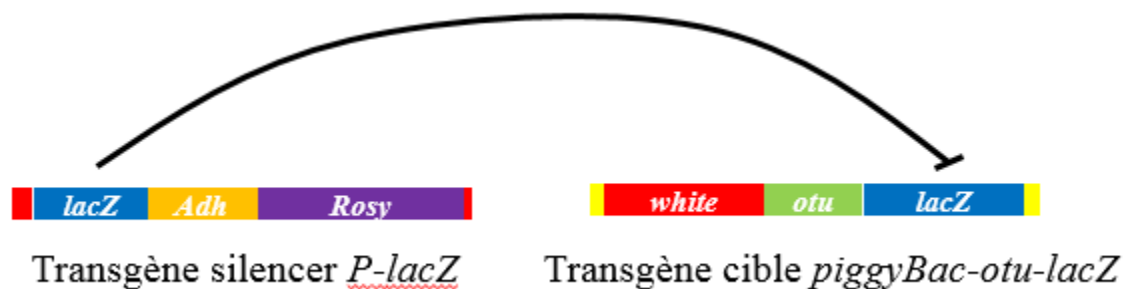
La lignée *T-1* possède 7 transgènes non télomériques *P-lacW* (séquences du premier exon de l'élément *P*, du gène *lacZ* et du gène *white*), insérés en tandem sur le bras droit du chromosome 2 (Ronsseray et al. 2001). Cette insertion de transgènes est létale homozygote, probablement car elle interrompt le gène *ago1*. Elle induit la formation d'un *locus* hétérochromatique localisé (Dorer et Henikoff 1994; Dorer et Henikoff 1997; de Vanssay et al. 2013) entraînant la production de piARN qui répriment des séquences euchromatiques homologues s'exprimant dans les ovaires. La lignée *BX2* qui possède les 7 mêmes transgènes au même site chromosomique n'a pas les mêmes propriétés épigénétiques: aucun piARN correspondant au *locus BX2* n'est produit (de Vanssay et al. 2013). Un traitement aux rayons X de *BX2* a permis d'obtenir la lignée *T-1*, ce qui en fait des outils précieux pour visualiser la répression médiée par les petits ARN.

Ces deux lignées ont également servi dans la démonstration du premier cas de paramutation chez l'animal : *T-1* et *BX2* possèdent la même insertion de transgènes mais ont des propriétés épigénétiques différentes. Il a été démontré au laboratoire que l'héritage maternel de piARN homologues à *BX2* provenant de la souche *T-1* était capable d'activer la production de piARN par *BX2*. Les descendants d'un croisement de femelles hétérozygotes *T-1* par des mâles *BX2* sont capables de réprimer des transgènes cibles homologues même lorsque le *locus* producteur original (*T-1*) n'est plus présent : les piARN transmis par le cytoplasme de l'ovocyte ont « réveillé » le *locus* dormant *BX2*. Cette nouvelle lignée appelée

*BX2\** produit des piARN *de novo* (de Vanssay et al. 2013). Elle est stable au cours du temps (plus de 100 générations) et est capable à son tour d'activer la production de piARN chez un *BX2* « naïf ». Une étude récente a montré qu'une des différences entre *T-1* et *BX2* est la présence de la marque H3K9me3, environ quatre fois plus présente sur *T-1*. De plus, la conversion de *BX2* en *BX2\** est suivie de l'apparition de H3K9me3 au niveau de ce *locus* (Le Thomas et al. 2014). La capacité à produire des piARN semble donc étroitement liée à la présence de cette marque dans les clusters double brins. Sans transmission maternelle de piARN, le *locus T-1* arrête de produire des piARN (de Vanssay et al. 2013) et la présence de H3K9me3 baisse de moitié (Le Thomas et al. 2013). Ceci suggère que des piARN hérités de la génération précédente induisent dans la descendance un changement dans l'état de la chromatine des séquences auxquelles elles sont homologues.

## B.2 Transgènes « silencer » et transgènes « cible »

L'utilisation de transgènes exprimant des protéines rapporteurs dans la lignée germinale grâce à un phénomène "enhancer-trap" couplée à la présence de transgènes homologues situés dans les TAS nous permet d'observer la répression médiée par les piARN (Fig. 11).



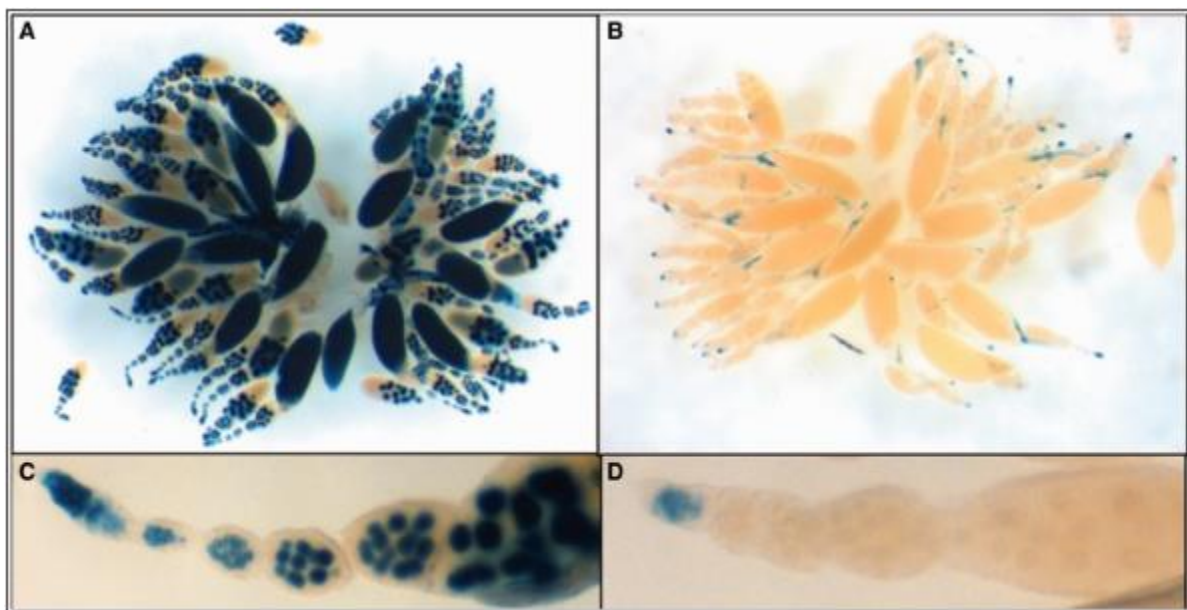
**Figure 11 - Répression d'un transgène cible *PBoL* par un transgène *P-lacZ*.** Une femelle portant un transgène silencer *P-lacZ* est capable de réprimer en *trans* un transgène cible *piggyBac-otu-lacZ* indépendamment de la présence de séquences de *P*. Ici, l'homologie de séquence entre les séquences de *lacZ* est suffisante pour induire la répression de la transcription du transgène cible.

Au laboratoire, nous utilisons ce système pour mimer la répression des éléments transposables et ainsi l'étudier plus facilement. Historiquement, ces transgènes étaient des *P-lacZ* (Fig. 11) contenant les deux premiers exons de l'élément *P* fusionnés en phase avec le gène *lacZ* d'*E. coli* et produisant la  $\beta$ -galactosidase lorsqu'insérés en contexte permissif (Roche et Rio 1998; Marin et al. 2000). Aujourd'hui nous utilisons également d'autres

protéines rapporteurs fluorescentes, telles que la GFP ou la RFP, comme cibles. Ce système repose donc sur deux familles de transgènes :

Les transgènes « silenciers » sont insérés dans les *loci* producteurs de piARN sub-télomériques hétérochromatiques, les TAS. Ce contexte chromatinien particulier leur permet d'échapper à la transcription canonique : la  $\beta$ galactosidase n'est pas produite, on n'observe aucune coloration après une coloration au X-gal d'ovaires provenant d'une lignée contenant un transgène silencer et ils produisent des piARN correspondant à toute leur séquence. Durant ma thèse, j'ai utilisé majoritairement deux lignées silencer différentes : *RS3* qui contient un transgène *P-FRT-white* dans le TAS du bras droit du chromosome 3 et *P-1152* possédant deux transgènes *P-lacZ* insérés au site 1A du chromosome X (exemple du transgène silencer de la Fig. 11).

Les transgènes « cibles » expriment une protéine rapporteur dans la lignée germinale, ils sont insérés dans l'euchromatine et sont donc transcrits de façon canonique. Après une coloration X-gal ou un immunomarquage, on observe une coloration des ovaires provenant d'une lignée cible (Fig. 12A-C).



**Figure 12 – TSE. (A-D) coloration X-gal.** Le transgène cible *BC69* induit la coloration des cellules nourricières et de l'ovocyte au niveau du noyau. (A, C) Ovaires et détail d'un ovariole de femelle *BC69* portant une copie d'un *P-lacZ* enhancer-trap inséré dans le gène *vasa*. (B, D) Ovaires de descendantes de croisement de mâles *BC69* avec des femelles *RS3* (Silencer *P-FRT-white* au TAS du 3R). Le marquage  $\beta$ -galactosidase est fortement réduit en présence du transgène silencer fort *RS3*, sauf dans le germarium (Dufourt et al. 2014).



Lorsqu'une souche de *Drosophile* possède un couple silencer/cible partageant de l'homologie, le transgène cible est réprimé par le transgène silencer : les ovaires n'expriment pas la protéine rapporteur après coloration (Fig. 12B). Il peut arriver que le transgène cible soit exprimé de façon modérée dans le germarium (Fig. 12B-D), on appelle ce phénomène le « Patterning » en raison de la localisation très stéréotypée de cette réexpression du transgène cible au milieu du germarium. Cette région est appelée Piwi-less Pocket (Dufourt et al. 2014). En effet on détecte une baisse modérée de l'expression de certaines protéines impliquées dans la voie des piARN à cet endroit et les conséquences de l'existence d'une telle région sont discutées dans la partie A.3.1 du Chapitre 4 qui décrit le germarium.

Les piARN provenant du *locus* silencer ont donc permis de réprimer l'expression du transgène cible par ciblage de son ARNm (PTGS) ou par répression transcriptionnelle au niveau de sa séquence ADN (TGS). Chez l'adulte, le TSE peut être estimé en comptant le nombre de chambres blanches et bleues au sein d'un ovaire (Josse et al. 2007). Durant le développement, la quantification des cellules germinales primordiales (PGC) grâce à un marquage *Vasa* et l'évaluation de la proportion de celles-ci exprimant la cible permettra d'établir un pourcentage de répression.

### **B.3 Propriétés et analyse du TSE**

#### ***B.3.1 Effet maternel***

Le TSE est efficace dans la lignée germinale femelle et présente un effet maternel car une forte répression du transgène cible n'est observée que lorsque le transgène silencer est transmis par la femelle (Fig. 13) (Ronsseray et al. 2001). En revanche, l'expression de la cible est indépendante de sa transmission maternelle ou paternelle.



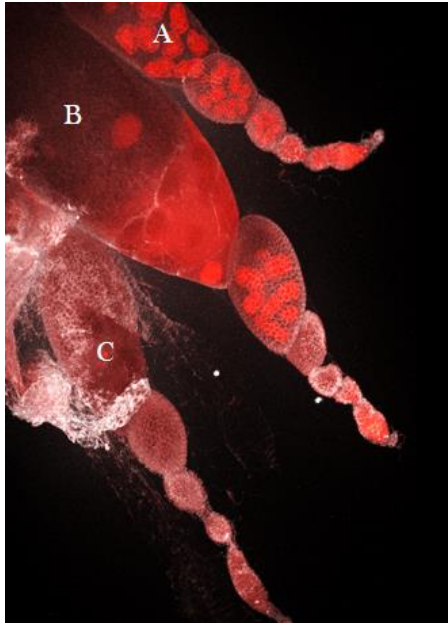
C'est dans l'article de T. Josse que la relation a été faite entre répression d'une cible et un facteur extra-chromosomique transmis de façon strictement maternel. Il a également montré que celle-ci était complètement perdue dans un contexte mutant pour plusieurs gènes impliqués dans l'ARN interférence : *aubergine*, *homeless*, *armitage*, *piwi* et *Su(var)205/HP1a*, rapprochant ce facteur extra-chromosomique des petits ARN. La mutation de *r2d2* n'ayant aucun effet sur le TSE, ils concluent que celui-ci est médié par les rasiARN ou piARN et nécessite des protéines impliquées dans la formation de l'hétérochromatine. A. Asif-Laidin a montré cette année que le pourcentage de répression observé par T. Josse est directement corrélé avec la quantité de piARN observés par Deep-seq à chacune de ces générations (Laidin, en préparation).

### ***B.3.2 Propriétés épigénétiques***

Les piARN hérités maternellement induisent chez la descendance un changement dans l'état de la chromatine des séquences auxquelles elles sont homologues. Au niveau du *locus* producteur de piARN, ce changement dans la chromatine est nécessaire pour la production de piARN à la génération suivante. Dans la discussion de l'article de T. Josse, cette option avait été envisagée : « [...] *maternally transmitted small ARN molecules may modify the chromatin structure of the chromosomal copy of the (paternally or maternally transmitted) telomeric transgenes in the embryo and thereby confer to the telomeric transgene the capacity to be a silencer.* ». Des études récentes ont prouvé cette hypothèse (de Vanssay et al. 2013; Le Thomas et al. 2014). La présence de piARN générés à partir du *locus* silencer a également un impact sur la séquences des transgènes cibles : une augmentation de la marque H3K9me3 (répressive) et une diminution de la marque H3K4me2 (activatrice) sur le *locus* cible *BC69* peuvent être observées en présence du silencer *P-1152* (Le Thomas et al. 2013). Cette augmentation des marques hétérochromatiques répressives sur la séquence cible ne se produit que quand il y a héritage de piARN homologues par la mère. En l'absence de piARN hérités, le niveau de H3K9me3 sur le silencer *P-1152* diminue et la répression n'est pas observée, indiquant que le TSE est le produit du silencing post-transcriptionnel (attaque des ARNm des cibles par les piARN par le *locus* silencer) et transcriptionnel (blocage de la transcription du *locus* cible) (Le Thomas et al. 2013).

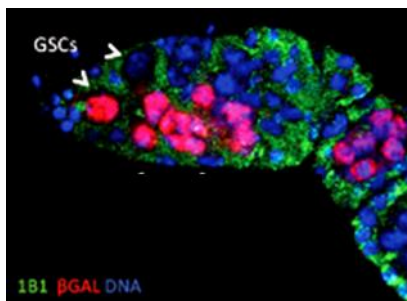
### B.3.3 Cas de répression incomplète : la Variégation

Dans certaines occasions, un TSE incomplet a été observé : au sein d'un même ovaire, certaines ovarioles montrent une répression totale (Fig. 15C) tandis que d'autres montrent l'expression du transgène qui s'exprime dans chacune de leur chambres à œuf (Fig. 15A). De la variégation est visible parmi certaines ovarioles présentant à la fois des chambres à œufs où le transgène *P-lacZ* est réprimé ou exprimé (Fig. 15B).



**Figure 15 - Variégation** – Immunomarcage sur ovaires adultes de la descendance d'un croisement d'une souche portant un silencer *P-lacZ* télomérique (*P-1152*) par une cible *P-lacZ* euchromatique (*BC69*). L'ovariole A montre l'expression de la cible *P-lacZ* ( $\beta$ galactosidase, en rouge) dans les noyaux des cellules nourricières de chacune de ses chambres à œuf alors que l'ovariole C ne montre aucune expression de la  $\beta$ galactosidase. De la variégation peut être observée dans l'ovariole B : certaines chambres continuent d'exprimer la cible  $\beta$ galactosidase tandis que d'autres la répriment.

Ce phénomène n'est pas courant et dépend du couple de transgène silencer et cible utilisé. Par exemple, la répression est généralement totale en utilisant le transgène silencer du chromosome 3, *RS3*. Dans les germariums provenant de mouches qui portent le transgène silencer *P-1152* et le transgène cible *P-lacZ* euchromatique *BC69*, on distingue également de la variégation (Fig. 16).



**Figure 16 - Détail Variégation** - Immunomarcage du germarium provenant d'ovaires *P-1152* x *BC69* comme précédemment. Une cellule souche germinale (GSC) montre l'expression de la cible *P-lacZ* ( $\beta$ galactosidase, en rouge) au niveau du noyau et cette expression semble s'étendre au cystoblastes et à ces cellules filles jusqu'à la chambre à œuf alors que l'autre GSC et ces descendantes ne montrent aucune expression de la  $\beta$ galactosidase.

Les mécanismes sous-jacents liés à la variégation sont encore inconnus et une partie de ma thèse a été consacrée à en découvrir les propriétés. On suppose qu'elle est associée à un état chromatinien du transgène cible différent selon les chambres à œufs à cause du phénomène d'effet maternel. En effet, on sait que l'héritage maternel de piARN est nécessaire

à la réactivation du transgène silencer à chaque génération. Comme on détecte de la répression liée au transgène *P-1152* dans la descendance de ces femelles, on peut penser que tous les descendants héritent de piARN, qu'ils proviennent ou non de chambres à œuf réprimant ou exprimant la cible.

### Chapitre 3 : La Biogenèse des piARN

En 2007, le laboratoire de G. Hannon a commencé à déchiffrer la biogenèse des piARN (Brennecke et al. 2007). Ils ont découvert que les piARN complémentaires montraient un overlap de précisément 10nt ce qui n'est pas compatible avec le mécanisme de fabrication des siARN par exemple. Grâce à des anticorps particuliers, ils ont pu immunoprécipiter les trois protéines PIWI en compagnie de leurs piARN mettant en évidence que chacune est associée à une population spécifique. Les piARN associés à Piwi sont les plus longs, ils font en moyenne 25.7 nt. Aubergine (Aub) prend en charge des piARN de 24.7nt en moyenne, Argonaute 3 (Ago3) préfère ceux d'une taille moyenne de 24.1nt (Brennecke et al. 2007). Les deux membres des paires de piARN ayant un overlap de 10nt sont complexés avec deux protéines différentes, généralement Aub et Ago3. Ces informations laissaient penser que les protéines PIWI elles-mêmes sont impliquées dans la biogenèse des piARN. Les protéines de la famille PIWI présentes dans le cytoplasme, Aub et Ago3 répriment les transposons au niveau de leurs transcrits par clivage de leur ARNm empêchant ainsi l'expression de transposases ou d'intégrases qui permettraient aux éléments transposables de se mobiliser. Chez *D. melanogaster* il y a une seule protéine PIWI nucléaire, Piwi, chargée de réprimer les éléments transposables transcriptionnellement, c'est-à-dire en induisant l'hétérochromatinisation des copies actives de ces ET. On appelle ce phénomène le TGS pour Transcriptional Gene Silencing.

Ces deux mécanismes complémentaires ont donc lieu dans deux compartiments cellulaires différents : le noyau et le cytoplasme. De nombreuses protéines sont impliquées dans la biogenèse des piARN dits primaires avant leur prise en charge par Aubergine et Piwi. Tout d'abord les protéines qui définissent l'identité des clusters de piARN et leur transcription non canonique puis celles impliquées dans la maturation des transcrits et enfin les différents partenaires nécessaires à l'amplification du nombre de piARN et à leur action régulatrice.

## **A. Les clusters de piARN**

Dans cette partie je détaillerai les différentes sources de piARN endogènes chez *D. melanogaster*.

### **A.1 Transcription des clusters de piARN**

Dans les cellules germinales, les dual-strand sont les producteurs principaux de piARN. Ils sont définis, entre autres, par la présence de la marque hétérochromatique H3K9me3 et des protéines qui la reconnaissent, en particulier Rhino.

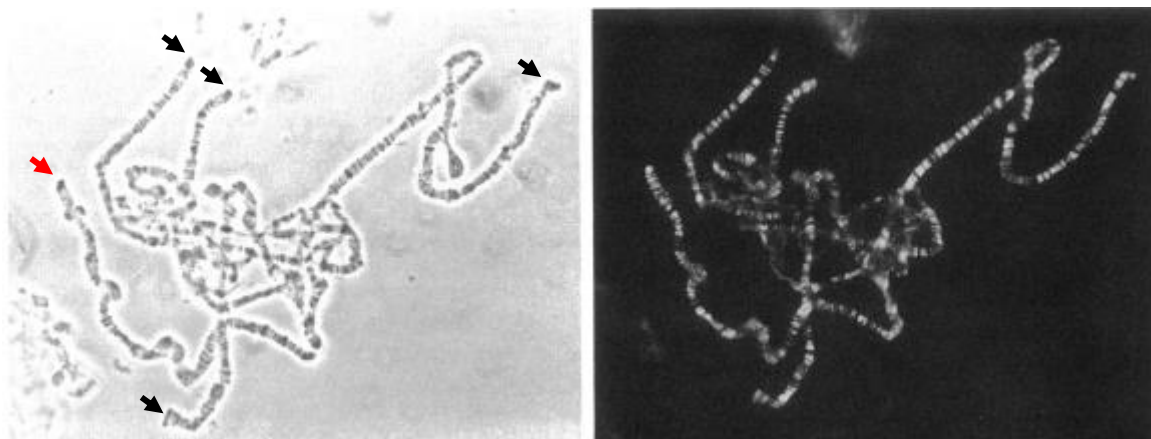
#### ***A.1.1 La chromatine***

L'unité de base de la chromatine est le nucléosome, octamère de protéines composé de deux exemplaires de chacune des histones H2A, H2B, H3 et H4 autour duquel s'enroulent 146 pb d'ADN. L'activation de la transcription génique nécessite que la machinerie transcriptionnelle puisse accéder à l'ADN et donc une diminution de la compaction de la chromatine. La modification des queues N-terminales des histones induit un changement de l'état de compaction de la chromatine autour du nucléosome. C'est un des processus permettant l'activation ou la répression de la transcription et les études récentes s'appuient souvent sur la présence de méthylation de l'histone H3 au niveau de la Lysine 9 (H3K9me3) sur un *locus* pour le catégoriser comme peu permissif à la transcription. Dans le cas des clusters producteurs de piARN, il a été montré qu'en dépit de la présence de cette modification d'histone sur leur séquence, leur transcription est très active (Brennecke et al. 2007; Mohn et al. 2014). La substitution des histones canoniques par certains variants semble également contrôler la transcription. La modification de la structure des nucléosomes mais aussi le remplacement des histones marquées par H3K9me3 par H3.3 permet l'activation immédiate des gènes qui sont réprimés par la méthylation de la Lysine 9 (Ahmad et Henikoff 2002) car celle-ci est extrêmement stable (Jenuwein 2001). H3.3 à 4 acides aminés de différence avec H3 et sa substitution peut être indépendante de la réplication (Ahmad et Henikoff 2002).

##### ***A.1.1.1 Les variants d'histones***

Les variants d'histones sont donc un mécanisme supplémentaire permettant de réguler l'accessibilité à la chromatine (pour revue (Talbert et Henikoff 2010; Baldi et Becker 2013).

Les variants de l'histone H2A sont hautement conservés à travers les espèces et jouent un rôle dans le contrôle de la transcription, la formation des bornes de l'hétérochromatine, le lignage cellulaire et la réparation de l'ADN. La plupart des eucaryotes possèdent au moins deux variants: H2A.Z et H2A.X, avec des sous-familles supplémentaires chez certaines espèces : H2A.Z1 et H2A.Z2 chez les vertébrés. Chez la *Drosophile*, H2Av est l'unique variant de H2A. Il possède des traits communs avec H2A.X et H2A.Z. Il est impliqué dans l'organisation des promoteurs, la formation de l'hétérochromatine et dans la surveillance des dommages à l'ADN (Baldi et Becker 2013). J'utiliserai l'appellation H2Av au fil de ce manuscrit mais ce variant est également connu sous les noms de H2A.V, D2, H2A.2, H2AvD, dH2A.Z et H2A.Z.X. H2Av est très exprimé durant les 12 premières heures de l'embryogenèse puis son niveau d'expression décroît (van Daal et Elgin 1992). Des expériences d'immunomarquage sur chromosomes polytènes de larves L3 montrent que celui-ci est localisé sur un grand nombre de régions chromosomiques mais pas de façon ubiquitaire (Fig. 17).



**Figure 17 – Localisation du variant d'histone H2Av sur chromosomes polytènes de larves de *Drosophile*.** Utilisation d'un antisérum spécifique de His2AvD. À gauche contraste de phase, à droite, immunofluorescence. On remarquera que H2Av semble présent aux télomères (flèches) incluant celui du chromosome X (flèche rouge) (van Daal et Elgin 1992).

Plus tard, chez l'adulte, il reste très présent chez la femelle (van Daal et Elgin 1992). Les mutants *H2Av* présentent une diminution de l'acétylation de la Lysine 12 de l'histone H4, H4K12ac, une baisse de H3K9me3 et une réduction du recrutement de HP1a centromérique (Swaminathan et al. 2005). Les expériences menées dans cet article dessinent une cascade ordonnée d'événements nécessitant le recrutement d'H2Av suivie l'acétylation de H4K12 avant l'établissement de la marque H3K9me3 et le recrutement de HP1a.



### A.1.1.2 HP1

HP1 (Heterochromatin Protein 1a) a été découverte chez la drosophile (HP1a). Elle est principalement associée à l'hétérochromatine, d'où son nom. Chez *D. melanogaster*, cinq paralogues ont été caractérisés: HP1a, codée par le gène *Su(var)205* et majoritairement hétérochromatique, HP1b, trouvée dans l'hétérochromatine et l'euchromatine, HP1c, principalement euchromatique, HP1d, codée par le gène *rhino* et exprimée uniquement dans la lignée germinale femelle et HP1e, exprimée dans la lignée germinale mâle. Dans ce manuscrit je me concentrerai sur HP1a et HP1d (appelée Rhino dans la suite du texte) qui semblent avoir un impact important sur les clusters producteurs de piARN.

HP1a est constituée d'un chromo-domaine (Paro et Hogness 1991) et d'un chromo-shadow domaine (Aasland et Stewart 1995). Ils sont respectivement nécessaires pour la liaison à la chromatine (reconnaissance des Lysines triméthylées) et pour les interactions entre protéines. HP1a est impliquée dans le « Position-Effect Variegation » (le PEV), la répression transcriptionnelle de gènes euchromatiques relocalisés vers des régions hétérochromatiques *via* un réarrangement ou une transposition (Eissenberg et al. 1990).

Des marquages sur chromosomes polytènes de larves ont permis de montrer que l'hétérochromatine  $\beta$  et le chromosome 4 sont recouverts de HP1a. Les marquages au niveau des télomères sont fréquents, en particulier sur les extrémités des chromosomes 2R, 3R et sur celles du chromosome X (James et al. 1989) ce qui inclut probablement les TAS (Yin et Lin 2007; Frydrychova et al. 2008). En effet, HP1a fait partie de la coiffe des télomères et sa présence s'étend sur toute la région *HTT*, composée des rétroéléments du télomère (*HeT-A*, *TART* et *TAHRE*) et TAS à un niveau plus modéré, ainsi que sur diverses régions des bras chromosomiques (Fanti et al. 1998; Frydrychova et al. 2008). Une mutation au niveau du chromo-domaine de HP1a entraîne la dérégulation des rétroéléments *HTT*, l'élongation du télomère et la perte H3K9me3 sans affecter sa présence à la Coiffe (Frydrychova et al. 2008a; Savitsky et al. 2002; Perrini et al. 2004).

Au niveau de l'euchromatine, l'analyse des cibles de HP1a et de Su(var)3-9 (une méthyltransférase H3K9) au niveau du génome entier a permis de définir respectivement 218 et 127 *loci* cibles pour ces deux protéines (Greil et al. 2003). Parmi les cibles de HP1a, 10 positions n'ont pas été annotées mais 56 *loci* contiennent des séquences répétitives, 46 correspondent à 27 éléments transposables uniques. Ceci concorde avec le fait que HP1a est généralement associée à un contexte chromatinien répressif. Néanmoins, les 152 *loci* restants

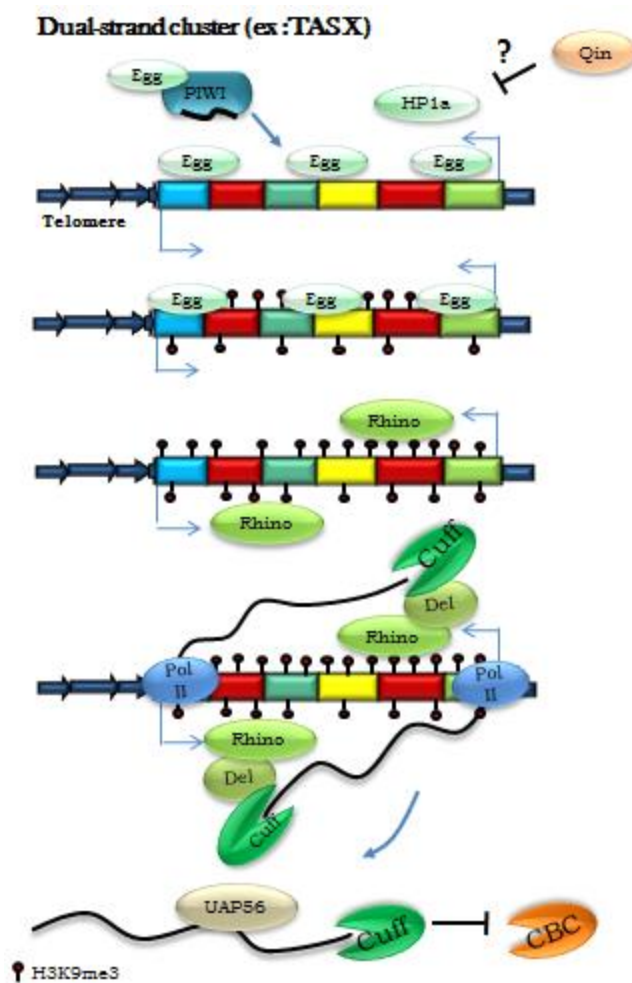
correspondent à des gènes uniques et dans cet article les auteurs mettent en avant qu'un grand nombre de gènes ciblés à la fois par Su(var)3-9 et HP1a présentent des niveaux d'expression modérés ou élevés: indiquant que le recrutement de ces protéines ne suffit pas à réprimer la transcription complètement. Il semble même que la présence de HP1a soit un frein à l'influence répressive de Su(var) 3-9 seule. De plus, l'analyse de trois gènes associés à HP1a (expériences de CHIP) a montré que certains nécessitent HP1a pour être exprimés. *Pros35* (codant la sous unité de 35-kd du protéasome), *CG5676*, et *cdc2* (codant une protéine kinase) montrent une baisse d'expression dans le mutant *Su(var)205*. Parmi ces gènes, *cdc2* est également associé à la marque H3K9me3 (Cryderman et al. 2005). Ce résultat suggère un rôle de HP1a dans l'expression de certains gènes euchromatiques.

Dans les cellules somatiques, Su(var)3-9 modifie les histones au niveau des régions péri-centromériques et télomériques ce qui attire HP1a, formant de larges blocs d'hétérochromatine. Dans les cellules germinales, Eggless/SETDB1 (une triméthyltransférase de H3K9) et son cofacteur Wendei sont guidés spécifiquement vers les clusters de piARN double brins par Piwi chargée par un petit ARN hérité maternellement (Yin et Lin 2007; Le Thomas et al. 2014). Ceci définit ces séquences comme des clusters producteurs de piARN : la présence de la marque H3K9me3 recrute alors Rhino (Rangan et al. 2011) en plus de son homologue HP1a. Ces deux protéines peuvent se lier aux séquences productrices de piARN *via* leur chromo-domaine qui reconnaît H3K9me3. Contrairement à HP1a qui est présente dans tous les tissus, Rhino est spécifique de la lignée germinale et son absence entraîne des défauts de développement similaires à ceux causés par la mutation de Piwi, montrant son rôle essentiel dans la transcription des clusters de piARN.

#### A.1.1.3 *Rhino*

Pour les chercheurs de l'article Mohn et al, 2014, la différence de transcription entre les uni-strand et dual-strand clusters est due à la présence de la protéine Rhino, homologue de HP1a. De façon intéressante, le transcriptome des mutants *rhino* ne montre pas de dérégulation au niveau des gènes (Klattenhoff et al. 2009). En contexte mutant pour Rhino, certaines familles de transposons montrent une dérégulation record : par exemple, *HeT-A* est 70 fois plus exprimé et 80% des piARN totaux sont perdus. Ce sont les clusters « dual-strand » qui montrent le plus de pertes. Par exemple, le cluster *42AB* perd 97% de ses piARN. En revanche, le cluster *20A* qui est transcrit dans un seul sens, ne montre aucune baisse de production. Des expériences de CHIP-seq montrent que Rhino est effectivement enrichie

uniquement sur les clusters producteurs de piARN transcrits dans les deux sens et que sa présence est nécessaire pour leur transcription (Mohn et al. 2014; Klattenhoff et al. 2009). Des expériences RTqPCR montrent une baisse de la quantité de transcrits des clusters dual-strand dans le mutant *rhino*. Enfin, la localisation d'autres protéines clés de la voie des piARN, les trois protéines PIWI, nécessite la présence de Rhino. Le recrutement de Rhino n'est que la première étape qui mène à la production de transcrits à partir des clusters de piARN double brin (Fig.18). En effet, c'est grâce à au chromo-shadow-domain de Rhino que les protéines Deadlock (Del) et Cutoff (Cuff) sont recrutées sur les dual-strand clusters (Mohn et al. 2014).



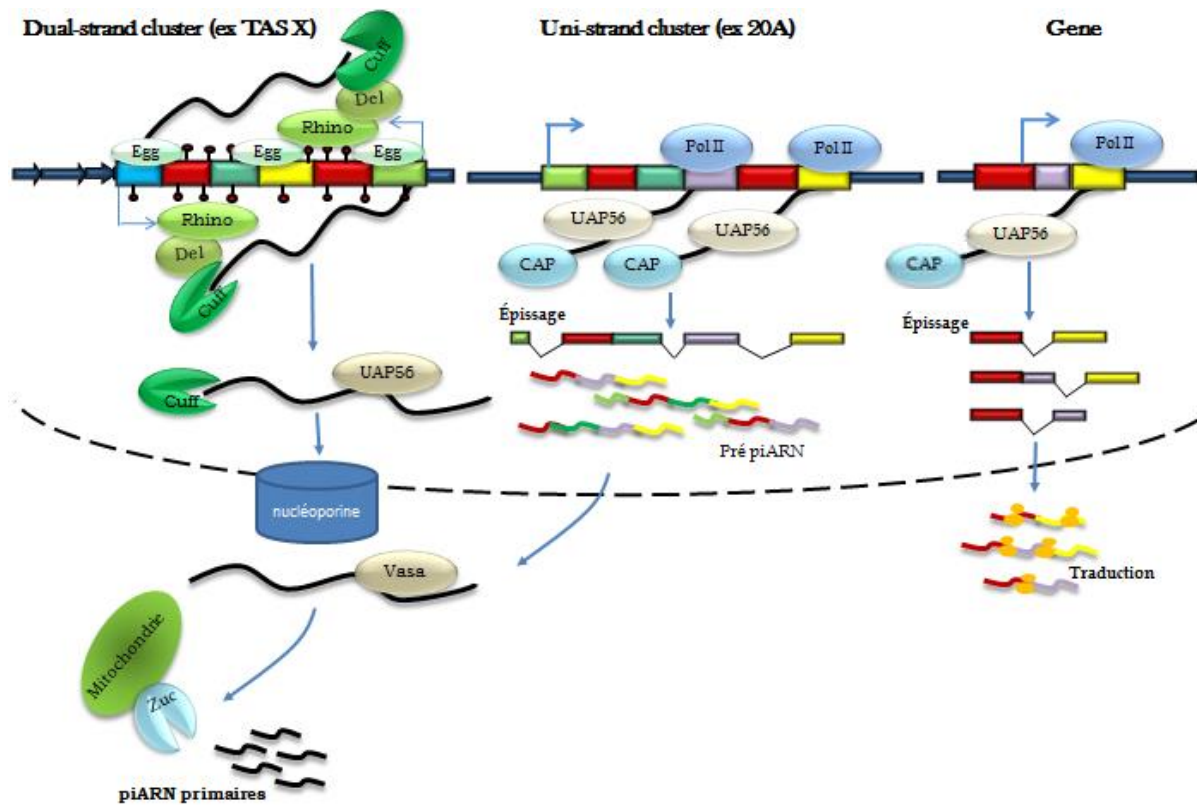
**Figure 18 - Définition d'un Dual strand cluster chez *D. melanogaster*.** Les piARN hérités maternellement et associés à la protéine Piwi sont capables de reconnaître des transcrits naissants homologues et de diriger Piwi vers les séquences qui les produisent. La protéine Qin pourrait empêcher la fixation de HP1a sur ces séquences, évitant leur répression transcriptionnelle liée à la mise en place de chromatine répressive (Anand et Kai 2012). En parallèle, la triméthyltransférase Eggless va ensuite modifier l'histone H3 en y apposant la marque H3K9me3 que va reconnaître Rhino *via* son chromo-domaine. La protéine Deadlock (Del) va alors interagir avec Rhino grâce à son chromo-shadow domain et recruter Cutoff (Cuff) qui va remplacer la coiffe sur les ARN précurseurs produits par le cluster, permettant d'empêcher la liaison du Cap Binding Complex (CBC) (Mohn et al. 2014; Chen et al. 2016). Ils éviteront ainsi la traduction.

Ce complexe appelé le RDC, pour Rhino-Deadlock-Cutoff, constitue à l'heure actuelle un facteur nécessaire pour la définition de l'identité des dual-strand clusters. Nécessaire mais peut-être pas suffisant : Rhino et H3K9me3 sont présents à d'autres *loci* au sein du génome qui ne sont pas transcrits sur les deux brins et ne donnent pas naissance à des piARN. Pour cette raison, A. Aravin a postulé qu'un autre facteur pourrait être impliqué avant le

recrutement du RDC. Pour l'instant on a observé que lorsqu'il est recruté, Rhino sert d'ancrage au complexe : son chromo-shadow domain se lie à Del qui fixe la protéine Cuff de façon transitoire (Mohn et al. 2014). Cuff se lie à la partie 5' du transcrit produit par le dual-strand cluster et le suit ensuite lors de son export. En empêchant le CAP binding complexe d'y accéder, Cuff évite la traduction des transcrits du cluster et en remplaçant la Coiffe, Cuff les protège de la dégradation (Fig. 18). De plus, la présence de Cuff permet à l'ARN Pol II d'ignorer les signaux de polyadénylation et de terminaison de transcription : ces dual-strand clusters semblent juste être transcrits en continu (Chen et al. 2016).

### ***A.1.2 Export des précurseurs de piRNA***

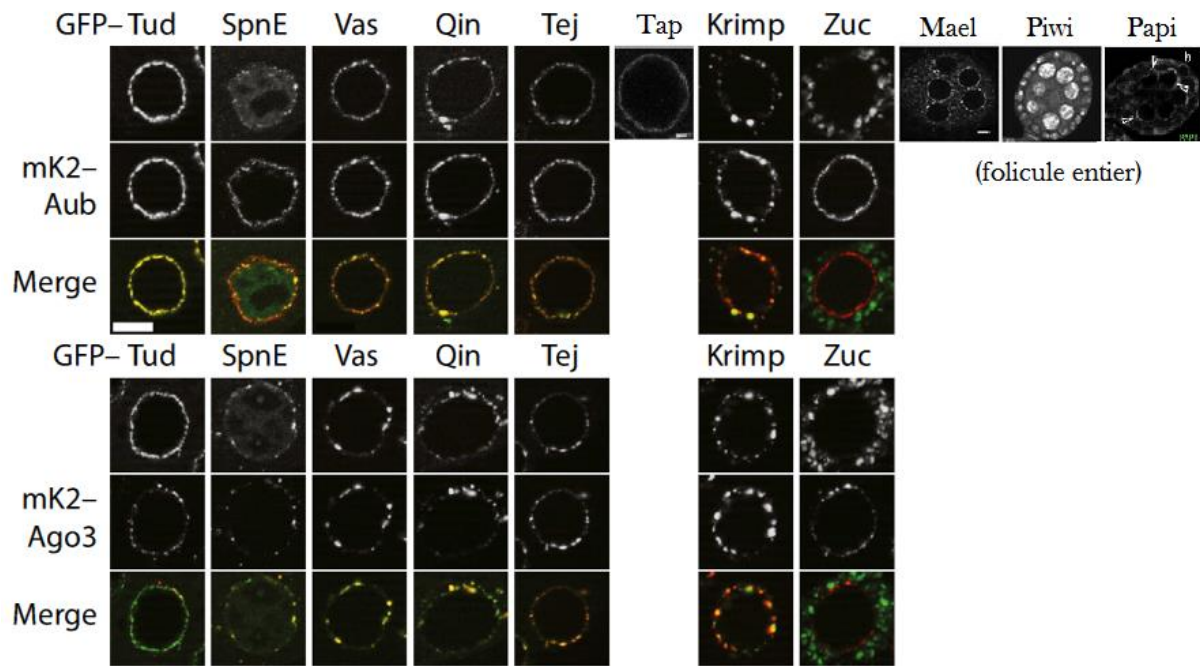
En cellules somatiques non ovariennes, la présence de Rhino recrute UAP56, une DEAD-box hélicase spécialisée dans l'export d'ARNm associés à l'Exon Junction Complex, l'EJC (Gatfield et al. 2001). Dans les cellules germinales, la prise en charge des transcrits de clusters de piARN par UAP56 a pour effet d'éviter l'épissage bien qu'il semble que les dual-strands clusters possèdent une structure exon-intron (F. Zhang et al. 2012; Z. Zhang et al. 2014). Ainsi, les transcrits précurseurs vont directement au pore nucléaire où la nucléopore favorise la libération du transcrit par UAP56 (Montpetit et al. 2011). Les transcrits du cluster de piARN traversent le pore nucléaire et sont associés à Vasa, une autre DEAD-box hélicase (Fig. 19). Les pré-piARN se retrouvent alors dans le nuage où d'autres protéines effectuent les étapes de maturation qui produiront les piARN primaires ou secondaires.



**Figure 19 - Transcription et export des ARNm de différents loci : Dual-strand et Uni-strand clusters de piARN et séquences géniques.** Les transcrits des Dual-strand clusters de piARN échappent à l'épissage dans les cellules germinales grâce aux hélicases UAP56 et Vasa tandis que les transcrits des Uni-strands clusters passent pas une étape supplémentaire de maturation dans le noyau.

## A.2 Le nuage

Le nuage est une structure électroniquement dense qui entoure le noyau des cellules germinales animales. Dans les cellules germinales en développement de *Drosophila melanogaster* le nuage est composé d'agrégats localisés dans le cytoplasme près de la membrane nucléaire (Mahowald 1971). Les Yb bodies, Chromatoïde bodies, pi-bodies, et P-bodies sont des organelles associées au nuage, ces noms sont utilisés en fonction de l'animal et du type de cellules étudiés (pour revue Voronina et al. 2011), les facteurs protéiques présents dans ces différentes structures sont proches (Hirakata et Siomi 2016). Chez la drosophile la majorité des facteurs cytoplasmiques de la voie des piARN y est concentrée (Fig. 20).



**Figure 20 - Nuage** : Colocalisation de différentes protéines de la voie des piARN dans le nuage des cellules nourricières avec Aubergine et Ago3 (ovariole adulte) barre d'échelle 10  $\mu\text{m}$  ((Webster et al. 2015; Anand et Kai 2012; Patil et al. 2014).

### A.2.1 Prise en charge des précurseurs et fabrication des piARN primaires

Une fois que le précurseur de piARN a passé le pore nucléaire, il est pris en charge par l'hélicase Vasa qui les transmet aux nucléases Zucchini (Zuc) (Pane, Wehr, et Schüpbach 2007; Ipsaro et al. 2012), Squash (Squ) (Pane, Wehr, et Schüpbach 2007) et Nibbler (Feltzin et al. 2015). Dans les ovaires, c'est majoritairement la nucléase Zucchini qui clive les précurseurs de piARN. On observe néanmoins une dérégulation des transposons *HeT-A* et *TART* dans ces tissus quand Squ ou Zuc est mutée (Pane et al. 2007). Contrairement à la majorité des membres de la voie des piARN, Zuc se trouve à la surface de la mitochondrie un peu plus loin de la membrane nucléaire (Fig. 20), son domaine nucléase pointé vers le cytoplasme (Ipsaro et al. 2012). Les piARN dont les extrémités viennent d'être découpées par Zuc ne correspondent pas tout à fait aux piARN trouvés dans les banques, leur extrémité 3' doit encore subir des modifications.

Les piARN primaires entrent ensuite dans le nuage et peuvent s'associer avec les protéines Aubergine ou Piwi. On ne connaît pas encore les facteurs qui gouvernent ce choix mais l'hélicase Armitage semble impliquée dans le chargement de piARN primaires par Piwi (Saito et al. 2010). Récemment, il a été montré que Nibbler (Nbr), une exonucléase 3'-5' chargée de raccourcir l'extrémité 3' des miARN chez la drosophile fait la même chose avec

les piARN germinaux. Des séquençages haut-débit de petits ARN montrent que des piARN plus longs (26-28nt) s'accumulent dans les ovaires de mouches mutantes pour *Nbr* tandis que les piARN courts (22-23nt) disparaissent (Feltzin et al. 2015). Ce phénomène n'est alors pas observé chez les piARN de *flamenco* mais H.Wang et collaborateurs le montrent un an après avec leur mutant *nbr<sup>cas9</sup>* qui n'a pas de domaine exoribonucléase (H. Wang et al. 2016). Néanmoins les mutants *Nbr* ne montrent pas de défauts graves du développement de l'ovaire ou de stérilité ce qui explique pourquoi les cribles sur la voie de biogenèse des piARN ne l'avaient pas identifié comme candidat (Czech et al. 2013; Handler et al. 2013; Muerdter et al. 2013). La dernière modification de l'extrémité 3' des piARN est une 2'-O-méthylation catalysée par Pimet, maintenant appelée Hen1 comme chez *A. thaliana* (Saito et al. 2007). Cette méthylation les protège justement contre les exonucléases comme Nibbler. Le mutant *Hen1<sup>cas9</sup>* a exactement l'effet inverse du mutant *nbr<sup>cas9</sup>*, les piARN sont plus courts en 3'. De façon amusante, les doubles mutants montrent des piARN ovariens de longueur normale et la dérégulation des ET observée dans les simples mutants disparaît dans le double mutant (H. Wang et al. 2016). Nibbler raccourcit donc probablement l'extrémité 3' des piARN chargés par une protéine PIWI jusqu'à être gêné par celle-ci, ce qui expliquerait les différentes longueurs de piARN associés à Aub, Ago3 et Piwi.

### ***A.2.2 Boucle d'amplification des piARN : le Ping-Pong***

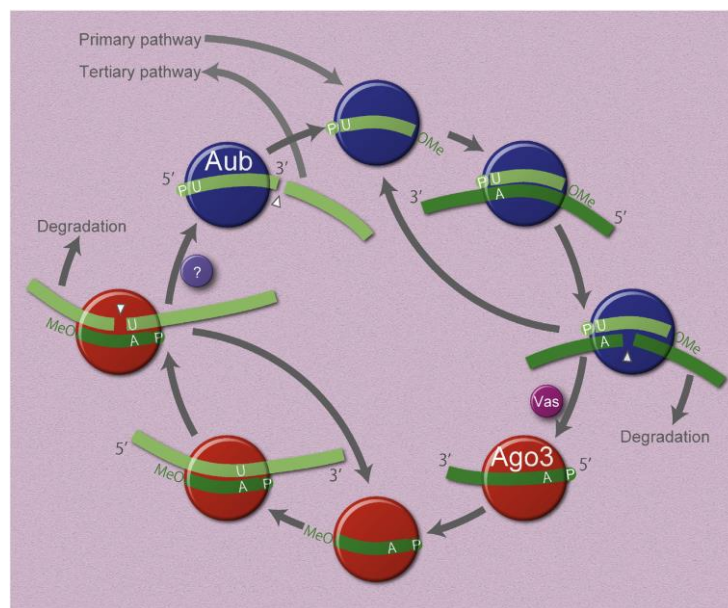
Le devenir des piARN pris en charge par la protéine Piwi sera traité dans la partie suivante, Piwi étant la seule protéine PIWI capable de retourner dans le noyau une fois chargée d'un piARN.

#### ***A.2.2.1 Composition du nuage***

Aubergine est une des 3 principales actrices de la voie des piARN dans l'ovaire adulte (Vagin et al. 2006) et dans les testicules : elle est nécessaire à la répression du *locus Stellate* (A. A. Aravin et al. 2004). Elle est aussi requise pour la formation des cellules germinales embryonnaires (Harris et Macdonald 2001). Aub s'accumule au pôle postérieur de l'ovocyte durant l'ovogenèse et participe à la protection du génome embryonnaire au début du développement ainsi qu'à la régulation de certains gènes (voir Chap.3 A.4.3). Les piARN qui se lient à la protéine Aubergine restent dans le cytoplasme, au niveau du nuage. En effet, des comparaisons GFP-Aub et mKate2-Ago3 (en rouge Fig. 20) ont montré qu'Aub peut être visualisée de façon très diffuse dans le cytoplasme et plus soutenue et homogène dans le nuage, contrairement à Ago3 qui est strictement localisée au nuage et forme des foci distincts.



C'est d'ailleurs grâce à la reconnaissance des ARN cibles par les piARN chargés par Aubergine que celle-ci est recrutée dans le nuage (Webster et al. 2015). Ces piARN sont donc capables de reconnaître le transcrite complémentaire d'un ET avant que celui-ci ne soit clivé par le domaine RNase de la protéine Aub. La coupure à lieu entre les nucléotides 10 et 11 de la partie homologue au piARN, générant une extrémité 5' à 10 nt de distance de son extrémité (Fig. 21). Le transcrite cible sera chargé par la troisième protéine PIWI, Ago3, et coupé à son tour pour donner un piARN secondaire sens, puisqu'il provient directement de l'ARNm d'un ET. Ce piARN sens secondaire (car défini par Aub et Ago3 et non pas Zuc) permettra le clivage d'un nouvel ARN précurseur, probablement par complémentarité avec les insertions antisens de transposons dans les clusters producteurs. On appelle ce phénomène d'amplification le Ping-Pong (Brennecke et al. 2007); Les protéines PIWI avec leur domaine RNase III en sont les principaux effecteurs mais il dépend d'un grand nombre d'autres protéines concentrées dans le nuage et interdépendantes les unes des autres.



**Figure 21 - Modèle du ping-pong chez la drosophile.** Les complexes Aub-piRNA sont générés à partir des transcrits des clusters avec un U en position 1. Ils ciblent les transcrits des ET (en vert foncé) par complémentarité de séquence générant un A en position 10. Le reste du transcrite est alors pris en charge par AGO3 et peut cliver à nouveau un transcrite issu d'un cluster de piRNA générant un U en position 1 qui sera alors être pris en charge par Aub. Ce cycle d'amplification génère ainsi des piRNA dits secondaires n'existant pas dans les clusters de piARN (Hirakata et Siomi 2016).

Tudor (Tud), Spindle-E (Spn-E) et Vasa (Vas), sont localisées comme Aubergine : leur distribution est homogène dans le nuage alors que Qin/Kumo (Qin), Tejas (Tej), Tapas (Tap), Papi et Krimper (Krimp) sont plus concentré sous forme de foyers comme Ago3 (Fig.

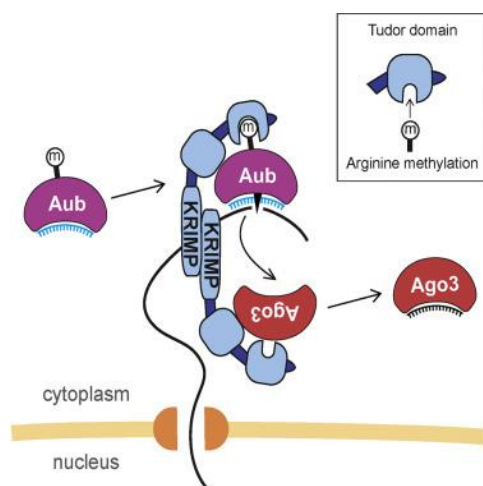


20). Papi, une protéine TUDOR qui interagit avec les trois protéines PIWI, a une préférence pour Ago3 et est nécessaire à sa localisation (Liu et al. 2011). La protéine Mael apparaît sous forme de foyers dans le cytoplasme et de façon concentrée dans le nuage. Enfin, Piwi est très concentrée dans le noyau et plus diffuse dans le cytoplasme (Anand et Kai 2012). Zuc montre une localisation sous forme de foyers plus loin du noyau, dans le cytoplasme (Webster et al. 2015), suggérant que le découpage des précurseurs de piARN a lieu hors du nuage. Les pré-piARN y retournent chargés par Aub ou Piwi pour être maturés.

Spn-E, une hélicase, définit la localisation de Vasa et de Maelstrom (HMG-box et domaine MAEL) dans le nuage. Elle est retrouvée sous forme de complexes avec Qin, Aub et Ago3 et intervient dans le Ping-Pong (Andress et al. 2016). Mael est un composant du réseau de microtubules (Sato et al. 2011). L'analyse de la lignée germinale et somatique ovarienne en condition mutantes pour Mael a montré qu'elle agit à deux niveaux. Mael est impliqué dans la régulation des ET au niveau transcriptionnel, elle ne participe ni à la biosynthèse des piARN ni à la localisation des protéines de la voie des piARN (Sienski et al. 2012). Elle semble également agir au niveau de la production des piARN dans la lignée germinale : la HMG-box pourrait se fixer aux précurseurs de piARN et Mael participerait à l'export vers le nuage (Matsumoto et al. 2015).

#### A.2.2.2 Les protéines Krimper et Qin : gardiennes du Ping-Pong

Les protéines Tudor, Krimp et Qin sont également impliquées dans le Ping-Pong. Si Aub est recrutée dans le nuage par son petit ARN, Ago3 est dépendante de la présence de Krimper pour sa localisation (Z. Zhang et al. 2011; Webster et al. 2015). Krimper fixe une protéine Ago3 non chargée et l'amène dans le nuage où une autre protéine Krimper associée à Aub va former un homodimère et rapprocher les deux protéines (Fig. 22).



**Figure 22 - Détail de la mise en place du Ping-Pong** Krimper forme un homodimère dont les deux partenaires sont associés aux protéines Ago3 et Aub pour favoriser le ping-pong hétérotypique (Webster et al, 2015)

Un des rôles de Krimper est de renforcer le biais antisens des piARN en empêchant Ago3 de se fixer à des transcrits de cluster antisens (Sato et al. 2015). L'action de cette protéine est spécifique au nuage où elle favorise le Ping-Pong hétérotypique Aub/Ago3. Dans le mutant *krimper*, les piARN ne se lient plus à Ago3 et un Ping-Pong homotypique Aub/Aub peut avoir lieu (Sato et al. 2015). Les chercheurs postulent que c'est le rapport de piARN sens/antisens et non seulement l'abondance de piARN antisens (ceux qui sont sensés cibler directement les ARNm d'ET) qui détermine l'efficacité de la répression des transposons. Le Ping-Pong homotypique diminuerait la proportion de piARN antisens et provoquerait donc la dérégulation des ET (Brennecke et al. 2007; Gunawardane et al. 2007; Malone et al. 2009).

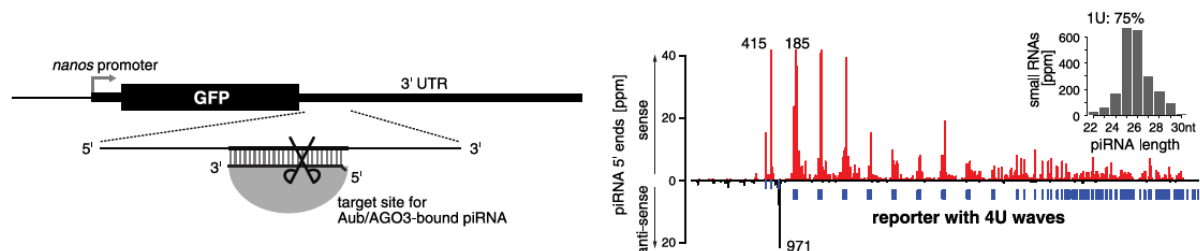
Du côté d'Aubergine, Qin/Kumo, une autre protéine Tudor, agit pour empêcher le couple Aub/Aub de cibler le même transcrit et ainsi conserver un Ping-Pong productif. L'allèle mutant *qin*<sup>1</sup> (insertion d'un transposon *piggyBac* dans le plus grand exon du *CG14303*) cause la stérilité chez les femelles, accompagnée de l'accumulation du variant d'Histone  $\gamma$ H2Av dans les noyaux des cellules germinales (Z. Zhang et al. 2011). Ce variant de l'Histone 2A est associé aux cassures de l'ADN suggérant qu'en absence de Qin, les transposons se mobilisent. En effet, 56 des 93 familles de transposons montrent une expression plus forte en contexte mutant *qin*<sup>1</sup> (1.5 à 3 fois plus). On détecte aussi des défauts de développement embryonnaires chez leur descendance, un phénotype comparable à d'autres mutants de la voie des piARN. Le fait que *Zam* et *Blood* ne sont pas dérégulés dans le mutant *qin* suggère que la protéine est spécifique des cellules germinales (Z. Zhang et al. 2011; Anand et Kai 2012).

Les mutants *qin*<sup>1</sup> ne montrent pas de baisse du niveau global de piARN, c'est la proportion des piARN sens/antisens qui change. Dans les conditions sauvages, les piARN sens représentent 25% des piARN et dans le mutant *qin* cette proportion augmente jusqu'à 41% (Z. Zhang et al. 2011). Les piARN antisens ciblant directement les ARN des ET, leur perte entraînerait donc une dérégulation de ces ET, qui est particulièrement forte pour une partie d'entre eux (*HeT-A*, *Burdock*, *TAHRE*, et l'élément *I*). Les données montrent que Qin agirait pour maintenir cette proportion : dans les ovaires mutants pour *qin* le Ping-Pong Aub/Aub remplace Aub/Ago3 (Z. Zhang et al. 2011). Mais la grosse différence entre les deux études est l'absence de perte des piARN chez Zhang et al, alors que la mutation *kumo*<sup>M41-13</sup> présente une perte d'environ 65% des piARN totaux. 60.2% de réduction des piARN sens

et 87% de réduction des piARN antisens correspondant aux dual-strands clusters disparaissent (Anand et Kai 2012). Des qPCR brin-spécifique montrent la baisse de transcription du *locus 42AB* : la région A présente une diminution de la transcription du brin sens d'un facteur 2.4 et du brin antisens d'un facteur 7.6. Ceci rejoint quand même l'effet sur la proportion des piARN sens/antisens noté chez Zhang. Ils montrent ensuite que la protéine Qin ne s'accroche pas directement aux clusters producteurs mais que la perte de Qin entraîne une plus forte expression de HP1 dans les stades précoces. HP1 est plus présente sur les dual-strand clusters (pas de changement pour les uni-strand) et se lie à Qin physiquement. Selon eux, Qin agirait aussi dans le noyau pour faciliter la transcription des dual-strands cluster en empêchant HP1 de s'y fixer (Anand et Kai 2012).

### A.2.3 Les piARN tertiaires

Une nouvelle source de piARN cytoplasmique a été découverte simultanément par plusieurs équipes en 2015 (Siomi et Siomi 2015; Mohn et al, 2015; W. Wang et al. 2015). Ils ont été appelés piARN tertiaires car ils viennent de transcrits ciblés par des piARN primaires ou secondaires chargés par Aub ou Ago3. Un piARN antisens s'accroche sur un long transcrit cible et le clive. Ensuite, des protéines Piwi prennent en charge le reste du transcrit en se fixant suivant leur affinité pour les U en 5'. Zuc clive alors ce transcrit entre les protéines Piwi. Les piARN produits sont raccourcis par Nibbler, produisant majoritairement des petits ARN de 26nt. Ceci donne un schéma très reproductible de découpage de certains transcrit nommé « phasing » car on peut clairement déduire la position des protéines Piwi sur ces transcrits (Fig. 23). On peut voir ces vagues de piARN produits par le phasing de Piwi (piARN rouges, sens) après un premier découpage par Aub/Ago3 (piARN noir, antisens) sur un transgène rapporteur GFP (Mohn et al. 2015).



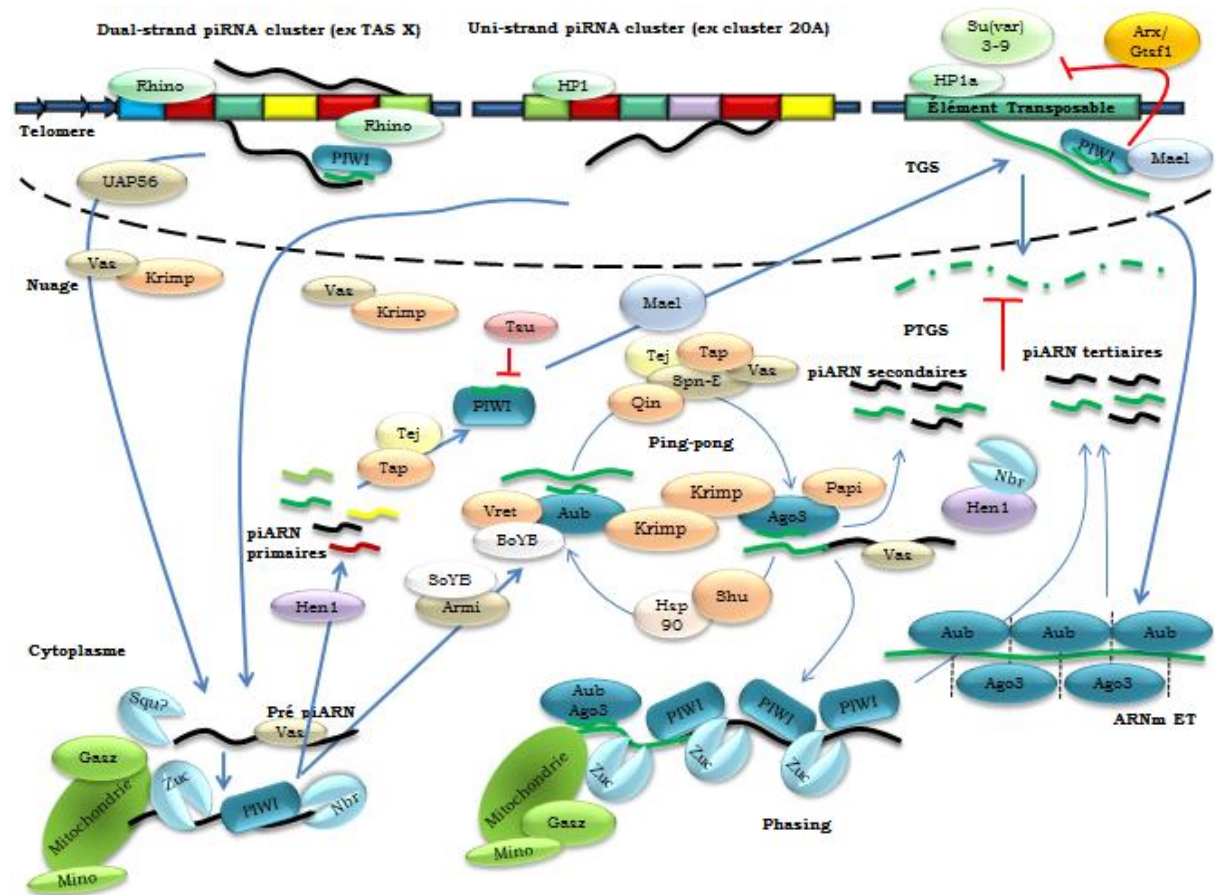
**Figure 23 - phasing**, patron de découpage d'un rapporteur *nosGFP* exprimé dans la lignée germinale et contenant un site de clivage pour Aub ou Ago3. Il est ensuite pris en charge par Piwi et va produire de nombreux piARN sens (rouge) dont la taille dépend de la fréquence de U sur le transcrit (Mohn et al. 2015).

D'après l'équipe de P. Zamore, une fonction d'Ago3 chargé d'un piARN sens, serait de détecter les transcrits de transposons et de produire les piARN antisens chargés ensuite par

Piwi, en plus d'assister Aub dans le Ping-Pong. En effet, ils voient que la perte d'Ago3 entraîne une diminution des piARN liés à Piwi plus importante que la perte d'Aubergine. Qin empêcherait ensuite les produits de clivage d'Aubergine de se lier à Piwi pour conserver des protéines Piwi chargées d'antisens et capables de faire du TGS (W. Wang et al. 2015). Dans cette perspective, les piARN liés à Aub antisens hérités maternellement doivent initier le Ping-Pong, conduisant à l'accumulation de piARN sens liés à Ago3 (Brennecke et al. 2008; Le Thomas et al. 2014). De plus, l'amplification des piARN sens liés à Ago3 permettrait le phasing et donc la production massive de piARN antisens liés à Piwi. L'implication d'Ago3 dans la fabrication des piARN tertiaires permet donc la fabrication de piARN de transposons qui ne seraient pas inclus dans les clusters (W. Wang et al. 2015).

### **A.3 Retour au noyau : le choix de Piwi**

Chez *D. melanogaster* il y a une seule protéine PIWI nucléaire, Piwi, chargée de réprimer les éléments transposables transcriptionnellement, c'est-à-dire en induisant l'hétérochromatinisation des copies actives de ces ET. Les piARN antisens primaires et tertiaires liés à Piwi ciblent les transcrits naissants de copies actives d'ET dans le noyau (Fig. 24). Puisque la localisation nucléaire de Piwi requiert la fixation d'un piARN (Saito et al. 2010; Olivieri et al. 2010), la première étape à lieu dans le nuage.



**Figure 24 - Résumé de la biogenèse piARN germinaux.** Schématisation des trois voies de biogenèse des piARN germinaux : les piARN primaires fabriqués à partir des transcrits des clusters, les piARN secondaires générés par la boucle d'amplification Ping-pong et les piARN tertiaires produits à partir des transcrits des ET.

La protéine Tejas et son paralogue Tapas, localisées dans le nuage (Fig. 20), sont nécessaires à la localisation nucléaire de Piwi dans les cellules germinales. Leur effet est cumulatif : la mutation de chacun de ces deux gènes donne lieu à des ovaires de taille réduite, mais le double mutant est totalement stérile ce qui est probablement dû à la perte des cellules souches germinales par apoptose (Patil et al. 2014). La mutation de *tejas* entraîne la délocalisation de Vasa (Patil et Kai 2010) et le double mutant montre également une délocalisation de plusieurs autres facteurs nuageux de la voie des piARN. De plus, Tap est impliqué dans la répression des ET en lignée germinale mâle et femelle (*TART*, *HeT-A*, *I-element*, *Stellate*, ...). Le modèle actuel postule que Tej et Tap sont impliquées dans le chargement de piARN par Piwi car le double mutant *tej/tap* n'a plus de Piwi nucléaire (Patil et al. 2014).

Asterix/Gtsf1 (protéine à doigts de Zinc CG3893) est une protéine spécifique de la lignée germinale dont la localisation nucléaire dépend de Piwi. Comme Mael, elle ne semble

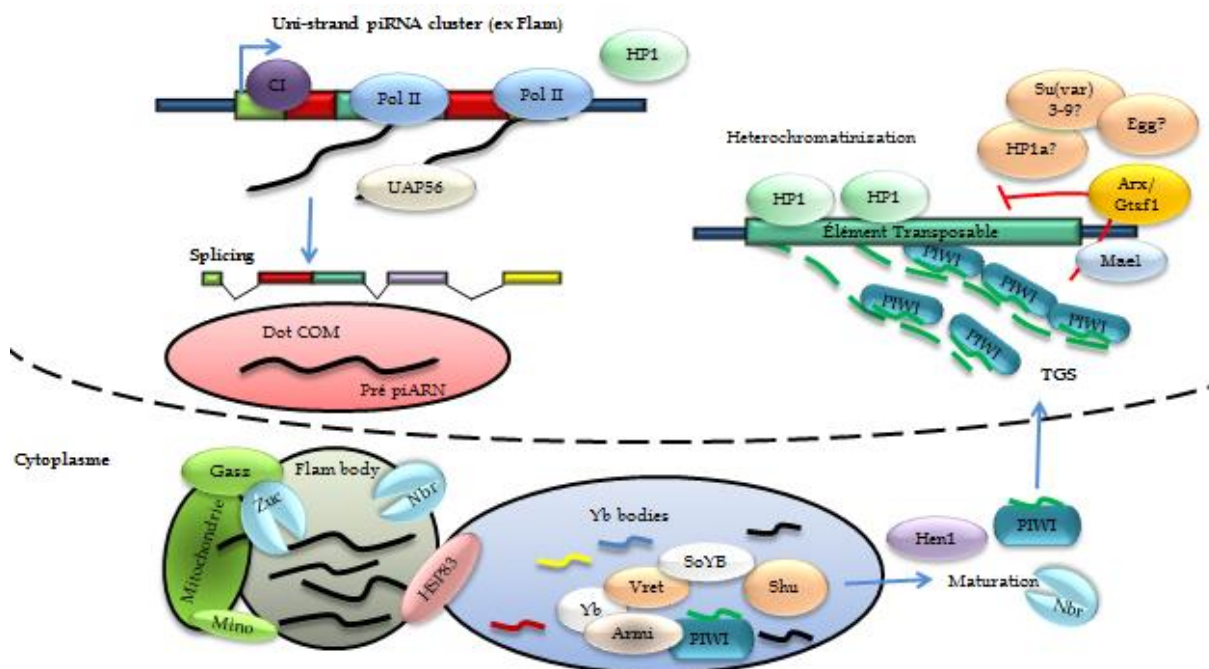
pas être impliquée dans la biogenèse des piARN mais influe sur la répression nucléaire (Dönertas et al. 2013). Dans l'article de Sienski et collaborateurs, les chercheurs montrent que le TGS médié par Piwi dans les cellules somatiques ovariennes est dépendant de la présence de Mael (Sienski et al. 2012). En effet, si l'activité RNase de Piwi n'est pas requise, le domaine Mael est indispensable, peut-être à cause de son activité RNase. Les piARN antisens de *flam* chargés par Piwi guident la protéine vers la copie active d'un ET et Piwi guide l'hétérochromatinisation de sa séquence *via* le dépôt de H3K9me3 par Eggless qui recrute HP1a (S. H. Wang et Elgin 2011). L'interaction directe entre le domaine N-ter de Piwi et HP1a est suggérée par un crible double hybride chez la levure validé sur du lysat d'embryon de Drosophile. La colocalisation a lieu au niveau de *loci* particuliers incluant les éléments transposables *F* et *1360* (Brower-Toland et al. 2007). Mael interviendrait pendant ou après cette étape *via* son domaine MAEL qui pourrait servir de support à Piwi (Sato et al. 2015). L'hélicase Armitage est également nécessaire à la localisation nucléaire de Piwi. Yb recrute le complexe Piwi-Armi dans les Yb bodies (Saito et al. 2010), des structures similaires au nuage germinale formées à la surface des mitochondries et qui contiennent les protéines fs(1)Yb (Yb), Armitage (Armi), Vreteno (Vret) (Zampanini et al. 2011), Shutdown (Preall et al. 2012) et Sister of Yb (SoYb) (Handler et al. 2011). Armitage est donc impliquée dans la biogenèse des piARN primaires (Yb bodies) et dans le TGS (chargement de piARN par Piwi).

#### **A.4 Les piARN somatiques**

##### ***A.4.1 Biogenèse des piARN somatiques ovariens***

Dans les cellules somatiques de l'ovaire chez la drosophile, seule la protéine Piwi est exprimée parmi les membres de la famille PIWI. Les précurseurs de piARN sont transcrits à partir de clusters uni-strand comme le *locus flamenco*. *Flam* héberge un grand nombre de transposons tronqués, dont la plupart sont orientés en antisens par rapport à la transcription du *locus* (Brennecke et al. 2007). Des piARN antisens aux transcrits des ET sont ainsi produits. Les transcrits primaires des clusters sont exportés vers le cytoplasme et éventuellement découpés en fragments de tailles intermédiaires (Fig. 25). Pour *Flamenco*, il a été observé par RNA FISH que les transcrits précurseurs s'accumulent dans le noyau (Dennis et al. 2013). Ce foci a été nommé Dot COM, il concentre également des transcrits précurseurs de plusieurs autres clusters somatiques de piARN et est retrouvé à proximité des Yb bodies dans lesquels ont lieu plusieurs étapes de la maturation des piARN.





**Figure 25- Biogenèse et fonctionnement des piARN somatiques.** Étapes de la production des piARN provenant de clusters somatiques simple brin et protéines impliquées dans leur action au niveau des séquences d'éléments transposables.

Zucchini découpe les précurseurs des clusters exprimés somatiquement dans le cytoplasme avec la participation de Minotaure (Mino) et GasZ localisées à la mitochondrie (Handler et al. 2013; Vagin et al. 2013). Nibbler est probablement aussi présente pour l'ajustement précis de la taille des piARN primaires et Hen1 pour la protection de l'extrémité 3'. Les étapes de maturation ont lieu dans des granules périnucléaires appelés Yb bodies. Yb semble avoir un rôle important de liaison à l'ARN: la présence de cette protéine permet la création de granules appelées Flam bodies qui contiennent les transcrits primaires de *flam* ou leurs intermédiaires (Murota et al. 2014). Ceci permet sans doute de les stocker en prévision de leur maturation en piARN fonctionnels qui a lieu dans les Yb bodies. Toutes ces protéines sont nécessaires à la production de piARN matures et au chargement des piARN dans Piwi (Olivieri et al. 2010). Armi se lie à Piwi et est nécessaire à sa présence dans le Yb body tandis que Yb sert à attirer Armi (Saito et al. 2010). Les piARN chargés par Piwi (pi-Piwi) sont ensuite importées vers le noyau pour réguler la transcription des ET.

Les pi-Piwi nucléaires dirigent le dépôt de H3K9me3 sur la chromatine au niveau des *loci* cibles pour réprimer la transcription des transposons (Klenov et al. 2011; Sienski et al. 2012; Huang et al. 2013). Astérix/DmGTSF1 s'associe avec le complexe pi-Piwi et Mael pour le dépôt de la marque répressive (Muerdter et al. 2013). Enfin, la participation de Egg,

Su(var) 3-9 et HP1a pour cette étape est en cours d'étude (Brower-Toland et al. 2007; Rangan et al. 2011).

#### ***A.4.2 Tissus somatiques non ovariens***

Les piARN retrouvés dans les cellules somatiques ovariennes sont généralement antisens car ils proviennent de clusters uni-strands où les fragments de transposons sont insérés majoritairement dans le sens contraire de la transcription. Mais qu'en est-il des piARN réellement somatiques provenant de tissus n'appartenant pas aux ovaires?

La présence de piARN dans le cerveau de la drosophile a été démontrée. Les neurones  $\alpha\beta$  ne sont pas très concentrés en protéines Aub et Ago3. On observe une expression très forte de nombreux ET qui y sont 5 à 15 fois plus exprimés comparés aux autres types de neurones. Les nombreuses nouvelles insertions des ET observées dans les neurones  $\alpha\beta$  suggèrent que les protéines Piwi et des piARN sont efficaces dans les autres neurones où moins d'insertions sont observées (Perrat et al. 2013). L'existence d'insertions de rétroéléments à LTR dans les gènes impliqués dans la mémoire et l'apprentissage pousse les auteurs à formuler l'hypothèse que l'hétérogénéité résultante aurait une fonction précise dans le cerveau (Perrat et al. 2013). Chez l'Aplysie, une fonction directe des piARN somatiques sur un gène a été trouvée : des petits ARN de 28nt sont présents dans le système nerveux central de ce mollusque gastéropode. Ils ont été classés comme des piARN : 60% présentent un U en 5'. Ils sont 2'O-méthylés, associés à une protéine PIWI exprimée dans le cerveau et montrent une répartition sous forme de clusters dans le génome. Ils auraient une action sur la mémoire grâce à la régulation négative du répresseur de transcription CREB2 en l'absence de Piwi, ce qui empêcherait la « Long-term facilitation ». L'injection d'anti-piARN spécifiques change la méthylation de l'ADN dans le cerveau de l'Aplysie, ce qui a une influence directe sur la transcription des régions touchées (Rajasethupathy et al. 2012).

D'autres modèles animaux ont servi à des analyses plus poussées sur la fonction des piARN somatiques. Chez les planaires, les néoblastes expriment les protéines de la famille PIWI et produisent des piARN. Ces cellules sont impliquées dans les facultés de régénération extraordinaires de ces animaux (Clark et Lau 2014). Chez *Schmidtea mediterranea*, les piARN (25-35nt, associés à SMEDWI-2 et SMEDWI-3 des homologues de PIWI, paires sens/antisens) ne sont pas restreints à la lignée germinale. La déplétion des protéines PIWI SMEDWI-2 et 3 a des conséquences physiologiques : absence de régénération, problèmes de maintiens de l'homéostasie et baisse de la quantité de néoblastes et de piARN. Ces derniers



seraient donc exprimés dans les cellules souches. Certains piARN ne sont produits qu'en condition de stress : 3876 piARN uniques sont retrouvés uniquement dans les animaux en train de régénérer, montrant l'implication de la voie des piARN dans ce phénomène (Palakodeti et al. 2008).

#### ***A.4.3 Les piARN embryonnaires somatiques : un héritage indispensable***

Une autre façon de voir l'impact des piARN sur le soma est l'étude de l'héritage des piARN durant l'ovogenèse : les cellules nourricières sont chargées de produire les piARN et ceux liés aux protéines Piwi et Aub sont accumulés dans l'œuf au pôle postérieur (Harris et Macdonald 2001). Au début du développement, les gènes ne sont pas exprimés par l'embryon lui-même. C'est la traduction des ARN messagers maternels accumulés lors de l'ovogenèse qui va assurer les fonctions essentielles durant les premières heures. Le contrôle du développement passe ensuite du génome maternel au génome zygotique lors de la « Maternal-to-Zygotic transition », la MTZ. Les messagers maternels sont alors dégradés activement et ceux du zygote commencent à être produits. Cette dégradation passe par le raccourcissement de la queue poly-A des messagers par la déadénylase CCR4, ceci déstabilise l'ARN et le conduit à la dégradation. Pour recruter CCR4, différents acteurs entrent en jeu, dont les piARN (Rouget et al. 2010). Le laboratoire de M. Simonelig a montré qu'en conditions mutantes pour Aub, Piwi, Ago3 ou Spindle E, l'ARNm de *nanos* n'est pas dégradé dans la partie somatique de l'embryon. Ils ont mis en évidence deux petits ARN correspondants aux ET *roo* et *412*, chargés par Aub qui interagissent directement avec le 3' UTR de l'ARNm de *nanos*. Lorsque cette interaction est perdue, grâce à la délétion des séquences d'interaction des deux piARN sur l'ARNm de *nanos*, celui-ci n'est plus dégradé, entraînant de graves défauts dans le développement de la tête. Ceci implique les piARN dans le développement somatique embryonnaire. En effet, des expériences récentes de iCLIP (Immunoprécipitation de la protéine Aub en compagnie des ARNm qu'elle cible) ont montré qu'Aubergine chargée d'un petit ARN s'associe avec des ARNm dans l'embryon précoce (Barckmann et al. 2013; Barckmann et al. 2015). 0 à 2h après la ponte (AEL pour After Egg Laying), 182 d'entre eux sont stabilisés dans le mutant Aub et seraient donc dégradés spécifiquement dans le soma suite à leur ciblage par un piARN chargé par Aub (Barckmann et al. 2015). Ils sont liés à des fonctions germinales (enrichissement en termes Gene Ontology tels que « germ cell development » ou « cellular processes involved in reproduction ») et grâce à leur dégradation dans le soma, ils sont uniquement présents au niveau du pôle postérieur de l'embryon. Si ces transcrits sont dégradés dans le soma, un facteur supplémentaire doit intervenir dans ce

processus pour les protéger dans le pôle postérieur. Les auteurs font l'hypothèse que cette dégradation sélective est due à une complémentarité imparfaite entre le piARN et l'ARNm ou à la présence d'un ou plusieurs autres acteurs protéiques. Sur le modèle de la régulation de *nanos*, le complexe incluant la déadénylase CCR4 pourrait être absent de la future lignée germinale. La régulation négative du gène *nanos* par les piARN en embryon précoce (Rouget et al. 2010) montre l'influence que peut avoir la voie des piARN sur les gènes. En outre, les piARN couplés à Aubergine ont été détectés dans le plasma germinal où ils interagiraient avec les 3'UTR d'ARNm en compagnie de la protéine Tudor pour les séquestrer au pôle postérieur. Cette séquestration des ARNm pourrait être nécessaire à la formation des cellules germinales (Vourekas et al. 2016)



## **B. Autres sources de piARN dans le génome**

### **B.1 Les piARN géniques**

Tous les piARN ne proviennent pas de clusters. Des piARN correspondants à des ARNm de gènes ont été observés dans les cellules germinales et somatiques. Les ARNm ne sont habituellement pas dirigés vers la machinerie des piARN. Néanmoins, la perte de certains facteurs de la voie des piARN peut entraîner une erreur d'aiguillage. La mutation des gènes *UAP56* et *Rhino*, par exemple, donne lieu à des petits ARN euchromatiques venant principalement de séquences 3'UTR de gènes (Z. Zhang et al. 2014).

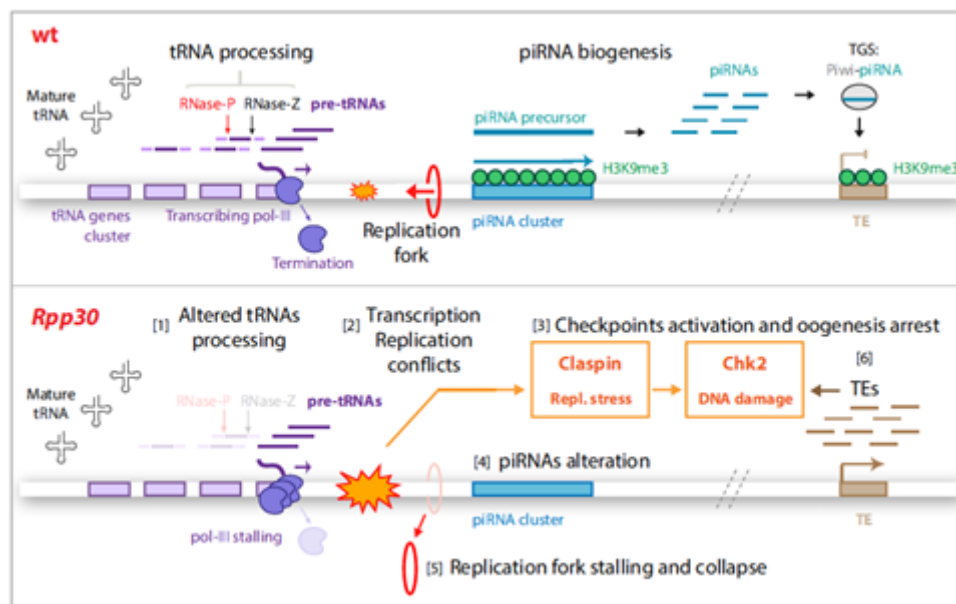
Mais il arrive parfois qu'on trouve de tels piARN géniques en contexte sauvage. Un exemple de piARN génique a été documenté en 2009 dans des OSS. Le gène *traffic jam (tj)* produirait des piARN sens venant de son 3' UTR. Ce long gène de la famille *MAF*, orthologue des facteurs de transcription *c-Maf* et *MafB/Kreisler* chez les vertébrés, produit un facteur somatique spécifique contrôlant la morphogenèse des gonades et nécessaire pour la fertilité chez les femelles et les mâles (Saito et al. 2009). Les piARN issus de ce 3'UTR viennent du brin sens et ne peuvent donc cibler le transcrit lui-même, mais l'analyse du génome de *D. melanogaster* a montré une complémentarité entre ces piARN et ceux d'un autre ARNm *FasciclinIII (FasIII)*. Ces piARN sont essentiels et ont une importance fonctionnelle : dans les mutants *tj* comme *piwi*, Fascicline III est surexprimée et les cellules somatiques ne parviennent pas à se mêler aux cellules germinales. De plus la perte de Tj abolit l'expression de Piwi dans les cellules somatiques des gonades. Ce gène aurait donc deux fonctions distinctes : la protéine Tj qui active l'expression de Piwi et les piARN guident Piwi vers le gène *fas III* (Saito et al. 2009).

### **B.2 Connexions entre les différentes voies de biosynthèse des petits ARN interférents**

Les ARN de transfert (ARNt) peuvent augmenter directement la quantité de piARN. Par exemple l'ARNt glu-CTC donne naissance à une paire de piARN très exprimés (Senti et al. 2015). L'extrémité 5' du premier piARN est coupée par la RNase P et est chargée par Aub tandis que le second piARN, immédiatement adjacent au premier piARN, est pris en charge par Piwi. Ils sont probablement reconnus par Aub grâce au 5' monophosphate de l'ARNt et clivés par Zucchini, leur présence étant dépendante de celle de la nucléase.

Une conséquence indirecte de la synthèse des ARNt sur la production des piARN a été découverte en 2015 (Molla-Herman et al. 2015). Un crible EMS a permis d'isoler une lignée donnant des mouches viables, aux ovaires atrophiés et stériles. La mutation affecte le gène *Rpp30* qui produit une sous-unité de la RNase P, un ribozyme impliqué dans le clivage du côté 5' des pré-ARNt. Le mutant *Rpp30*<sup>18.2</sup> présente un arrêt de l'ovogenèse au stade 3-4 : cette catégorie de phénotype est fortement associée à des mutations dans la voie des piARN. *Rpp30* semble essentielle dans la lignée germinale adulte (l'expression de *Rpp30*-GFP rétablit l'ovogenèse), sa mutation entraînant l'accumulation de pré-ARNt conservant leur région 5'. Paradoxalement, le niveau d'ARNt matures est normal dans les mouches mutantes. Le phénotype n'est donc pas lié à un manque d'ARNt. Comme pour la mutation de certains acteurs de la voie des piARN (*spindle-E* ou *armitage* Chap. 1 B.3.4) l'absence de *Rpp30* activerait les points de contrôle de dommage à l'ADN qui entraînent des défauts de développement : l'utilisation d'un double mutant *Rpp30*<sup>18.2</sup> avec *mnk*<sup>p6</sup> ou *chk2* restaure le phénotype sauvage (Molla-Herman et al. 2015).

Les auteurs relient l'encombrement causé par la transcription par l'ARN Pol III des ARNt, qui sont "clustérisés" dans le génome, à la gêne de la réplication de l'ADN. Cela déstabiliserait la fourche répllicative et induirait des lésions de l'ADN activant les points de contrôle (Clelland et Schultz 2010). De façon intéressante, la transcription d'au moins 30 clusters de piARN est fortement dérégulée dans les ovaires mutants pour *Rpp30*, or les clusters d'ARNt et certains *loci* producteurs de piARN sont très proches dans le génome (Fig. 26). Le stress réplificatif aurait donc un effet négatif sur la transcription de ces derniers. La baisse de production des piARN entrainerait alors la dérégulation des ET causant encore plus de dommages à l'ADN (Molla-Herman et al. 2015).



**Figure 26 - Mise en parallèle du mécanisme de fabrication des ARNt dans des ovaires sauvages et mutants pour *Rpp30*, et modélisation de la cascade d'évènements entraînant la dérégulation des piARN :** les gènes des ARNt (violet) sont transcrits par l'ARN pol III et leurs transcrits sont maturés par les RNases P et Z. Lorsque des clusters de gènes des ARNt sont localisés à proximité de clusters de piARN (turquoise), un conflit entre l'ARN pol III immobilisée par la mutation de *Rpp30* et de l'ADN polymérase cause un stress de la réplication. Les points de contrôle des dommages à l'ADN (boîtes jaunes) sont alors activés, et la fabrication des piARN perturbée. En conséquence, les ET (brun)ne sont plus réprimés (Molla-Herman et al. 2015).

Finalement, les auteurs proposent un modèle évolutif qui énonce que la proximité entre les clusters de gènes d'ARNt et de piARN serait conservée à travers le vivant car elle faciliterait l'insertion de copies d'ET à ADN dans les clusters de piARN, les incluant d'office dans le catalogue des piARN afin d'induire leur répression. Cette insertion aurait lieu grâce à la perte de vitesse de la fourche de réplication due à la présence des gènes d'ARNt en amont de ceux-ci. Il a été montré que l'élément *P* par exemple, favorise les origines de réplication (ORI) pour s'insérer. Cela lui permettrait d'augmenter son nombre de copie s'il parvient à le faire au niveau de multiples ORI en même temps (Allan C. Spradling et al. 2011). Ceci expliquerait donc en partie pourquoi on trouve des catalogues d'ET fragmentés dans les régions difficiles à répliquer, comme les centromères et les télomères (Molla-Herman et al. 2015).

Des connexions entre la voie des miARN et des piARN ont été remarquées : dans les cellules germinales, le 5'UTR de l'ARNm du gène *pasha* mime la structure d'un pré-miARN. La configuration en épingle à cheveux partiellement complémentaire de cet ARN donne lieu à deux piARN reconnus par Aub et Piwi. La délétion de l'une ou l'autre de ces protéines

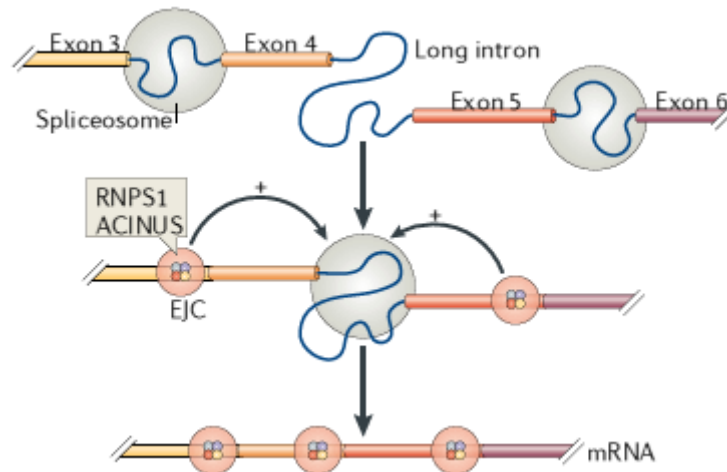
entraîne la perte du premier ou du deuxième piARN respectivement tandis que l'absence d'Ago3 n'a aucun effet. Zucchini est également nécessaire à la présence des deux piARN (Senti et al. 2015). Dans les cellules germinales, un grand nombre d'ARN messagers donnent lieu à des piARN qui ne sont pas restreints aux 3'UTR de gènes : l'analyse de données de Deep-seq montre que la reconnaissance d'un site de coupure dans un ARNm entraîne ensuite la production de piARN tertiaires par phasing à partir du reste du transcrit ciblé (Mohn et al. 2015).

### **B.3 Impact de l'Exon Junction Complex sur la voie des piARN**

L'Exon Junction Complex (EJC) est un complexe protéique hautement conservé qui lie l'ARN pendant l'épissage et y reste associé durant l'export vers le cytoplasme (pour revue (Le Hir et al. 2016)). Il aide à la maturation des ARN messagers en favorisant l'épissage des pré-ARNm, processus qui élimine les introns des transcrits primaires et lie les exons adjacents. Il est déposé sur les messagers environ 20 à 24 nucléotides en amont des jonctions exon-exon après l'épissage et reste associé lors de l'export nucléaire pour agir lors d'événements ultérieurs. Le cœur de l'EJC est composé d'eIF4AIII (DEAD-box hélicase ARN) qui s'accroche directement à l'ARN au début de l'épissage, Mago Nashi et Tsunagi qui sont recrutés durant l'excision (Ballut et al. 2005).

L'absence d'expression des protéines Mago et Tsunagi n'entraîne pas de baisse de la quantité de AGO3 ni de Aub mais ces protéines semblent partiellement relocalisées dans le cytoplasme ce qui suggère une modification de la composition du nuage. En revanche la quantité de protéine Piwi diminue fortement, en germinale comme en somatique (Sienski et al. 2012). De nombreuses autres protéines interagissent de manière transitoire avec l'EJC pour effectuer des régulations post-transcriptionnelles supplémentaires (Tange et al. 2005). La sous-unité RnpS1 (protéine de liaison à l'ARN) nous intéresse pour son rôle dans la voie des piARN, elle est nécessaire chez l'adulte à la fois dans les cellules germinales adultes et dans les cellules folliculaires. En effet, en l'absence de RnpS1, la quantité de protéine Piwi chute, ce qui entraîne une dérégulation massive des TE en même temps qu'une baisse globale des piARN (Malone et al. 2014; Le Hir et al. 2016). Le ping-pong ne semble pas affecté. En revanche, le niveau de transcrits de *piwi* lui-même montre une baisse de 40% à 80% (RT-qPCR) avec un effet plus important observé sur l'amplicon couvrant les exons 4 et 5 du transcrit (Fig. 27). Lorsque le 4ème intron de l'ARNm de *piwi* est retenu cela engendre une baisse du niveau de production de protéines Piwi correctement conformées et donc l'arrêt de

la régulation des transposons dans les gonades. *piwi* est l'ARNm le plus sensible à une délétion de membres de l'EJC mais *mei-P26* (un suppresseur de tumeur), *eIF4G2* (facteur d'initiation de la traduction), *Nipped-A* (histone acétyltransférase), *CG6479* et *CG9870* sont également sensibles à sa disparition (Malone et al. 2014).



**Figure 27 - La fonction du complexe de jonction d'exon dans l'épissage de l'ARNm de *piwi*.** Chez *D. melanogaster*, l'épissage correct du pré-ARNm de *piwi* est dépendant de l'EJC. Sa présence au niveau des jonctions intron/exon voisines peut faciliter (+) l'épissage d'un intron "faible" ce qui est le cas de l'intron 4 de *piwi*, particulièrement long (Le Hir et al, 2016).

Il est intéressant de noter que des mutations de *mago nashi* et de *tsunagi/Y14*, conduisent à des défauts de développement comparables à ceux générés par la mutation de *UAP56* et d'autres mutants de la voie des piARN. Cela peut refléter une fonction pour l'EJC, avec *UAP56* chargée de l'export des précurseurs de piARN dans le noyau (F. Zhang et al. 2012) ou via *Piwi* directement. Enfin, durant l'ovogénèse, *Tsunagi*, *MagoNashi* et *Ranshi* restreignent l'identité de l'ovocyte à une seule cellule dans les chambres à œufs (Lewandowski et al. 2010) et aident à la localisation de l'ARNm *oskar* (Mohr et al. 2001).





## Chapitre 4 : La Lignée Germinale, un Bastion à Protéger

### A. Description de la lignée germinale femelle chez la drosophile: de l'œuf à la mouche.

#### A.1 Mise en place des territoires embryonnaires

##### A.1.1 Polarité de l'embryon

Immédiatement après la ponte de l'œuf, le noyau du zygote commence à se diviser pour former l'embryon syncytial. 13 cycles de divisions sont observés avant le début de la formation des membranes, la cellularisation (Zalokar 1976). Les gradients d'ARNm et de protéines déposés par la mère lors de l'ovogenèse vont jouer un rôle important dans la mise en place des territoires de l'embryon : en particulier ceux déposés aux deux extrémités de l'œuf. Les ARNm de *bicoid* sont déposés au pôle antérieur de l'embryon, un gradient de protéine est formé, permettant le développement de la tête et du thorax en activant la production du gène *hunchback* dans les noyaux de cette région (Driever et Nüsslein-Volhard 1989). Pour le développement de l'abdomen de l'autre côté de l'embryon, les ARN de *nanos* vont s'accumuler au pôle postérieur où la protéine va réprimer *hunchback* (Tautz 1988; Tautz et Pfeifle 1989). E. Gavis et R. Lehmann ont montré en 1992 que le 3'UTR de l'ARNm *nanos* est suffisant pour sa bonne localisation, comme pour *bicoid* au pôle antérieur. Elles ont remplacé la région 3'UTR de *nanos* par celle de *bicoid* et observé des embryons avec un abdomen à chaque pôle (Gavis et Lehmann 1992). La localisation différentielle de ces ARNm permet l'activation des gènes nécessaires à la segmentation de l'embryon le long de l'axe antéro-postérieur. Vasa a une action directe sur le patterning de l'embryon. Elle active la transcription de *gurken* (*grk*) dans l'embryon via son domaine DEADbox et son C-ter acide (acidic C-ter), ce qui permet la différenciation de l'œuf le long de l'axe dorso-ventral. Ses régions N-ter et C-ter sont importantes pour sa localisation et sa partie N-ter est indispensable pour la formation des Cellules Germinales Primordiales, les PGC (Dehghani et Lasko 2015).

Pendant ma thèse je me suis plus particulièrement intéressée au pôle postérieur, où se forment les PGC. L'accumulation de l'ARNm *nanos* dépend d'un nombre important de gènes. Les travaux de R. Lehmann, Wieschaus et T. Schupbach ont catalogué Oskar, Tudor et Vasa comme composants du plasma germinale et Cappuccino, Spire et Staufén comme responsables de la localisation de ces derniers au pôle postérieur, entre autres grâce à leur action sur le

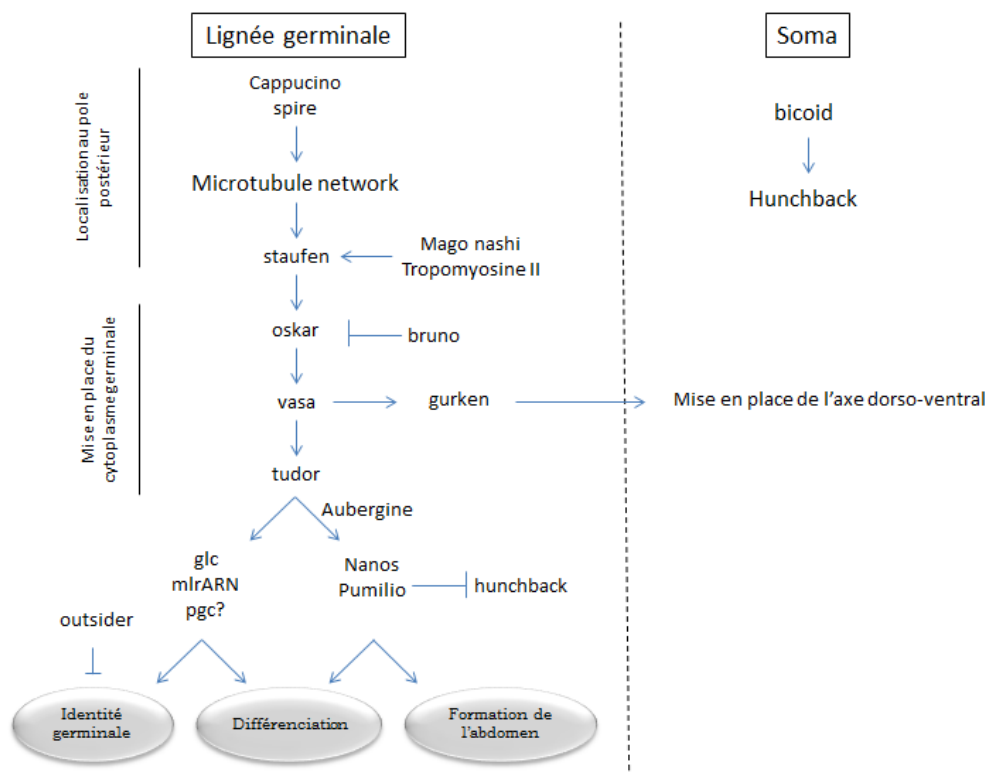
cytosquelette (pour revue (Ephrussi et Lehmann 1992)). Les deux répresseurs traductionnels Nanos (nos) et Pumilio (Pum), qui répriment l'expression de diverses protéines dans les cellules souches sont particulièrement importants. Nos et Pum recrutent CCR4 sur des ARNm spécifiques, comme celui de *mei-P26* (Joly et al. 2013).

### ***A.1.2 Définition du cytoplasme germinale***

Les premières phases du développement embryonnaire se produisent en l'absence de transcription *de novo* et sont commandées par des ARNm et des protéines d'origine maternelle. Au cours de cette période, les cycles cellulaires sont synchrones et rapides, ce qui empêche les gènes zygotiques particulièrement longs d'être transcrits (M. T. Lee et al. 2014). En effet, les gènes transcrits précocement ont tendance à être à courts. Par exemple, *knirps* est transcrit tôt en raison de sa petite taille (3 kb) alors qu'un de ses homologues de 20kb montre une transcription avortée (Rothe et al. 1992). On observe aussi la transcription défectueuse de *ultrabithorax* (77 kb) suivie de sa dégradation jusqu'au cycle 14 (Shermoen et O'Farrell 1991). La formation de l'hétérochromatine semble également jouer un rôle très précocement dans l'embryon, à partir du moment où les noyaux embryonnaires deviennent transcriptionnellement actifs. HP1a n'est pas présente dans les noyaux de l'embryon syncytial mais se met en place au cycle 10 (James et al. 1989). Dans tout l'embryon, les marques activatrices de la transcription H3K4me3 ne sont pas présentes avant la cellularisation, même parmi les gènes exprimés précocement. Le relais est donc pris par le zygote vers le stade 14/15 dans le soma (Lu et al. 2009) mais qu'en est-il des cellules germinales ?

Le plasma germinale est un cytoplasme spécialisé localisé à l'extrémité postérieure de l'œuf pendant l'ovogenèse. Il gouverne la transcription à cette période du développement. Il est particulièrement riche en protéines Vasa, Oskar, Tudor et Aubergine, et en ARNm d'*oskar*, *nanos*, *polar granule component (pgc)* et *germ cell-less (gcl)* (Schéma récapitulatif Fig. 28). La localisation de ces ARNm est également permise par leur séquence 3'UTR (Ephrussi et Lehmann 1992) qui interagit avec la chaperonne Hsp83, l'homologue de Hsp90 chez la drosophile (Y. Song et al. 2007). Dans une étude récente, la structure de la protéine Oskar a été dévoilée : son domaine N-ter LOTUS permettrait la dimerisation et des interactions avec Vasa (*in vitro*) tandis que le domaine C-ter permettrait l'interaction avec des ARNm (Jeske et al. 2015), comme ceux de *nos*, *gcl* et *pgc* qui sont connus pour interagir avec Oskar. Valois (Vls) est requis pour la localisation de Tudor (Tud) dans le nuage et au pôle postérieur de l'embryon *via* son interaction avec Oskar (Anne 2010). Ces transcrits et

protéines serviront à définir l'identité des cellules germinales primordiales et à assurer le maintien de leur état indifférencié jusqu'à la prise de relai par le génome zygotique.



**Figure 28 – Schéma de la spécification du plasma germinale.** Le diagramme montre les interactions génétiques décrites dans cette partie. Les flèches indiquent une influence sur l'activation ou la localisation, mais n'impliquent pas une interaction directe entre les protéines. Les ovals indiquent le destin cellulaire. Abréviations: mlrRNA, mitochondrial large ribosomal RNA (Tautz 1988; Ephrussi et Lehmann 1992; Rongo et al. 1997; Anne 2010; Barckmann et Simonelig 2013).

Le fait que les cellules germinales sont fortement réprimées transcriptionnellement est acquis depuis 1976, avec les travaux de Zalokar et Erk. Dans leur étude, la synthèse *de novo* d'ARN et de protéine dans des embryons précoces de drosophile a été étudiée avec des nucléotides et des acides aminés radioactifs (Zalokar 1976). Les résultats de ces expériences ont montré que les cellules somatiques des embryons incorporent une quantité importante d'uridine<sup>[3H]</sup> (un précurseur de l'ARN), à partir de la migration à travers l'intestin moyen. Aucune transcription n'est détectée avant les stades 9/10. En revanche, la synthèse de protéines a été montrée dans les cellules polaires.

Il semblerait que la répression transcriptionnelle soit un phénomène actif puisque l'introduction d'un pilote artificiel comme le GAL4VP16 n'est pas suffisant pour activer la transcription de transgènes UAS avant les premiers stades larvaires (M. Van Doren,

Williamson, et Lehmann 1998a). Seuls certains ARNm spécifiques y échappent. Les ARNm de *pgc*, initialement caractérisés comme non-codants, sont également exprimés dans les cellules germinales en formation dès le stade 4 (Nakamura et al. 1996). Ils produisent un petit peptide de 71 acides aminés exprimé à partir du stade 6 qui réprime la phosphorylation du domaine C-Terminal (CTD) de l'ARN pol II (au niveau de la Sérine 2) et donc la transcription dans ces cellules (Hanyu-Nakamura et al. 2008). *gcl*, dont les ARNm sont déposés au pôle postérieur durant l'ovogenèse (Jongens et al. 1994; Leatherman et al. 2002) pourrait également être impliqué dans cette répression transcriptionnelle. La présence de *gcl* est cruciale pour la formation de PGC au pôle postérieur mais son induction ectopique n'est pas suffisante pour créer des cellules germinales. Après cette période, la transcription zygotique commence : l'influence maternelle sur le développement commence à diminuer, les facteurs de transcription déposés maternellement sont dilués et le ralentissement des cycles de réplication permet la transcription des longs gènes zygotiques (Lu et al. 2009; M. T. Lee et al. 2014).

## **A.2 Développement des gonades**

De nombreuses revues ont été écrites sur le sujet. Dans cette partie je ferai une synthèse de Santos 2004, Kunwar 2006 et Lesch 2012 pour l'embryogenèse et des travaux et revues de L. Gilboa pour le développement des gonades larvaires femelles.

### ***A.2.1 Formation des cellules germinales primordiales***

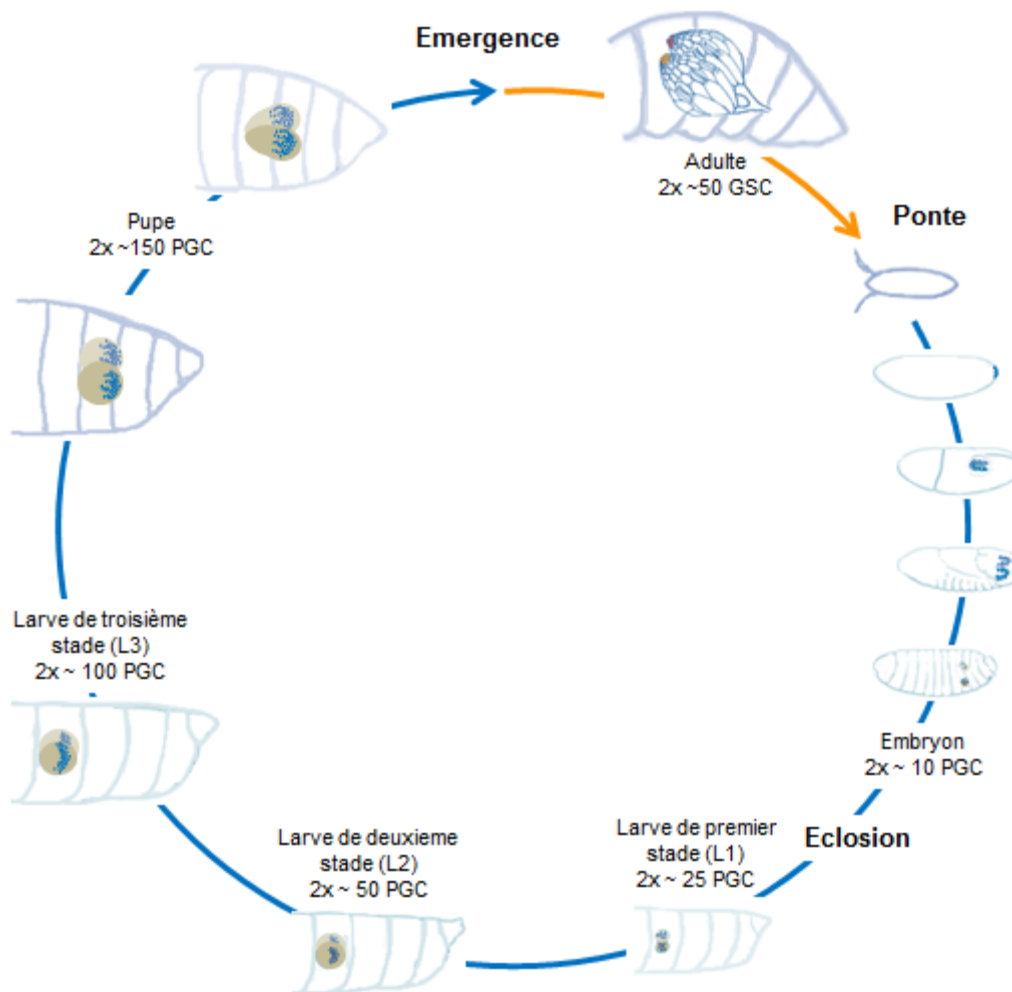
Encore aujourd'hui le mécanisme de formation des PGC est nébuleux. Les noyaux qui se dirigeront vers le pôle postérieur deviendront ceux des cellules germinales primordiales (les PGC). Celles-ci sont les premières à se cellulariser 90 minutes après la fertilisation. *gcl*, *pgc*, *pumillio* et *nanos* maintiennent les PGC embryonnaires transcriptionnellement inactives pendant la première partie de leur migration (Williamson et Lehmann 1996) et la transcription de *outsider* dans celles qui se perdent en route permet leur élimination (Coffman et al. 2002). Un des rôles de Piwi est la maintenance des cellules germinales adultes pour les deux sexes, elle permet la répression des rétrotransposons LTR dans les précurseurs des cellules germinales mâles (Kalmykova et al. 2005). Chez la femelle, on retrouve dans les mutants *piwi* un petit nombre de gamètes (environ égal au nombre de cellules souches formées à la fin de la métamorphose) et plus aucune cellule souche (Lin et Spradling 1997). On ne sait pas si ces fonctions nécessitent la présence d'un petit ARN.

### **A.2.2 Embryogenèse**

Une fois formées, les PGC se différencient des cellules somatiques : environ trois heures après la ponte de l'œuf, elles commencent leur migration. D'abord de façon passive : elles s'accrochent à l'intestin moyen puis le traversent grâce à la dissolution partielle des connections entre les cellules épithéliales (Callaini et al. 1995; Jaglarz et Howard 1995). Elles migrent ensuite dans le mésoderme, à ce stade, la plupart des transcrits maternels sont rapidement dégradés et la transcription zygotique des PGC commence au stade 9 (Seydoux et Dunn 1997), soit une heure plus tard que dans les cellules somatiques de l'embryon (Simonelig 2012). Des gènes zygotiques régulent ensuite la migration chimiotactique des PGC dans l'embryon : *wunen* et *wunen-2* repoussent les PGC loin des tissus adjacents (ils sont exprimés dans le système nerveux central, l'ectoderme et l'intestin moyen ventral). Elles sont alors attirées vers le mésoderme par *Hmgcr*, *fpps*, *columbus*, *quema* et  $\beta$ -*ggt1* (pour revue voir (Santos et Lehmann 2004; Kunwar, Siekhaus, et Lehmann 2006; Lesch et Page 2012)).

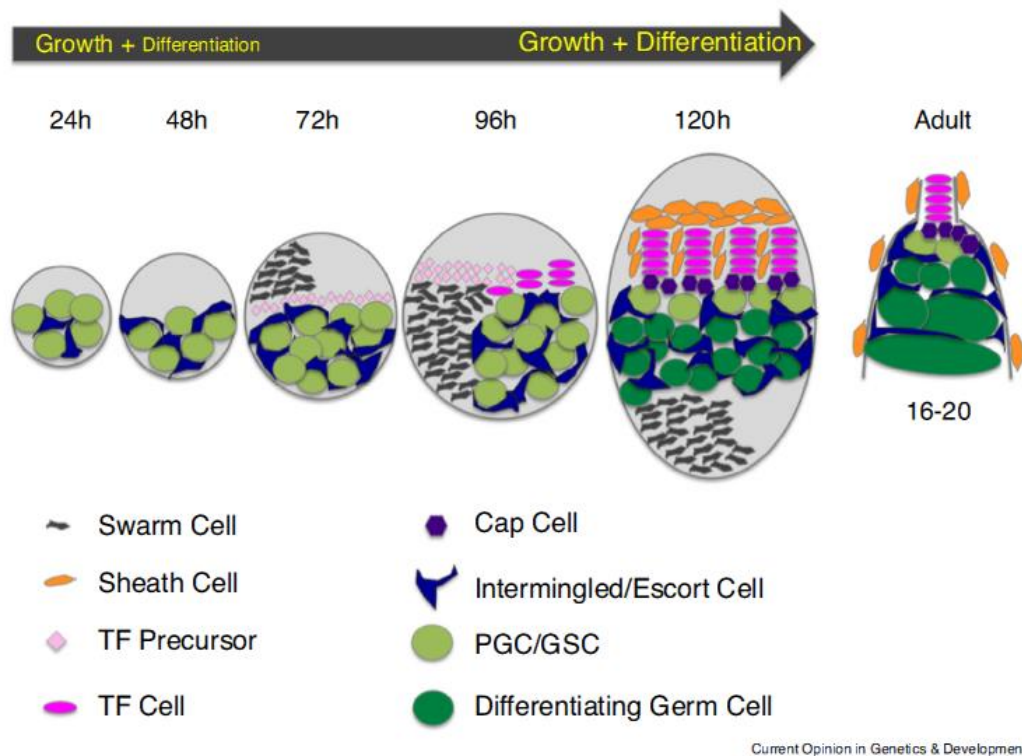
### **A.2.3 Organogenèse de la gonade larvaire**

Quand les PGC arrivent dans le mésoderme, elles s'agrègent avec les « mesodermal somatic gonadal precursor cells » (SGP) qui expriment *HMGC<sub>o</sub>Ar* ce qui stoppe leur migration. Leur morphologie change, elles deviennent rondes et non mobiles puis les deux types de cellules forment une paire de gonades compactes : c'est la coalescence (Jaglarz et Howard 1995) dirigée par *E-Cadherin/shotgun* (*shg*) et *fear of intimacy* (*foi*) (Jenkins et al. 2003; M. Van Doren et al. 2003). L'embryon est alors au stade 14. À ce moment, l'identité sexuelle des cellules germinales est déterminée : les testicules ou les ovaires embryonnaires sont établies (Murray et al. 2010). Dans les testicules embryonnaires, un petit sous-ensemble de PGC donne lieu aux cellules souches adultes (les GSC) de la lignée germinale mâle et occupe la niche de cellules souches, le hub (Le Bras et Van Doren 2006). Dans la larve femelle, les PGC subissent quatre mitoses asynchrones et forment des ovaires larvaires composés d'environ 150 PGC et plusieurs types de cellules somatiques (Gilboa et Lehmann 2006). Dans ces travaux je me suis intéressée exclusivement à la lignée germinale femelle (Fig. 29). D'une part, parce que la femelle produit et transmet l'information nécessaire à la répression des ET dans sa descendance, qu'elle soit mâle ou femelle. D'autre part, parce que le système de read-out de la répression par les piARN que nous utilisons, le TSE, ne fonctionne pas chez le mâle.



**Figure 29 – Schéma du développement de la lignée germinale femelle chez *D. melanogaster*.** Les PGC sont formées au pôle postérieur de l'embryon. Elles migrent durant la gastrulation et se séparent en deux groupes pour former les gonades qui contiennent également des cellules somatiques. Les cellules germinales subissent environ 4 mitoses asynchrones durant les 3 phases larvaires (L1, L2 et L3) pour donner des gonades pupales contenant approximativement 150 PGC.

Au début de la vie larvaire, (stades L1-L2), les cellules germinales et somatiques subissent une phase de prolifération. Les PGC sont entourées de cellules somatiques dès le stade L1 : les Intermingled cells (IC) et les Swarm Cells (SC) forment alors la partie postérieure de la gonade. Un peu avant le stade L3, les cellules somatiques précurseurs des filaments terminaux, (TF, Fig. 30) sont les premières à se différencier grâce à l'expression de facteurs de transcription comme Bric-à-brac ou engrailed.



**Figure 30 - Représentation schématique du développement de la lignée germinale femelle de *Drosophila melanogaster*.** Antérieur vers le haut. De gauche à droite : stades L1 (24h), L2 (48h), Jeune L3 (72h), Vieille L3 (96h), Puce (120h) et partie antérieure du germarium adulte (Gilboa 2015).

Ils formeront la niche somatique contenant les GSC dans le germarium adulte. *pumillio* (*pum*) est exprimé dans les cellules des filaments terminaux. Il est impliqué dans la division asymétrique des PGC (dès la fin du stade L3). Cette division asymétrique est nécessaire à la future organisation de l’ovaire (Lin et Spradling 1997). Au stade L3, les cellules germinales primordiales s’alignent au milieu de la gonade pour l’organisation des ovarioles futures. À la fin du stade larvaire (L3-Puce), les TF vont s’empiler et recruter des (Intermingled Cells (IC) *via* la voie Notch : celles-ci deviendront les « Cap Cells » (Fig. 30). Avant ce stade, *dpp* est produit par les cellules somatiques et maintient les PGC indifférenciées tout en favorisant leur prolifération (Gilboa et Lehmann 2004). Les PGC qui se retrouvent éloignées de la niche commencent à exprimer Bam et leur différenciation débute 112h After Egg Laying (AEL) : les premiers cystoblastes sont observés au stade pré-pupal (Ohlstein et McKearin 1997). À la fin de la métamorphose, des cellules somatiques (Sheet Cells) migrent entre les empilements de TF pour séparer la gonade en 16 à 20 ovarioles distincts.



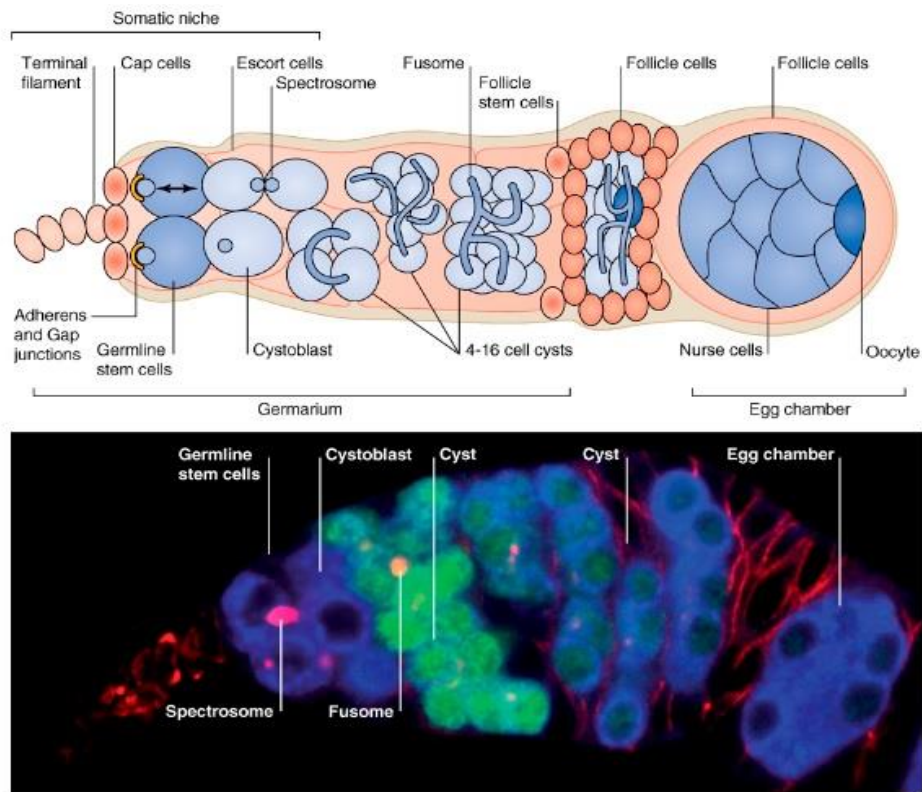
Au stade L3, *pumillio* et *piwi* sont impliqués dans le maintien de l'état indifférencié des cellules germinales et dans leur organisation. Sans *pum*, les spectrosomes sont réduits, de taille variable, voire même absents. Les PGC se divisent de façon symétrique, forment des clusters entourés de cellules somatiques, elles n'ont pas de canaux cytoplasmiques (structure reliant les cellules sœurs après une division) et restent diploïdes. Dans le mutant *piwi*, aucune GSC n'est présente dans les ovarioles ni dans les testicules : le développement de l'ovaire semble se passer normalement jusqu'en L3 mais les PGC ne s'alignent pas correctement dans la gonade. A la place, elles se différencient et peuvent donner quelques chambres à œuf défectueuses avec un nombre de cellules nourricières aberrant et des défauts de polarité (Lin et Spradling 1997; Cox et al. 1998). En revanche les cellules somatiques de l'ovariole sont toujours présentes.

### **A.3 Stade adulte : l'ovogenèse**

Au stade adulte, les deux ovaires contiennent entre seize et vingt ovarioles, les « chaînes de montage » de l'ovocyte. Il s'agit d'une succession de follicules ovariens (ou chambres à œufs) à différents stades de développement organisés selon l'axe antéro-postérieur : la position des TF et des PGC dans la gonade larvaire est conservée pendant la métamorphose. Le germarium, qui contient la niche abritant deux ou trois GSC et les cystes de cellules germinales (Fig. 31), est du côté antérieur de l'ovariole et le follicule le plus mature est placé du côté postérieur.

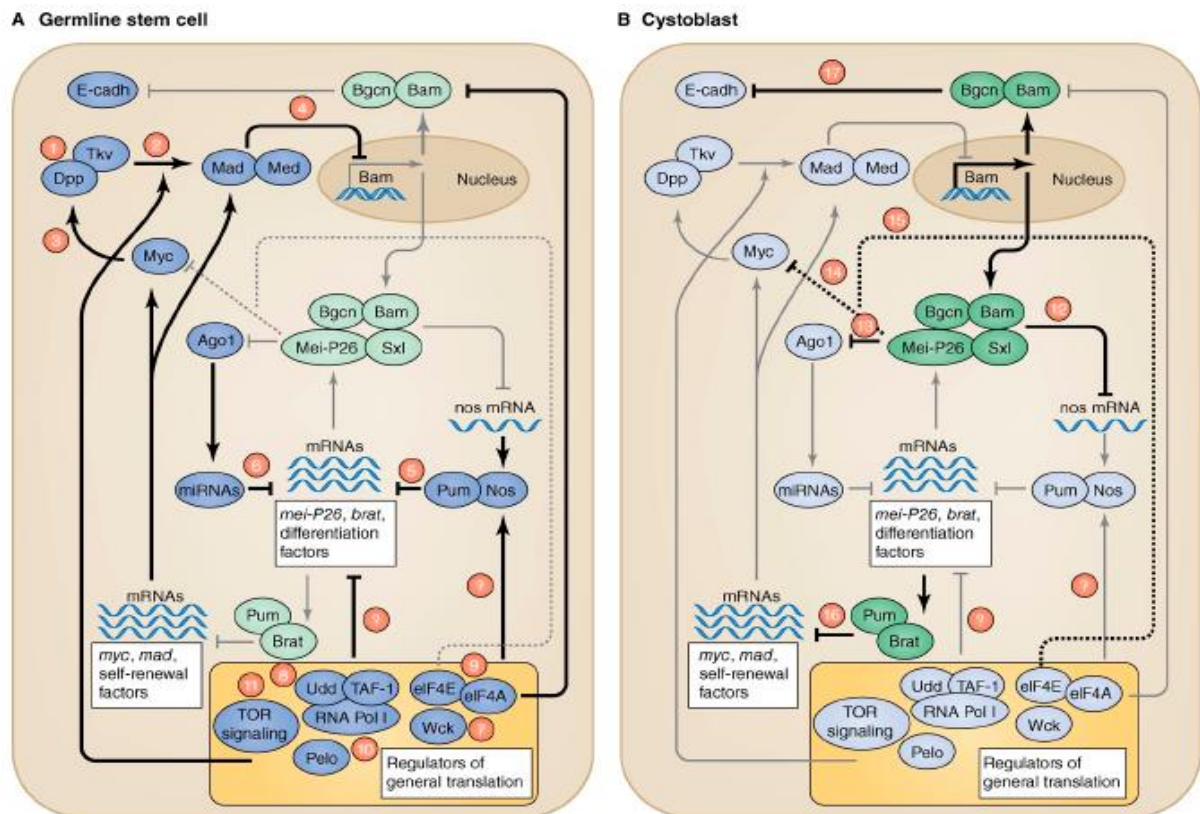
#### **A.3.1 Le germarium**

Dans la niche, les deux ou trois GSC (Fig. 31) se divisent de façon asymétrique : la cellule produite située du côté antérieur du spectrosome reste dans la niche et la deuxième cellule fille (coté postérieur) va initier un programme de différenciation à cause de la perte de contact avec les Cap Cells : cette cellule deviendra le cystoblaste.



**Figure 31 – Germarium.** En haut, schéma d'un germarium d'ovaire adulte de drosophile et de la première chambre à œuf. Les tissus somatiques sont indiqués en rose et les tissus germinaux en bleu. En bas, immunomarquage d'un germarium. Vasa en bleu (cellules germinales). En vert : territoire d'expression de Bam qui dirige la différenciation des cystoblastes et l'entrée en méiose (Bam,GFP). En rouge, spectrosomes dans les GSC et fusome (1B1) dans les cystes ainsi que les membranes des cellules somatiques. Antérieur à gauche (Slaidina et Lehmann 2014).

En effet, le comportement des GSC est régulé par les cellules somatiques qui les entourent : un germarium transplanté dans un mâle continuera à produire des cystoblastes et tous les stades folliculaires (Lin et Spradling 1993). Slaidina et Lehmann ont résumé cette boucle de rétrocontrôle de la différenciation en 2014 (Fig. 32). Dans les GSC, Myc induit la production de Dpp qui active avec Tkv l'expression de *Mad*. Celui-ci réprime alors *Bam* dans les PGC. Ceci empêche l'activation de *Mei-P26* et *Sxl* dans ce tissu pour éviter la différenciation. *Myc* et *Mad* sont donc impliqués dans le renouvellement des GSC, appuyés par la voie TOR et différents répresseurs de transcription, gardant les GSC indifférenciées. Si *Mei-P26* et *Sxl* sont activés, ils répriment la traduction de *nanos*. Lorsque Pum est lié à Brat elle permet la répression de la traduction de *Myc* et *Mad* et donc *Dpp* et *Tkv*.

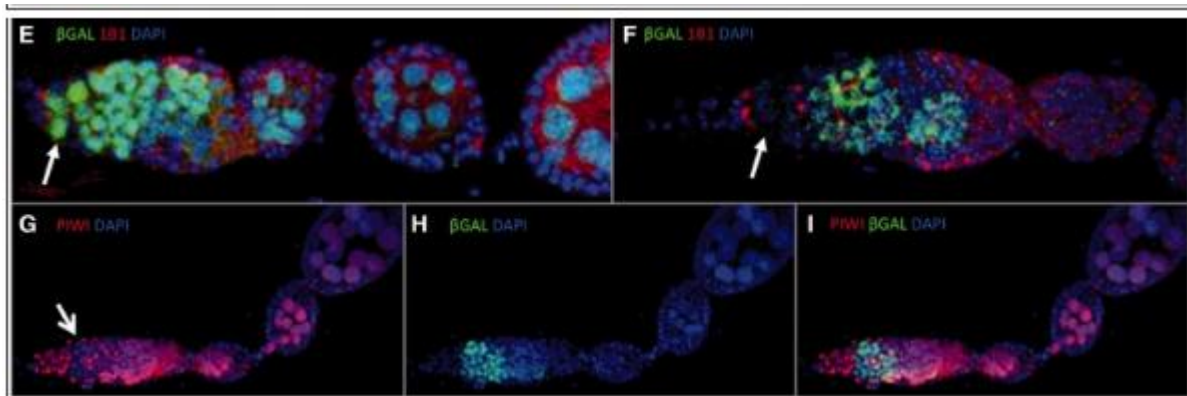


**Figure 32 - Balance entre les facteurs impliqués dans la maintenance des GSC et dans la différenciation en cystoblaste.** (A) Réseau des gènes impliqués dans le maintien de l'état indifférencié des GSC. (B) Schématisation de la différenciation des GSC en cystoblastes. Les facteurs impliqués dans le renouvellement des GSC sont en bleu et les facteurs induisant la différenciation sont en vert. Les facteurs actifs dans le tissu concerné sont en bleu/vert foncé et les facteurs inactifs en bleu/vert clair. Les flèches indiquent une régulation positive ou négative. Abréviations : E-cadh, E-cadherin; Pelo, Pelota; Tkv, Thickvein; Udd, Underdeveloped. (Slaidina et Lehmann 2014).

La transition vers la méiose se produit lorsque les cellules quittent la niche somatique et s'éloignent des signaux qui maintiennent cette boucle. Au moment où le cystoblaste quitte la niche, les niveaux de protéine Bam augmentent entraînant la répression traductionnelle de *Nanos* (Fig. 32). Pum se lie à Brat et bloque la production de Mad et Myc, des facteurs de renouvellement des GSC. Le cystoblaste n'étant plus en contact avec les signaux de la niche, il subit quatre séries de divisions incomplètes pour former un cyste germinale (Fig. 31) de 16 cellules interconnectées par le fusome, un organe dérivant du spectroso.

Dans les cystes germinaux il y a également une modification de la quantité des protéines impliquées dans la répression des ET. L'étude de la répression du rétrotransposon *Idefix* dans la lignée germinale adulte a permis de montrer que Aub, Vasa et SPN-E sont nécessaires dans les toutes premières étapes au niveau du germarium. Elles semblent être moins importantes pour l'efficacité de la répression des ET dans les stades folliculaires tardifs.

En revanche Piwi, Ago3 et Mael sont nécessaires à toutes les étapes de l'ovogenèse (Dufourt et al. 2014). Étonnamment, il a été remarqué que l'expression de transgènes *Idefix*-GFP contenant le rapporteur GFP couplé à un fragment du gène *gag* d'*Idefix* cloné en antisens, échappe au contrôle de l'hôte entre les stades cystoblaste et cyste à 16 cellules. On observe une expression forte de la GFP uniquement dans les cystes du germarium. Cette atténuation de la répression est corrélée avec une diminution localisée de l'expression de la protéine Piwi (Fig. 33 G- I).

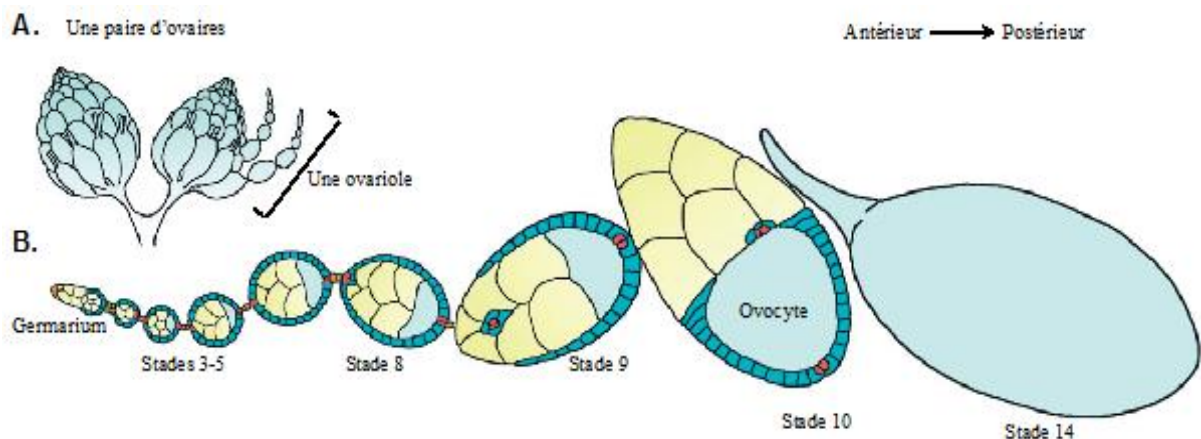


**Figure 33 - Piwi less Pocket** (E) Double immunomarquage d'un germarium de femelle *BC69* et (F) germarium de descendant du croisement d'une femelle *RS3* par un mâle *BC69*. 1B1 (rouge) marque les spectrosomes de E-G (flèches), et la  $\beta$ -galactosidase (vert) est la cible non réprimée. L'ADN est marqué au DAPI (bleu). Lorsque le transgène *RS3* est hérité maternellement (F), l'expression de la  $\beta$ -galactosidase est partiellement réprimée dans la zone des cystes germinaux. (G-I) Double immunomarquage d'un germarium pour Piwi (rouge) et  $\beta$ -galactosidase (vert). DAPI (bleu). La répression incomplète colocalise avec une faible présence de Piwi dans le germarium (marqué par une flèche G) (Dufourt et al. 2014).

Les transgènes *P-lacZ* sont également fortement dérégulés dans ces cystes, même en présence d'un répresseur fort comme *RS3* (Patterning, Fig. 12 E-F). Cette région particulière a été appelée «Piwi-less Pocket» ou PILP, et permettrait un léger relâchement de la répression des ET. De cette façon, de nouvelles insertions peuvent se produire et seraient transmises à la génération suivante, contribuant ainsi à la dynamique du génome (Dufourt et al. 2014). Il s'agit en effet du tissu où ce type de « loterie génétique » a le moins d'impact sur l'organisme. Les cellules souches répriment fortement les ET et si il y a mobilisation dans le cystoblaste ou dans une cellule du cyste entraînant une mutation trop délétère, l'œuf résultant ne donnera pas d'individu viable. En revanche les GSC répriment toujours les ET donc « la chaîne de production » continuera à fournir des œufs.

### A.3.2 Les chambres à œufs, follicules ovariens

Les cystes sortent ensuite du germarium et s'entourent progressivement de cellules folliculaires (Fig. 31, stade 3). L'une des cellules du cyste est alors spécifiée comme l'ovocyte. Les 15 autres cellules de la chambre, appelées cellules nourricières, serviront de support à l'ovocyte. Elles produisent les protéines et les transcrits nécessaires au développement précoce de l'œuf et les déversent dans l'ovocyte grâce aux ponts cytoplasmiques qui les connectent (Fig. 31). Le cytoplasme germinal qui est mis en place au pôle postérieur de l'ovocyte permet la création de nouvelles PGC et donc la pérennité du cycle.



## **B. Maintien de la répression des ET au cours de la vie des cellules germinales**

### **B.1 La protection du génome embryonnaire, une responsabilité maternelle.**

Les mécanismes de protection de la lignée germinale embryonnaire sont très divers à travers l'arbre du vivant, mais toujours présents. La suppression de l'activité des transposons dans les gamètes est la première défense pour assurer la stabilité du génome en empêchant l'activation de la transposition dans la génération suivante.

#### ***B.1.1 Plantes***

Chez *A. thaliana*, des siARN sont synthétisés par l'ARN pol IV à partir de *loci* incluant des éléments transposables et des séquences répétées suite à la dérégulation sélective des ET dans l'endosperme (Fig. 4) (Gebert et Rosenkranz 2015). Ces siARN s'accumulent dans l'endosperme et guident la méthylation de séquences ADN homologues dans la cellule œuf de façon analogue aux sncARN de la paramécie *Tetrahymena* (Mochizuki et Gorovsky 2004; Schoeberl et al. 2012). Bien que ce mécanisme fonctionne grâce à la suppression de la répression des transposons dans l'embryon, cette dérégulation n'est pas nuisible puisque l'endosperme n'est pas hérité par la génération suivante (Ito 2012). Cette hétérochromatinisation des copies euchromatiques d'ET permet donc la répression des séquences dangereuses dans les cellules germinales.

#### ***B.1.2 Insectes***

Chez *D. melanogaster*, il n'y a pas besoin de déréguler les transposons pour activer la production de piARN spécifiques. Ceux-ci sont produits tout au long de l'ovogenèse par les cellules nourricières et déversés dans l'œuf par des canaux cytoplasmiques une fois chargés par Piwi ou Aub. De nombreux exemples de l'efficacité de ces piARN maternels ont été découverts, en particulier pour la suppression de la dysgénésie des hybrides liée aux éléments *P* et *I*.

Dans un article de 1993, S. Ronsseray a observé que les descendants de femelles hétérozygotes portant un transgène *P* au télomère sont capables de réprimer la dysgénésie des hybrides, qu'ils aient reçu le transgène régulateur ou non (Ronsseray et al. 1993). Ceci montre que l'héritage cytoplasmique de piARN est efficace dans la génération suivante alors que la



seconde génération des mouches sans transgène régulateur présente de la dysgénésie. Ces résultats ont été répétés et interprétés plus finement en 1996 par M. Simmons sans pour autant les lier aux piARN encore inconnus à cette époque (Simmons et al. 1996). Ils sont généralisables à d'autres éléments transposables puisque l'héritage maternel d'ARN produits par une mère hétérozygote pour un transgène contenant un fragment de l'élément *I* montre une descendance protégée de l'activité *I* pendant une génération (environ 53% de répression de l'activité *I*) (Jensen et al. 1999).

### *B.1.3 Mammifères*

Le rôle de la voie des piARN pendant l'ovogenèse des mammifères est resté mystérieux assez longtemps. En effet, le modèle *M. musculus* ne montrait pas le même schéma que la *Drosophile*: dans les premières études on trouvait peu de piARN dans les ovocytes de souris et ils ne semblent pas nécessaires pour leur développement alors que beaucoup d'études chez le mâle ont révélé un impact important lors de la spermatogenèse. Néanmoins, l'analyse récente des petits ARN de l'homme, la souris, le macaque et la vache a révélé que des piARN sont également exprimés dans la lignée germinale femelle de ces espèces (Rosenkranz et al. 2015). L'observation de l'expression dynamique de PIWIL1, PIWIL2 et PIWIL3 pendant l'ovogenèse des mammifères et l'embryogenèse précoce montre des différences d'une espèce à l'autre. Les ovocytes dans l'ovaire humain adulte expriment PIWIL1 et PIWIL2, alors que ceux de l'ovaire bovin expriment seulement PIWIL1. Dans les ovaires (humains, bovin, macaque) ils ont observé des piARN qui ressemblent à ceux des testicules. Les ovocytes folliculaires bovins contiennent d'abondants piARN relativement courts qui ciblent préférentiellement les éléments transposables accompagnés de la protéine PIWIL3 ainsi que d'autres composants de la voie des piARN. Une partie de ceux-ci est transmise aux embryons précoces, révélant un probable impact sur l'embryogenèse chez les mammifères (Roovers et al. 2015). Les ARNm et la protéine Miwi ont été observés dans les ovaires 1 et 2 semaines postpartum, respectivement. Les ARNm de *miwi* bien que présents dans les follicules primaires et secondaires diminuent lors de la maturation (X. Ding et al. 2013). Pour autant, l'impact fonctionnel de ces piARN de mammifères dans la lignée germinale femelle est encore en cours d'étude.

## **B.2 Activation des clusters de piARN**

En plus de la protection du génome au début du développement, l'effet maternel observé lors de l'étude du TSE chez la *Drosophile* montre qu'un héritage maternel de piARN

est indispensable pour la répression d'une cible homologue dans la descendance. Cela signifie que les clusters de piARN sont activés par les piARN déposés dans l'embryon. Le moment précis de l'activation desdits clusters au cours du développement est une question encore sans réponse. L'objet de mes trois ans de thèse a donc été de tenter d'y répondre en mettant en évidence la répression des éléments transposables par les piARN dans les gonades en développement et en caractérisant les partenaires protéiques de cette régulation.





## **II – Résultats**



## **Visualiser la répression d'un transgène rapporteur dans les cellules germinales au cours du développement**

Au laboratoire, nous étudions la répression médiée par les piARN grâce à des transgènes. Des transgènes dits « Silencer » sont insérés dans les *locus* producteurs de piARN où ils acquièrent les propriétés de ces *locus* en quelques générations (A. Laidin, en préparation). Nous utilisons également des transgènes dits « Cibles » qui permettent l'expression d'une protéine rapporteur dans la lignée germinale, généralement la  $\beta$ galactosidase ou une protéine fluorescente telle que la GFP ou la RFP. Nous privilégions les transgènes cibles dont l'expression est restreinte à la lignée germinale grâce à un phénomène d'enhancer-trap ou de transgènes inductibles. Les transgènes cibles possédant des séquences UAS couplés à un transgène pilote qui exprime une protéine activatrice de la transcription, GAL4, de façon tissu-spécifique en fonction de leur promoteur, sont un outil précieux. Pour la lignée germinale les transgènes pilotes possédant le promoteur des gènes *nanos* ou *vasa* sont très efficaces.

Notre premier objectif a donc été de trouver la combinaison de transgènes permettant l'expression d'une cible dans les cellules germinales au cours du développement embryonnaire, larvaire et pupal. En nous basant sur les analyses de protéomes et de transcriptomes de PGC embryonnaires de Najeeb U. Siddiqui et collègues, nous avons cherché des gènes exprimés tôt au cours du développement et spécifiques de la lignée germinale (Siddiqui et al. 2012). La protéine Vasa est déposée dans l'œuf par la mère au niveau du pôle postérieur et le gène est exprimé par le zygote dès le début de la transcription et de la traduction zygotique (3 heures après la ponte). Dans cet article, des cellules germinales d'embryons exprimant la protéine de fusion VasaGFP ont été récupérées par FACs et après vérification de la pureté et la spécificité de la population de PGC, le profil d'expression d'un grand nombre de gènes a pu être obtenu grâce à des expériences de RNAseq. Les transcriptomes et protéomes d'embryons de 0-3h, 3h-5h et 5h-7h ont été comparés. Les gènes de la *Drosophila* ont été classés en neuf classes selon l'évolution de leur expression au cours du temps. Pour notre expérience nous avons choisi les gènes dont la quantité de transcrits augmente durant les 7 premières heures du développement appartenant aux catégories VI, VIII et IX (Siddiqui et al. 2012). Ces trois catégories englobent 140 gènes, parmi lesquels, nous avons trouvé 39 souches possédant une insertion de type enhancer-trap (*lacZ* ou *gal4*) disponibles au Bloomington Stock Center ou au DGRC de Kyoto. Nous avons

commandé ces lignées et j'ai testé l'expression de ces transgènes par coloration X-GAL, soit directement pour les lignées enhancer-trap *lacZ*, soit par croisement avec une lignée rapporteur *UASplacZ* pour les lignées enhancer-trap *gal4*. J'ai également testé différents pilotes *gal4* spécifiques de la lignée germinale adulte disponibles au Bloomington Stock Center: *vasaGal4*, *matatubulineGal4*, *nosGal4VP16* et les *nosGal4green-eye 32180* et *32179*.






## **1. Résultats du crible basé sur les transcriptomes obtenus par Siddiqui et al. 2012**

### **1.1 Crible des lignées enhancer-trap *lacZ***

Dans un premier temps, les embryons des souches candidates ont été directement colorés. Certaines souches expriment les transgènes durant l'ovogénèse, donnant ainsi de jeunes embryons (0-3h) intégralement bleus. Afin de savoir s'il s'agit d'un dépôt maternel ou d'une expression zygotique précoce, un croisement de mâles « enhancer-trap *lacZ* » par des femelles Canton (contrôle) a été réalisé. Le transgène hérité paternellement conduit à une disparition de la coloration. Il s'agit donc d'un dépôt maternel de protéine: pas d'expression zygotique précoce du transgène.

Si durant les trois premières heures du développement on ne détecte pas d'expression de la  $\beta$ galactosidase, ni somatique ni germinale, après ce laps de temps les embryons des différentes lignées présentent un grand nombre de profils d'expression somatique différents (Bilan et exemples Tableau 1). Parmi les 19 lignées testées qui possèdent un transgène enhancer-trap *lacZ*, aucune n'a montré d'expression spécifique de la cible dans les PGC. Cependant, 4 lignées enhancer-trap *lacZ* (*12179*, *11877*, *12259* et *12253*) ont montrées une expression somatique importante pouvant masquer l'expression de la  $\beta$ galactosidase dans les PGC (Tableau 2 « ubiquitaire» Annexes).

**Tableau 1 - Résultat du crible**

Lignées (44)	« Ubiquitaire »	Groupes de cellules isolées	Paterné	Tête, intestin, glandes salivaires	Pas ou peu d'expression
<i>gal4</i> (22)	3	3	3	13	0
<i>lacZ</i> (19)	4	4	4	4	3
Exemples					

### 1.2 Lignées enhancer-trap *gal4*

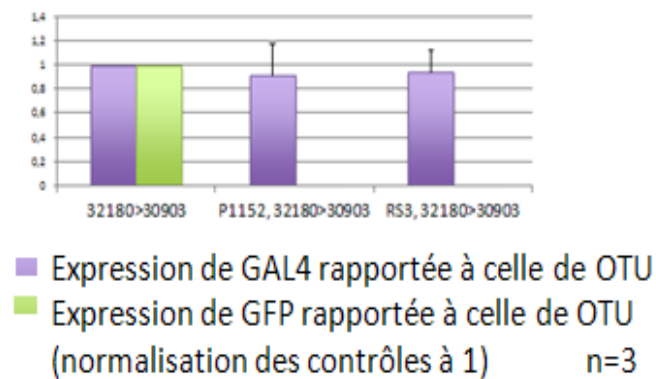
Pour révéler l'activité de la protéine Gal4, un transgène rapporteur *UASpCible* (GFP ou  $\beta$ galactosidase) est nécessaire. J'ai donc croisé les mâles des lignées possédant un transgène enhancer-trap *gal4* avec des femelles *UASplacZ* et testé les 22 souches possédant une insertion *gal4* listées dans le Tableau 2 (Annexes). Aucune n'a montré d'expression spécifique du rapporteur dans les cellules germinales embryonnaires mais comme pour les lignées enhancer-trap *lacZ*, certaines (103569, 104811 et 112347) ont montré des expressions somatiques importantes pouvant masquer une expression dans les PGC (Tableau 1 et 2 « ubiquitaire »).

Une coloration X-gal sur les embryons issus de croisements femelles *UASplacZ* par les mâles de ces lignées ne permet pas d'observer l'expression de la cible au niveau cellulaire. Nous avons donc recherché par immunodétection un comarquage éventuel dans les PGC avec un anticorps anti-Vasa et un anticorps anti-GFP ou anti- $\beta$ galactosidase au microscope confocal. J'ai observé que la GFP et la  $\beta$ galactosidase s'expriment dans les embryons de ces lignées au moins au stade 14 mais uniquement dans certaines cellules somatiques. L'augmentation de la quantité des ARN mesurée par RNAseq dans les cellules germinales d'embryons jeunes (Siddiqui et al. 2012) suggère fortement que la transcription débute. En revanche, aucune des lignées testées n'a montré d'expression zygotique de  $\beta$ galactosidase ou de GFP dans les PGC. La conclusion est que la traduction des différents rapporteurs semble bloquée au moins jusqu'au stade larvaire 1 (L1) pour tous les transgènes testés. Ceci est probablement dû à l'existence de 3'UTR spécifiques nécessaires à la traduction pour ces stades (Rangan et al. 2009). En effet, il existe une régulation fine de la traduction dans les

PGC embryonnaires, les protéines et peptides issues de plusieurs gènes dont *nanos*, *pumillio* et *pgc* répriment activement la traduction pour maintenir ces cellules indifférenciées (voir Intro Chap. 4). Dans les 44 lignées testées, aucun transgène ne contient de 3'UTR actif dans la lignée germinale précoce.

## 2. Test des constructions pilotes *gal4* spécifiques de la lignée germinale adulte

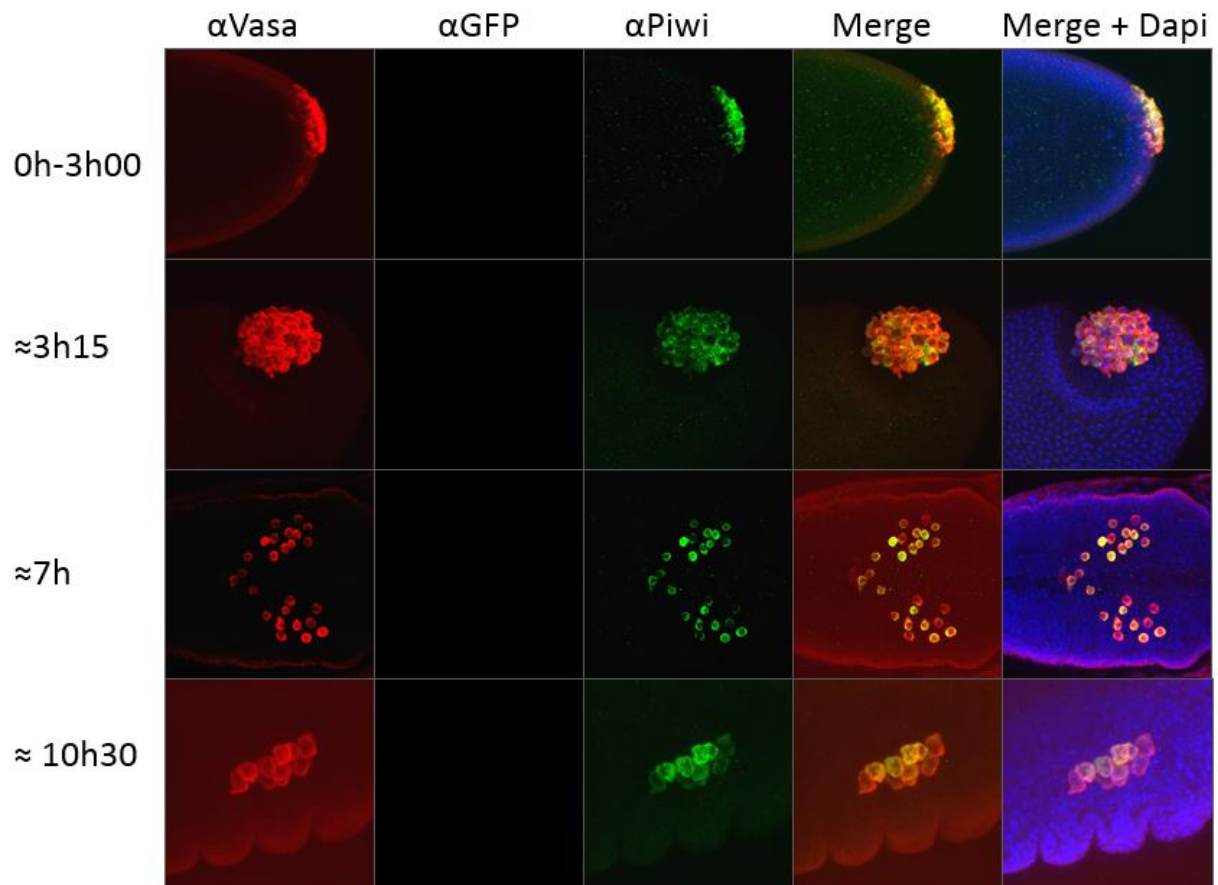
Nous avons testé ces lignées en croisant des mâles pilotes *gal4* par des femelles *UASpGFP<sup>30903</sup>*. Nous avons d'abord vérifié en ovaire adulte que dans le système *Gal4/UASpGFP<sup>30903</sup>/Silencer*, seule l'expression la cible *UASpGFP<sup>30903</sup>* est éteinte par le Silencer et pas celle du transgène pilote *nosGal4*. À titre d'exemple, les résultats de la souche *nosGal4<sup>32180</sup>* sont présentés (Fig. 35).



**Figure 35 - Résultats de RTqPCR** effectuée sur les ARNm de Gal4 et GFP en présence des Silencers *P-1152* et *RS3* sur des ovaires adultes. En ordonné : normalisation des contrôles à 1.

Une transcription inverse a été faite sur l'ARN total extrait d'ovaires adultes. L'analyse en qPCR des séquences OTU (référence), *gal4* et *gfp* montre que les transcrits *gfp* (en vert) disparaissent en présence des Silencers *RS3* et *P-1152*. En revanche, les transcrits *gal4* (en violet) ne sont pas affectés, le pilote *nosGal4<sup>32180</sup>* a donc été validé pour notre expérience. La souche *RS3* contient un transgène *P-FRT-white* inséré au site 100E3 (Bloomington #123282), elle produit des piARN de *P*, *w+* et de *FRT* (Dufourt et al. 2014) et la souche *P-1152* possède deux transgènes *PlacZ-rosy* produisant des piARN de *P*, *lacZ* et *rosy* (de Vanssay et al. 2012). Elles possèdent donc toutes les deux des séquences de l'élément *P*, environ 500pb, qui sont homologues avec une partie du transcrit du transgène cible *UASpGFP<sup>30903</sup>*.

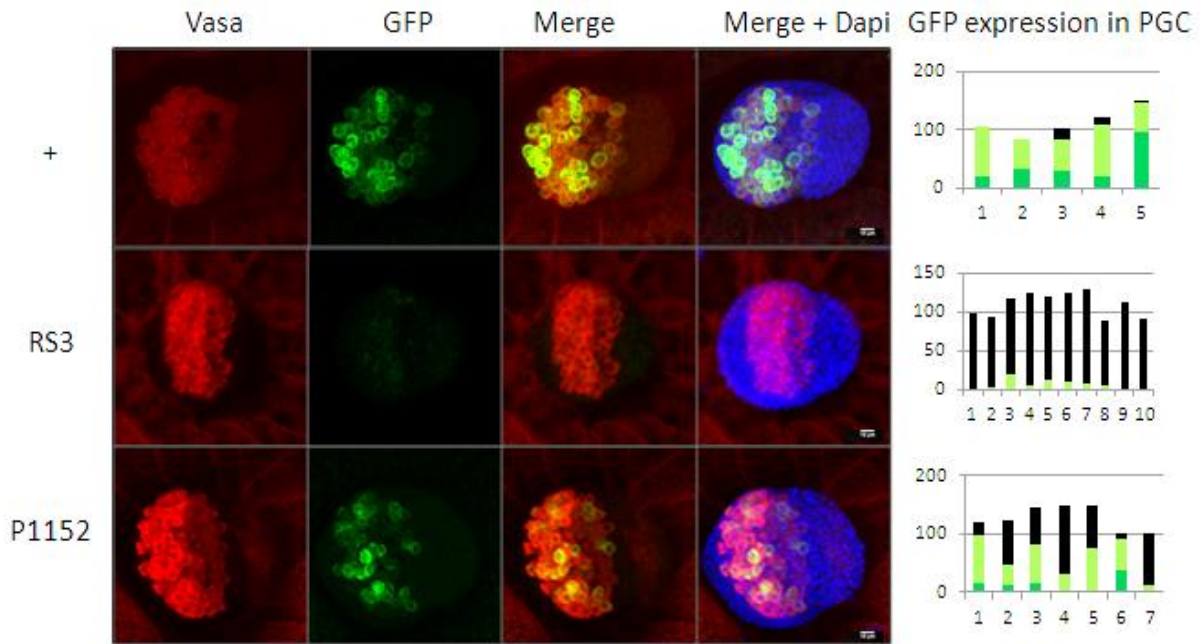
Les croisements de femelles *UASpGFP<sup>30903</sup>* par des mâles *nosGal4<sup>32180</sup>* ont été réalisés en pondoirs et les œufs ont été récupérés à différents stades représentatifs de l'embryogénèse. Ils ont ensuite été immunomarqués ( $\alpha$ Vasa,  $\alpha$ Piwi,  $\alpha$ GFP Fig. 36).



**Figure 36 - Triple immunomarcage d'embryons issus du croisement de femelles  $UASpGFP^{30903}$  par des mâles  $nosGal4^{32180}$ .** On peut voir que la protéine Piwi en vert (un des partenaires protéiques essentiels de la répression par les piARN) colocalise avec Vasa, en rouge. La GFP est également en vert et le Dapi (ADN) en bleu. Il n'y a pas d'expression de la GFP dans les cellules germinales embryonnaires. La protéine Piwi est fortement présente dans les cellules germinales primordiales durant la vie embryonnaire (au moins à partir de ~1 à 2h AEL).

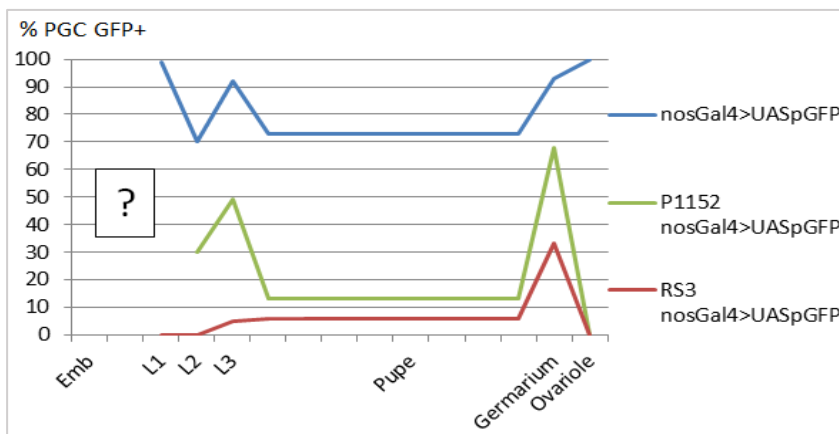
Les souches possédant les transgènes télomériques silencer *P-1152* ou *RS3* et la cible  $UASpGFP^{30903}$  ont été construites, mais on ne détecte aucune expression de la cible GFP dans les PGC à ces stades. En absence d'expression d'une cible, observer la répression durant l'embryogénèse n'apparaît pas possible. Néanmoins la cible  $UASpGFP^{30903}$  s'exprime à partir du stade L1. Des immunomarquages de gonades femelles de larves à partir du stade L1 jusqu'au stade pupal ont été réalisées et la répression a été quantifiée pour chaque stade avec les deux transgènes silenciers (exemple du stade L3 Fig. 37).





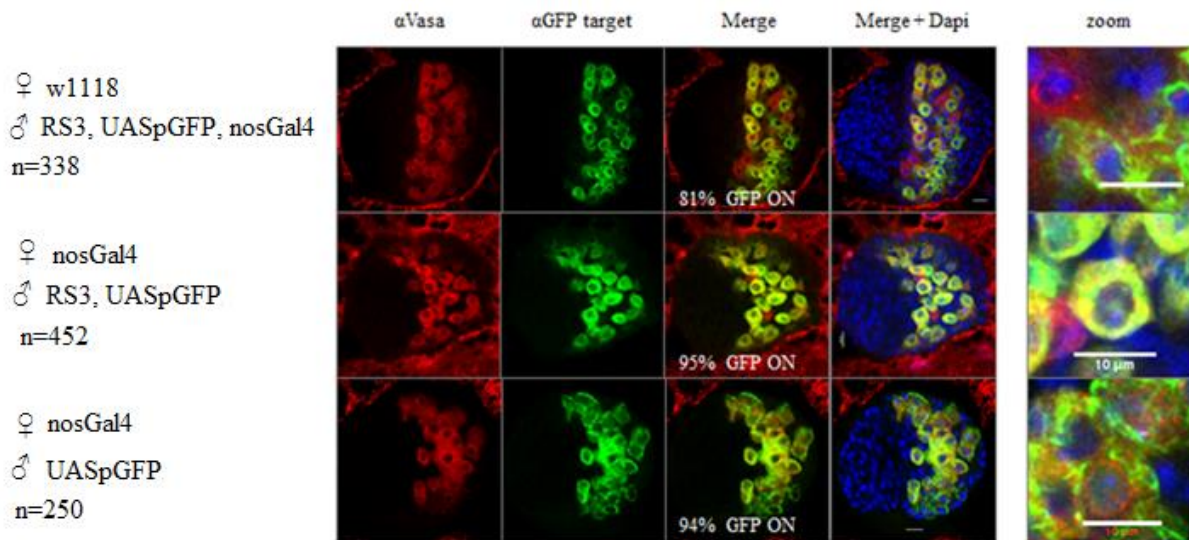
**Figure 37 - Double immunomarcage de gonades de larves L3 femelles** issues des croisements suivants  $\pm$  : ♀ *UASpGFP<sup>30903</sup>* par ♂ *nosGal4<sup>32180</sup>*, *RS3* : ♀ *RS3,UASpGFP<sup>30903</sup>* par ♂ *nosGal4<sup>32180</sup>* et *P-1152* : ♀ *P-1152;UASpGFP<sup>30903</sup>* par ♂ *nosGal4<sup>32180</sup>*. Histogrammes : En abscisse, chaque barre représente une gonade. En ordonnée, nombre de PGC comptées. PGC [GFP-] en noir, PGC [GFP+] en vert clair, PGC [GFP++] en vert foncé.

La GFP ne s'exprime pas dans totalité des cellules germinales du contrôle, au stade L3 94% de cellules germinales sont [GFP+] (Fig. 37) mais pour les autres stades le pourcentage est moins élevé. En L3, elle est presque entièrement éteinte par *RS3* (93% cellules éteintes,  $X^2$   $p < 0.001$ ). La descendance du croisement avec *P-1152* présente plus d'hétérogénéité: seulement 50% des cellules germinales sont éteintes. Ce résultat reste statistiquement très différent du contrôle,  $X^2$   $p < 0.001$ . Nous confirmons donc de manière quantitative les résultats précédemment décrits dans (Josse et al. 2008) et les étendons à une grande partie du développement larvaire. La répression est bien fonctionnelle en L3 et selon la force du silencer, de la variabilité peut être observée. Le bilan du développement est montré dans la Figure 38.



**Figure 38 - Répression du transgène *UASpGFP<sup>30903</sup>* au cours du développement de la lignée germinale femelle.**

Nous avons cherché à améliorer le niveau d'expression du contrôle afin de mieux caractériser la répression partielle exercée par le silencer *P-1152*. Le transgène *UASpGFP<sup>7373</sup>* montre une meilleure expression de la GFP dans les cellules germinales, en particulier quand le pilote *nosGal4* est aussi apporté maternellement (Fig. 39).



**Figure 39 - Pourcentage d'expression du transgène *UASpGFP<sup>7373</sup>* en fonction de l'héritage maternel ou paternel du pilote *nosGal4<sup>32180</sup>* et en présence ou non du silencer *RS3*. On remarque que la GFP n'est pas exprimée dans toutes les cellules germinales larvaires lorsque le pilote *nosGal4<sup>32180</sup>* est apporté par le père (première ligne 81% d'expression). Lorsqu'il est apporté par la mère, on obtient 95% des PGC exprimant la GFP (deuxième ligne). La présence du silencer *RS3* amené par le père n'a pas d'incidence sur l'expression de la GFP: le même pourcentage d'expression est quantifié entre les deuxième et troisième lignes avec et sans le silencer *RS3*.**

Cet effet est probablement dû au dépôt maternel de protéine Gal4 dans l'embryon, au niveau du germlasme (Williamson et Lehmann 1996; Van Doren et al. 1998). De plus, il est exprimé dès la fin de l'embryogenèse (~stade 14). C'est ce transgène cible qui a été utilisé pour l'article central de ma thèse.



## **Article 1 : « From egg to adult: piRNA-mediated silencing throughout germline development in *Drosophila* »**

### **Présentation de l'article**

Le premier objectif de cette thèse était de caractériser la répression par les piARN au cours du développement de la lignée germinale femelle et d'étudier son mécanisme moléculaire. Nous voulions voir s'il était différent de celui de l'adulte. En effet, on sait que les transposons sont régulés à plusieurs niveaux dans la lignée germinale femelle adulte : les *loci* sont réprimés transcriptionnellement par la mise en place de marques hétérochromatiques répressives et les transcrits qui échapperaient à cette répression sont ciblés dans le cytoplasme par des petits ARN complémentaires guidant une protéine à domaine RNase H capable de les cliver. C'est dans les ovaires adultes que la plupart des informations connues sur les piARN ont été démontrés. En raison de la forte activité transcriptionnelle des cellules nourricières et du fait de l'héritage de piARN maternels au début de son développement, cette production massive de petits ARN et de protéines impliquées dans la voie des piARN semble essentielle pour réprimer les transposons. Néanmoins, on peut penser que durant la mise en place des cellules germinales, au début de l'embryogenèse, celles-ci n'ont pas besoin de produire zygotiquement des piARN, d'autant qu'elles sont transcriptionnellement inactives à ce stade. Par exemple, le dépôt maternel pourrait suffire à assurer temporairement la répression des ET. Notre but était donc de visualiser la mise en place de la répression et de voir si elle était maintenue tout au long de la vie des cellules germinales au même niveau d'efficacité. La découverte de fenêtres de relâchements de la régulation autorisant une certaine mobilisation des éléments transposables, comme la poche Piwi-less (PILP) du germarium adulte aurait été intéressante (Dufourt et al. 2014).

La répression partielle d'un transgène *P-lacZ* euchromatique, *BC69*, dans les gonades de larve de troisième stade avaient été obtenus et publiés (Josse et al. 2008). Mon projet de thèse a été d'étendre ce travail à tout le développement de la lignée germinale femelle en mettant d'abord en place les outils nécessaires à la visualisation de la répression fonctionnelle de transgènes rapporteurs fluorescents. Dans cet article, nous montrons que la répression d'un transgène peut être observée depuis le stade embryonnaire jusqu'à l'adulte. Cette répression peut être totale ou partielle suivant le silencer utilisé et les cellules qui ne répriment pas la cible semblent prendre cette décision au moment de l'embryogenèse et la maintenir durant le

développement. J'ai ensuite analysé le mécanisme moléculaire de la répression grâce à un criblé gène candidat évaluant l'impact de l'absence de certaines protéines de la voie des piARN au cours du développement. Enfin, l'analyse des petits ARN interférents produits au troisième stade larvaire en conditions sauvages et mutantes pour les protéines Rhino et HP1a montre qu'une production zygotique de piARN existe. Ces résultats confirment le rôle de la protéine HP1a dans la biogenèse des piARN télomériques évoqué dans l'article de A. Todeschini et collègues (Todeschini et al. 2010).

Cette étude a été soumise pour une publication dans la revue G3.

## **From egg to adult: piRNA-mediated silencing throughout germline development in *Drosophila***

Pauline MARIE<sup>1</sup>, Stéphane RONSSERAY<sup>1\*</sup> and Antoine BOIVIN<sup>1\*</sup>

<sup>1</sup> Sorbonne Universités, Université Pierre et Marie Curie (UPMC), CNRS, Institut de Biologie Paris-Seine (IBPS), Laboratoire de Biologie du Développement, F-75005, Paris, France

\* co-corresponding authors

### **Running title**

piRNA-mediated silencing during development

### **Keywords**

piRNA / development / germline / silencing / *Drosophila*

### **Corresponding authors:**

Antoine BOIVIN and Stéphane Ronsseray

Laboratoire Biologie du Développement UMR7622

Institut de Biologie Paris-Seine

Bât. C, 7<sup>ème</sup> étage, porte 703

Boîte courrier 24

9, Quai St Bernard

75252 Paris Cedex 05 FRANCE

antoine.boivin@upmc.fr or stephane.ronsseray@upmc.fr

Tel: (33) 1 44 27 22 88 / Fax: (33) 1 44 27 34 45

## **Abstract**

In metazoan germ cells, transposable element activity is repressed by small noncoding PIWI-associated RNAs (piRNAs). Numerous investigations in *Drosophila* have enlightened the mechanism of this repression in the adult germline. However, very little is known about piRNA-mediated repression during germline development. Nevertheless, to maintain the integrity of the genome, repression should occur throughout the lifespan of germ cells. Here we show that piRNA-mediated repression is active in the female germline, from late embryonic to pupal primordial germ cells, and that genes related to the adult piRNA pathway are required for repression during development. *rhino*-dependent piRNAs, exhibiting the molecular signature of the piRNA pathway "ping-pong" amplification step, are detected in larval gonads, arguing for *de novo* biogenesis of functional piRNAs during development. We also show that production of telomeric piRNAs depends on Heterochromatin Protein 1a (HP1a). Furthermore, as in adult ovaries, we observe an incomplete, bimodal and stochastic repression resembling variegation at all developmental stages. Clonal analyzes of this incomplete silencing strongly suggest that a cellular memory of an early repression decision is initiated in embryonic germ cells and further maintained until the adult stage. Taken together, the data presented here show that piRNAs and their associated proteins are epigenetic components of a continuous repression system throughout germ cell development.

## Introduction

PIWI-interacting RNAs (piRNAs) are a class of small RNAs, ranging from  $\approx 23$  to 29 nucleotides, which ensure repression of transposable element (TE) activity in germ cells of metazoans (ARAVIN *et al.* 2007; IWASAKI *et al.* 2015). In *Drosophila melanogaster*, most piRNA sequences are complementary to a small number of heterochromatic genomic loci located near centromeres or telomeres. Such loci, called piRNA clusters, are heritable repositories of ancient or recent TEs (BRENNECKE *et al.* 2007; SENTI AND BRENNECKE 2010). In the germline, the Rhino-Deadlock-Cutoff (RDC) complex ensures non-canonical dual-strand transcription of most piRNA clusters that can bypass splicing and prevent RNA pol II termination (MOHN *et al.* 2014; ZHANG *et al.* 2014). Consequently, any new sequence (e.g. active TE or artificial transgene) inserted into these RDC-dependent piRNA clusters will be incorporated into long chimeric piRNA precursor transcripts, that are further matured into piRNAs in the nuage, an optically dense cytoplasmic region surrounding nurse cell nuclei (MUERDTER *et al.* 2012; IWASAKI *et al.* 2015). piRNAs are then loaded onto RNA slicing-competent Argonaute proteins (Aubergine (Aub) and Ago-3) and post-transcriptionally neutralize the expression of active TE copies through sequence complementarity. From the sliced TE mRNAs, new complementary sense piRNAs are generated that in turn direct cleavage of antisense precursors from piRNA clusters, a cycle termed "ping-pong amplification" (BRENNECKE *et al.* 2007; GUNAWARDANE *et al.* 2007). This posttranscriptional gene silencing (PTGS) is reinforced by transcriptional gene silencing (TGS), mediated by nuclear Piwi-bound piRNAs and cofactors, that target complementary nascent transcripts, thereby leading to chromatin modifications repressive for transcription (WANG AND ELGIN 2011; SIENSKI *et al.* 2012; LE THOMAS *et al.* 2013; ROZHKOVA *et al.* 2013; PEZIC *et al.* 2014; SIENSKI *et al.* 2015; YU *et al.* 2015).

In *Drosophila*, piRNA-mediated repression has been well studied in ovaries, the adult stage of the germline, or in ovary-derived cells, whereas very little is known about TE repression in developing germinal cells. Primordial Germ Cells (PGCs) are the first cells formed, at the posterior pole of the syncytial embryo. During embryogenesis, PGCs migrate, separate into two groups, and coalesce with somatic gonadal precursor cells to form two gonads that acquired sexual identity without dividing (DANSEREAU AND LASKO 2008). After hatching, female PGCs undergo four non-synchronous rounds of mitosis during larval and pupal development (GILBOA AND LEHMANN 2006). About 10 PGCs per gonad in the late embryo give  $\approx 150$  PGCs per pupal gonad (see cycle in Fig S1). In pupae, PGCs in contact with the niche (somatic cells that structure the adult ovary at the germarium tip) remain undifferentiated and become adult germline stem



cells (GSCs) (DANSEREAU AND LASKO 2008). In adult germariums, each GSC divides asymmetrically to produce a new GSC and a cystoblast that undergoes four rounds of mitosis with incomplete cytokinesis to form a 16-cell egg chamber (HUYNH AND ST JOHNSTON 2004). The *Drosophila* ovary is composed of 16–20 ovarioles, each containing a chain of progressively more mature egg chambers.

Are transposable elements repressed throughout *Drosophila* germline development, and which are the functional and molecular properties of this repression? Using transgene-based assays, previous results from our laboratory indicated that repression occurs in the female third instar larval gonad (JOSSE *et al.* 2008). Here, we demonstrate that piRNA-mediated repression is active throughout female germline development, from the late embryonic PGCs to the pupal PGCs. We show that 14 genes implicated in the adult piRNA pathway are also required for repression during germline development. piRNAs harboring a ping-pong signature are detected in late third instar larval gonads, and analyses of zygotic *rhino* knock down show that maternally inherited piRNAs are not sufficient to establish complete repression capacities. We establish that HP1a is necessary for telomeric piRNA biogenesis. Finally, incomplete repression resembling variegation is observed. Clonal analyses strongly suggest that this incomplete silencing results from a decision taken in the embryonic germ cells and stably maintained throughout development.

## MATERIALS AND METHODS

### *Drosophila* stocks

Flies were raised at 25°C. Stocks come from the *Bloomington Drosophila Stock Center* (#7373, #32180, #32249, #6420 and TRiP lines) and from the *Kyoto Drosophila Genomics and Genetic Resources* (#123282). *BC69* bears a *P-lacZ ry+* (*P{A92}*) enhancer trap transgene inserted in the *vasa* gene. Homozygous females are sterile but homozygous males are fertile.

### Immunostaining

Embryos were collected on petri dishes 18 to 22h AEL. After rinsing in tap water, embryos were dechorionated in 50% bleach for 4 minutes and rinsed again in water. Eggs were fixed in 3.7% formaldehyde in PBS containing 0.8% Triton X-100 (PBT) for 5 minutes sonicated for 2x7 sec at maximum intensity with a BIORUPTOR (Diagenode), and kept in the fixating solution for 15 minutes longer. Embryos were washed 15 minutes three times in PBT (0.4% Triton), prehybridized in PBT containing 1% Bovine Serum Albumin (PBTB) for 20 minutes, and incubated overnight at 4°C with primary antibodies diluted in PBTB. After 3x15 minute washes in PBT 0.4%, embryos were incubated in PBTB with secondary antibodies for three hours minimum. After 2x15 minute washes with PBS and one with DAPI (Sigma) (1:1000 in PBS) for 20 minutes, specimens were mounted in Citifluor (Biovalley). First instar larvae were collected on petri dishes 24 to 28h AEL and treated as embryos, except for sonication (2x12 sec). Second and third instar larvae were sexed and female fat bodies were hand dissected. They were treated like embryos, but with no sonication, and fixation was in PBT 0.4% instead of 0.8%. Gonads were observed under a Leica TCS SP5 reverse confocal microscope. Z-stacks of the PGCs-containing gonad sections were acquired and analyzed using Fiji software.

### Antibodies

Primary antibodies were from DSHB: rat anti-Vasa (1:1000), mouse anti-HP1a (1:1000), mouse anti-1B1 (1:1000), and from Rockland immunochemicals Inc.: rabbit anti- $\beta$ galactosidase (1:1000). Rabbit anti-Piwi (1:1000) and rabbit anti-Ago3 (1:1000) were a kind gift from T. Kai, rabbit anti-Aub (1:1000) and guinea-pig anti-Rhino (1:500): were a kind gift from B. Theurkauf. Secondary antibodies were either from Invitrogen: Alexa Fluor 633 goat anti-rat IgG (H +L) (1:1000), Alexa Fluor 647 goat anti-rat IgM ( $\mu$ chaine) (1:1000), Alexa Fluor 594 goat anti-rat IgM ( $\mu$ chaine) (1:1000), Alexa Fluor 568 goat anti-rabbit IgG (H+L) (1:1000); from Life Technologies: Alexa Fluor 594 goat anti-rabbit IgG (H+L) (1:1000), Alexa Fluor 594 goat anti-mouse IgG (H+L) (1:1000), Alexa Fluor 594 goat anti-guinea-pig

IgG (H+L) (1:1000); or from Chromotek: GFP-Booster\_Atto488 (1:1000).

#### RNA extraction and Deep-Sequencing

Experiments were performed as in (HERMANT *et al.* 2015) except that hand-dissected fat bodies or carcasses from one hundred third instar female larvae were used instead of adult ovaries.

Strains are available upon request. Raw data from deep sequencings are available at: <https://mississippi.snv.jussieu.fr/u/anbo/h/marie2016>.

## RESULTS

*A transgene-based assay reveals homology-dependent repression in PGCs during development*

To detect homology-dependent *trans*-repression in PGCs during development, we used a transgene combination that leads to reliable and specific expression of a reporter protein in the germline. A  $P\{UASp\text{-}GFPS65C\text{-}\alpha\text{Tub}84B\}$  construct (hereafter  $UASpGFP$ , Fig 1A) driven by a maternally inherited  $PBac\{GreenEye.nosGAL4\}$  transgene (hereafter  $nosGAL4$ , Fig 1B) strongly expresses GFP in almost all germ cells from embryonic to pupal stages (99,4% GFP-positive cells,  $n= 2601$ , Fig 1Cb-Db-Eb-Fb-Gb). To test if  $UASpGFP$  expression could be repressed by homologous piRNAs, we used two transgenic lines sharing sequence identity with  $UASpGFP$  and inserted into subtelomeric regions. Strain  $P\text{-}1152$  contains two insertions of a  $P\text{-}lacZ$  construct ( $P\{lArB\}$  Fig 1A) in a subtelomeric piRNA cluster at the X-chromosome tip (1A site), and produces abundant piRNAs homologous to  $P$ ,  $lacZ$  and  $rosy$  sequences (DE VANSSAY *et al.* 2012; MUERDTER *et al.* 2012). piRNAs generated from telomeric  $P\{lArB\}$  copies can silence in *trans* other  $P\text{-}lacZ$  transgenes inserted into euchromatic loci, a phenomenon called *trans*-silencing effect (TSE) (ROCHE AND RIO 1998; RONSSERAY *et al.* 2003; JOSSE *et al.* 2007; JOSSE *et al.* 2008). The second line,  $RS3$ , contains an insertion of construct  $P\{RS3\}CB\text{-}0686\text{-}3$  (Fig 1A) into a subtelomeric piRNA cluster at the 3R tip (100E3 site). It also produces abundant piRNAs and can silence  $P\text{-}lacZ$  transgenes in *trans* in adult ovaries (DUFOURT *et al.* 2014; HERMANT *et al.* 2015). Sequence identity between the silencer transgenes ( $P\text{-}1152$  and  $RS3$ ) and the targeted transcript ( $UASpGFP$ ) consists in 500 bp-long  $P$  sequences (Fig 1A). Immunostainings of gonads harboring maternally inherited telomeric silencer and  $nosGAL4$  transgenes and a paternally inherited  $UASpGFP$  transgene (Fig 1B) reveal that both telomeric transgenes can silence expression of  $UASpGFP$  in PGCs at all developmental stages, from late embryos to pupae (Fig 1C-G). For each gonad, we determined the number of PGCs with anti-VASA staining, and the proportion of PGCs expressing GFP (Fig 1H). Repression is partial with  $P\text{-}1152$  but complete with  $RS3$  (Fig 1C-H). Interestingly, the partial repression observed with  $P\text{-}1152$  resembles variegation, a phenomenon we will address below.  $RS3$ -mediated repression shows a maternal effect since in reciprocal crosses, with a paternally inherited  $RS3$  telomeric transgene, no significant silencing was observed (94.8% GFP-positive cells,  $n=272$ ). This maternal effect is characteristic of the *trans*-silencing phenomenon. Indeed, as the male gamete deposits no piRNAs, F1 females are unable to produce sufficient *de novo* piRNAs to ensure silencing (JOSSE *et al.* 2007; BRENNECKE *et al.* 2008). In conclusion, we show that a canonical *trans*-silencing phenomenon occurs in female germ cells at the embryonic, larval and pupal stages, and that incomplete silencing can be observed as early as in the embryonic PGCs.

*Developmental silencing is sensitive to germline knockdown (GLKD) of piRNA-mediated silencing genes*

In adults, a number of genes required for piRNA-mediated silencing have been characterized (for reviews, see (IWASAKI *et al.* 2015; HIRAKATA AND SIOMI 2016). To test whether these genes were also required for repression throughout development of the gonad, we used TRiP lines (NI *et al.* 2011). Under a GAL4-dependent promoter, such lines express a modified miRNA (shRNA) that can induce specific knockdown of a given gene. In our experiments, the *nosGAL4* driver ensured specifically shRNA-mediated germline knockdown (GLKD). We immunostained L3 gonads carrying a maternally inherited *RS3* telomeric silencer, *nosGAL4* driver and *UASpGFP* reporter transgene and different paternally inherited TRiP transgenes, to compare GFP repression levels, and identify TRiP lines efficiently affecting *trans*-silencing. *white* GLKD gonads, in which *RS3* mediates strong GFP repression, served as negative control (Fig 2c, h, m, r, w). By contrast, GLKD of core piRNA genes (*ago3*, *aubergine*, *rhino*, *piwi*) resulted in disappearance of the corresponding protein as well as expression of GFP (Fig 2b', c', g', h', l', m', q', r', v', w'). This was particularly clear for *ago3*, *aubergine* and *rhino*, whose expression is restricted to the PGCs. Expression of *piwi* was also detected in the somatic-origin cells that surround PGCs (Fig 2q, q'). In a *piwi* GLKD context, VASA staining (Fig 2s', t') confirmed that disappearance of the protein was specifically observed in the PGCs. Hence, *ago3*, *aubergine*, *rhino* and *piwi* are required for GFP repression, probably through production of piRNAs. Interestingly, similar results were obtained with GLKD of *Su(var)205* that encodes HP1a, a conserved eukaryotic chromosomal protein implied in gene silencing through interaction with di- and tri-methylated histone 3 Lys9 (H3K9me<sub>2,3</sub>) ((GREWAL AND JIA 2007), Fig 2v', w', x', y'). HP1a's role in telomere capping (FANTI *et al.* 1998) and/or its presence on the subtelomeric regions (FRYDRYCHOVA *et al.* 2008) could possibly be necessary for the ability of telomeric silencers to produce piRNAs. Agreeing with this, our lab previously showed that a half dosage of HP1a abolishes *P* element repression capacities mediated by telomeric *P* copies inserted in subtelomeric heterochromatin (RONSSERAY *et al.* 1996). Furthermore, using RNase protection assays, we found that production of subtelomeric transgenic small RNAs in adult ovaries is sensitive to HP1a dosage (TODESCHINI *et al.* 2010). HP1a could, also, be required at the *UASpGFP* target locus to establish transcriptional gene silencing in larvae, as described for endogenous TEs in adults (WANG AND ELGIN 2011; LE THOMAS *et al.* 2013; SIENSKI *et al.* 2015).

To avoid off-target effects and false positives, we tested, whenever possible, different TRiP lines inactivating each tested gene (Table 1). Numbers of anti-VASA staining PGCs and

proportion of GFP-positive PGCs were again calculated for each gonad (Fig 2 first column and Table 1). With this quantitative approach, we can detect variation in response levels of different TRiP lines targeting a given gene (Table 1). When GLKD of a gene was correlated to GFP appearance, we concluded that this gene is involved in repression, especially when different TRiP lines behaved similarly. An extensive study of 33 TRiP lines targeting 18 genes showed that 14 out of 15 tested genes implicated in the piRNAs pathway are required for repression. Hence, in addition to *ago3*, *aubergine*, *piwi*, *rhino* and *Su(var)205*, the genes *armitage*, *cutoff*, *H2Av*, *qin*, *maelstrom*, *tejas*, *tsunagi*, *vasa* and *zucchini* are also required to fully maintain repression during the third larval instar (Table 1 and Fig S1). As the three *eggless* TRiP lines behaved differentially and two did not break repression, we cannot conclude whether *eggless* is involved in repression in female L3 gonads (Table 1). Finally, we found that GLKD of *ago1*, *ago2* or *dicer2*, involved in miRNA and siRNA pathways, had no effect on repression (Table 1). Altogether, these results show that the main actors of all the different steps of the piRNA-mediated silencing identified in adult ovaries participate in silencing during development of the germline.

#### *Larval PGCs contain piRNAs with a ping-pong signature*

We looked for small RNAs in larval PGCs from reciprocal crosses between *RS3* and *w<sup>1118</sup>*. In adults, piRNA production and repression capacities of a telomeric silencer depend on maternally inherited homologous piRNAs (JOSSE *et al.* 2007; BRENNECKE *et al.* 2008). Hence, transgenic piRNAs should be detected only when the transgene is maternally inherited, whereas piRNAs produced by endogenous clusters present in all females should be detected in progeny of both crosses. Larval gonads are embedded in a fat body. The PGC-containing fat body from  $\approx 100$  third instar female larvae was hand dissected from carcasses, and small RNAs (smRNAs) were extracted from each tissue (heads were removed from carcasses). smRNAs from both tissues were deep-sequenced (Table S1). Numerous smRNAs complementary to the 42AB region, a strong piRNA-producing locus in the adult germline, and some complementary to the 3R telomere region, were detected in gonad-containing fat body fraction (Fig 3A-B). These smRNAs from both regions harbored a ping-pong signature (a high number of pairs of sense and antisense piRNAs which overlap exactly by ten nucleotides) and presented a typical size profile from 23 to 29 nt that are fully compatible with those of piRNAs (Fig 3A-B). Interestingly, 42AB smRNAs, but not smRNAs from the 3R telomere, presented a clear uridine bias at the 5' end (1U bias), a piRNA characteristic (BRENNECKE *et al.* 2007), Fig 3A-B). On the opposite, carcass fractions contained no large numbers of smRNAs, and these smRNAs were totally devoid of ping-pong signature or 1U bias (Fig 3A-B). Sense and antisense 23-29-nt smRNAs homologous to the *RS3* telomeric

transgene were specific to progeny with maternally inherited *RS3*, fitting well with the idea that these smRNAs are indeed piRNAs (Fig 3C). They also presented a strong 1U bias but their number was too low to detect a ping-pong signature.

*Global larval piRNA production relies on Rhino, while telomeric piRNAs specifically depend on HP1a*

To determine whether the detected piRNAs were *rhino*-dependent and came from PGCs, we analyzed smRNAs extracted from *rhino* GLKD larval fat bodies that had maternally inherited *RS3*. Compared to the *w* GLKD control, in which we found *bona fide* piRNAs as in the *RS3* maternal cross, the *rhino* GLKD context presented a dramatic decrease in 23-29nt smRNAs complementary to the *RS3* transgene (16,1% remaining smRNAs, Fig 4C) and 3R subtelomeric sequences (20,3% remaining smRNAs, Fig 4B). 42AB locus smRNAs were less decreased (50,4% remaining smRNAs, Fig 4A). As *rhino* is specifically expressed in PGCs (Fig 2I), and as the major part of the smRNAs coming from *RS3* and the 3R tip appear to be *rhino*-dependent, we conclude that these smRNAs likely come from the PGCs. Half of the 23-29-nt smRNAs from the 42AB region are also *rhino*-dependent and could thus come from PGCs (Fig 4A). However, we find that *rhino*-independent smRNAs do exist and that they could originate from germinal or somatic tissues. Since these smRNAs also present a strong ping-pong signature (Fig 4A), we favor the possibility that they come from PGCs. Taken together, our data strongly suggest that *bona fide* piRNAs are produced in third instar larval PGCs and that the loss of repression in the *rhino* mutant context (Fig 2m') is correlated with the loss of *RS3* transgenic piRNAs.

Next, we analyzed smRNAs extracted from *Su(var)205* GLKD larval fat bodies that had maternally inherited *RS3*. In this mutant context, 23-29nt smRNAs matching the 42AB locus appeared little affected (86,6% remaining smRNAs compared to *w* GLKD control) whereas telomeric 23-29nt smRNAs strongly decreased (19,9% remaining *RS3* transgene smRNAs, Fig 4C, and 27,3% remaining 3R subtelomeric sequence smRNAs, Fig 4B, compared to *w* GLKD control). Our results thus indicate that HP1a is required for the production of telomeric piRNAs but not of pericentric ones. To date, HP1a has mainly been shown to act as an effector of the piRNA pathway, required for locking target expression through a transcriptional gene silencing pathway (WANG AND ELGIN 2011; SIENSKI *et al.* 2015). However, as mentioned above, production of subtelomeric transgenic small RNAs in adult ovaries is sensitive to HP1a dosage (TODESCHINI *et al.* 2010). Our present results thus confirm that HP1a may act at the level of

piRNA production. Interestingly, we show that HP1a and Rhino proteins mainly colocalize in PGC nuclei but that the presence of one does not depend on that of other (Fig S2).

In order to extend the analysis to endogenous TEs, we compared the amounts of 23-29-nt antisense small RNAs matching known TEs in control GLKD and *rhino* GLKD contexts. Based on comparison of ovarian and embryonic piRNAs, TEs have been classified as producing piRNAs from the soma or the germline (MALONE *et al.* 2009). Figure 4D shows that most piRNAs originating from germline TEs were sensitive to *rhino* GLKD (in red in Fig 4D). By contrast, all piRNAs coming from TEs having somatic sources were highly expressed and insensitive to *rhino* GLKD (in green and yellow in Fig 4D). As piRNAs coming from these somatic TEs are known to be depleted in the embryo compared to the ovaries (MALONE *et al.* 2009), our results indicate that *de novo*-produced somatic piRNAs are present in large amounts at the L3 stage. Interestingly, when we compared amounts of 23-29-nt small RNAs matching known TEs in control GLKD and *Su(var)205* GLKD contexts, we observed that piRNA production of only 3 TEs were sensitive to HP1a (Fig 4E). *HeT-A* and *TAHRE* elements constitute telomere ends while piRNA production from the *I-element*, a recent invader of the *Drosophila* genome, could have different sources, including telomeric ones. In addition to the strong loss of piRNAs matching 3R subtelomeric sequences in *Su(var)205* GLKD (Fig 4B), these observations thus reinforce the idea that HP1a is specifically required for telomeric piRNA production.

#### *Incomplete silencing reveals a cellular memory from embryonic PGCs to adult GSCs*

The first observations of incomplete piRNA-mediated silencing were made in adult ovaries (RONSSERAY *et al.* 2003; JOSSE *et al.* 2007). Ovarioles presented fully repressed or non-repressed egg chambers, which were apparently randomly distributed (Fig 5A). Still, inside a given egg chamber, all nurse cells were almost always homogenous, suggesting that the decision to repress (or not) the target is taken at a one-cell stage, and is then maintained throughout formation and maturation of the egg chamber. In the present analyses of piRNA-mediated repression during development, we clearly show that incomplete silencing can occur as early as in the embryo (Fig 1Cf) and is detected at all subsequent developmental stages. These observations raise the question whether PGCs acquire a stable piRNA-mediated repression state of the target in the embryo, that could be ON or OFF, and that is maintained through development until the adult stage (*epigenetic lock* hypothesis). Alternatively, is the repression state plastic during development, changing from one state to the other, showing repression in embryos then expression at later stages, or vice versa (*plastic repression* hypothesis)? To discriminate between these hypotheses, we developed a genetic system to visualize the repression state of cells derived



from one or two embryonic PGCs: we coupled a heat-shock-induced clonal cellular lineage, revealed by flipped-out GFP expression, to *P-1152*-mediated incomplete silencing detected by  $\beta$ -Galactosidase ( $\beta$ GAL) expression of a target transgene (Fig 5B). Under the *epigenetic lock* hypothesis, the repression state is expected to be homogeneous among progeny of a single embryonic PGC, while under a fully *plastic repression* hypothesis, the proportion of  $\beta$ GAL+ versus  $\beta$ GAL- cells is expected to be the same among GFP+ clonal cells as in the total PGC population of the gonad.

We first heat shocked embryos for 15 to 20 minutes at 37°C and let them develop at 25°C until the late third larval instar (Fig 5C). Immunostainings of L3 gonads identified PGCs (VASA-positive cells), and among these GFP-positive (GFP+) cells were counted. Since PGCs undergo three mitoses between embryo and late third larval instar, the number of GFP+ cells detected in late L3 allows to estimate the number of PGCs that had been flipped-out in the embryo: for example, eight GFP+ PGCs in late L3 would derive from a single flipped-out PGC in the embryo. We examined 23 late L3 gonads, heat-shocked in embryo and presenting incomplete  $\beta$ GAL repression. Among ten gonads with eight GFP+ cells, three presented 100%  $\beta$ GAL+ staining of GFP+ cells and seven presented 100%  $\beta$ GAL- staining of GFP+ cells. These data definitely argue for homogeneity of repression among progeny of a single embryonic PGC. Note that under a fully *plastic repression* hypothesis, i.e. independence of cells for their repression state, the probability of such a distribution is  $p < 1,9 \cdot 10^{-22}$ . Thirteen gonads had more than eight but less than seventeen GFP+ PGCs, thus likely derived from two flipped-out embryonic PGCs. Among these, two had 100%  $\beta$ GAL+ staining among GFP+ cells (see Fig 5D), four had 100%  $\beta$ GAL- staining among GFP+ cells, and five presented 50%+/50%-  $\beta$ GAL staining, favoring again repression homogeneity. Of the two last gonads, one presented five  $\beta$ GAL+ cells and nine  $\beta$ GAL- cells, and the other, two  $\beta$ GAL+ cells and eight  $\beta$ GAL- cells. Three embryonic flipped-out PGCs might sometimes give rise to less than 17 PGCs in late L3, thus explaining the first distribution, and some PGCs might divide at a lower frequency, potentially explaining the second distribution. Alternatively, the repression state could be plastic in these 2 gonads. Taken together, the results show that 21 L3 gonads out of 23 behaved in a way that is fully compatible with repression homogeneity among cells derived from a single embryonic PGC, supporting the *epigenetic lock* hypothesis.

Next we asked whether this epigenetic lock of germline piRNA-mediated repression could persist until the adult stage. To test this, late third instar larvae were heat-shocked for 15 to 20 minutes at 37°C – long enough to induce flip-out in approximately 50% of PGCs - and left to develop at 25°C until emergence (Fig 5C). Each ovariole contains two or three germline stem cells

(GSCs) (WIESCHAUS AND SZABAD 1979), and the proportion of germaria containing 2 *versus* 3 GSCs varies from 20 to 50% (WIESCHAUS AND SZABAD 1979; MARGOLIS AND SPRADLING 1995). We counted PGC numbers by 1B1 immunofluorescence - that specifically label the spectrosome, a dotted structure present in GSCs and cystoblasts - and estimated the proportion of 2-GSC containing germaria to about 50% in our experimental conditions (n=61). Thus, about half of the ovarioles we tested had egg chambers derived from only two GSCs. Under a fully *plastic repression* hypothesis, no correlation is expected between heat-shock induced GFP+ cells and silencing. On the contrary, under the *epigenetic lock* hypothesis (i.e. if L3 PGCs maintain and transmit their repression state until the egg chamber), we expect specific patterns of heat-shock induced GFP+ cells in ovarioles with heterogeneity for both GFP and  $\beta$ GAL expression. First, GFP/ $\beta$ GAL staining "exclusion" is expected if one of the GSC is positive for GFP and negative for  $\beta$ GAL while the other GSC is negative for GFP and positive for  $\beta$ GAL. Alternatively, GFP/ $\beta$ GAL staining "overlap" is expected if one PGC is positive for both GFP and  $\beta$ GAL and the other one is negative for both. We focused on 61 ovarioles that present simultaneously heterogeneity for GFP and  $\beta$ GAL staining. More than the 50% expected 2-GSC containing ovarioles presented specific patterns (63,9%, "overlap" n=28 plus "exclusion" n=11), which again supports the *epigenetic lock* hypothesis (Fig 5E). The remaining ovarioles presented mixed pattern (n=22). The mixed pattern ovarioles could still be due to plastic repression, but we favor the hypothesis that they reflect the high proportion of germaria with 3 GSCs (about 50%) that could be heterogeneous. Moreover, under the *plastic repression* hypothesis, ovarioles with four GFP/GAL combinations should occur, but were never observed. Taken together, these studies performed from embryo to L3 and from L3 to adults suggest strongly that a cellular memory of the repression state exists from embryonic PGCs to adult GSCs.

## DISCUSSION

Here, we investigated piRNA-mediated repression throughout female germline development, from embryonic to pupal PGCs in *Drosophila*. Using transgenes inserted into telomeric piRNA clusters, we observed silencing in *trans* of partially homologous reporter transgenes located elsewhere in the genome. We show that typical piRNA genes are required for reporter gene repression in larval PGC while si- or miRNA genes are not. In addition, small RNAs corresponding to piRNAs are likely present in larval gonads.

In the GLKD experiments we present here, the expression of the modified miRNA that will knock down the gene of interest depends on a paternally inherited *nos-gal4* transgene. The knock down is thus zygotically induced. The fact that some of the larval piRNAs were sensitive to

a zygotically-induced GLKD of *rhino*, a gene required for the production of primary piRNAs, strongly suggests that these piRNAs are zygotic piRNAs. By extension, we propose that transgenic piRNAs, which were found sensitive to *rhino* GLKD, are also *de novo* piRNAs. By contrast, the piRNAs that were insensitive to *rhino* GLKD could be of maternal or zygotic origin. Importantly, the fact that under *rhino* GLKD the repression is abolished (Fig 2m') demonstrates that maternally inherited piRNAs are not sufficient to establish a complete repression in L3 PGCs.

Interestingly, *H2Av* and *Su(var)205* GLKDs exhibit a strong derepression of the GFP target transgene (FIG 2 and Table 1). HP1a, encoded by the *Su(var)205* gene, could play a role at telomeres to produce piRNAs, and/or a role on euchromatin to silence the target transgene through local heterochromatinization. We show here that HP1a is required to specifically produce telomeric piRNAs, reinforcing the first proposition (Fig 4C, E). However we cannot exclude that HP1a is also required for transcriptional gene silencing of the euchromatic transgene. Immunostaining of polytene chromosomes of third instar larvae revealed that H2Av localized at telomere tips (VAN DAAL AND ELGIN 1992; RONG 2008). Mutations of *H2Av* suppress position effect variegation (PEV), reduce repressive chromatin marks such as H3K9me3 or H4K12Ac, and reduce HP1a recruitment to centromeric regions (SWAMINATHAN *et al.* 2005). As for HP1a, we propose that H2Av could be required for the establishment of telomeric heterochromatin, a necessary step for transcriptional activation of telomeric piRNA clusters.

A striking result was that we consistently observed incomplete silencing during development using *P-1152* as a silencer. Our clonal analysis revealed that incomplete silencing likely follows the epigenetic lock hypothesis. Incomplete silencing appears to rely on two steps. The first is whether to establish, or not, target repression. This step likely occurs very early during embryogenesis and could depend on the amount of maternally inherited homologous piRNAs. It could also depend on the strength of the targeted promoter, or on genomic location of the target; indeed, we previously showed that the level of incomplete silencing of a given telomeric silencer changes with its euchromatic target in adult ovaries (JOSSE *et al.* 2008). The second step is the "lock" of the repression state: repression is either active (ON) or not (OFF). Since we observed that both the ON and the OFF state are maintained through PGC development, there seems to be no cumulative repressive effect of piRNAs over time during development, as this would have resulted in tardive establishment of repression. The data presented here thus suggest that the lock is already in place in late embryos. What could be the molecular nature of this lock? Does it concern telomeric piRNA-producing loci that become unable to produce piRNA precursor transcripts, thereby being unable to repress targets? Or are targets epigenetically locked by PIWI-piRNA complexes during larval stages? We confirm here that HP1a is required to produce wild

type amounts of telomeric piRNAs (TODESCHINI *et al.* 2010), Fig 5C, F). It is thus tempting to propose that *P-1152* resides in a telomeric region where the level of HP1a fluctuates, mimicking PEV. PEV occurs when a euchromatic sequence is relocated accidentally next to heterochromatic regions (ELGIN AND REUTER 2013). The expression of relocated genes could then be subject to a stochastic and bimodal (ON/OFF) expression due to the extension of heterochromatin proteins, such as HP1a, over the new adjacent gene sequences. This phenomenon is cell autonomous, thus possibly producing a variegated phenotype. By extension, we could imagine that in cells where HP1a is present over the telomeric transgene, enough piRNAs are produced to establish repression of its euchromatic target, whereas in others PGCs, a telomeric transgene devoid of HP1a could produce less piRNAs, resulting in non repression of its euchromatic target. Such a "piRNA production variegation" phenomenon could fit well with the variegating repression phenotype we observed with the *P-1152* silencer. However, was this hypothesis true, then progeny derived from OFF egg chambers could not inherit transgenic piRNAs and should thus be devoid of silencing properties. As  $\approx 50\%$  of the egg chambers are OFF, this suggests that  $\approx 50\%$  of the embryos should be devoid of silencing capabilities at the adult stage. But during years of maintaining various *P-1152* stocks we have never observed such instability: 100% of the progeny show  $\approx 50\%$  (variegated) repression. This suggests that piRNAs are produced in all egg chambers, and that piRNA production capacity is not bimodal (ON or OFF). Rather, different thresholds of piRNAs could be required to determine a locus as piRNA producer or to establish euchromatic target repression. We thus propose that the lock might involve early-established chromatin modifications of the euchromatic target itself, such as H3K9me3 and HP1a which have been observed on many euchromatic TE insertions in correlation with piRNA-mediated silencing (WANG AND ELGIN 2011; SIENSKI *et al.* 2012; LE THOMAS *et al.* 2013; ROZHKOV *et al.* 2013; PEZIC *et al.* 2014; SIENSKI *et al.* 2015; YU *et al.* 2015). Deposition of these marks could depend on a piRNA threshold that, in the case of *P-1152*, would depend on HP1a.

Our data suggest that piRNAs are produced in all egg chambers. Despite their presence, some targets remain unrepressed (blue egg chambers). What would be particularly interesting is the identification of specific chromatin factors that maintain the repression OFF despite the presence of homologous piRNAs. The products of *nanos* and others genes required for germ cell development are selectively decayed through Aub interactions in the somatic part of the embryo, while they are protected from piRNAs in the germlasm (ROUGET *et al.* 2010; BARCKMANN *et al.* 2015). Our data might help to understand how such genes can remain transcriptionally active despite the presence of numerous homologous piRNAs.

ACKNOWLEDGMENTS

We thank Toshi Kai, Bill Theurkauf and Jean-René Huynh for sharing reagents. We also especially thank J.R. Huynh for helpful suggestions. We thank Richard Schwartzmann for help with confocal microscopy. Confocal work was carried out at the Institute of Biology Paris-Seine Imaging Facility. We are grateful to Christophe Antoniewski and Marius van den Beek for developing Galaxy platform (CNRS, Institut de Biologie Paris- Seine, ARTbio Bioinformatics Analysis Facility). We thank Ritha Zamy for technical assistance, members of the Ronsseray lab for discussions and Neel Randsholt for critical reading of the manuscript. This work was supported by a fellowship from the *Ministère de l'Enseignement Supérieur et de la Recherche* to P.M. and by grants from the *Fondation pour la Recherche Médicale* (FRM DEP20131128532), the *Association pour la Recherche contre le Cancer* (ARC, SFI2012205921, SFI20131200470) and from the *Agence Nationale pour la Recherche* (ANR, project "plastisipi") to S.R.

#### COMPETING INTERESTS

The authors declare that no competing interests exist.

#### AUTHOR CONTRIBUTIONS

P.M., S.R. and A.B. designed the experiments, interpreted the results and wrote the manuscript, P.M. performed the experiments, A.B. analyzed the smRNA data.

## REFERENCES

- Aravin, A. A., G. J. Hannon and J. Brennecke, 2007 The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318: 761-764.
- Barckmann, B., S. Pierson, J. Dufourt, C. Papin, C. Armenise *et al.*, 2015 Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo. *Cell Rep* 12: 1205-1216.
- Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis *et al.*, 2007 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128: 1089-1103.
- Brennecke, J., C. D. Malone, A. A. Aravin, R. Sachidanandam, A. Stark *et al.*, 2008 An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322: 1387-1392.
- Dansereau, D. A., and P. Lasko, 2008 The development of germline stem cells in *Drosophila*. *Methods Mol Biol* 450: 3-26.
- de Vanssay, A., A. L. Bouge, A. Boivin, C. Hermant, L. Teyssset *et al.*, 2012 Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* 490: 112-115.
- Dufourt, J., C. Dennis, A. Boivin, N. Gueguen, E. Theron *et al.*, 2014 Spatio-temporal requirements for transposable element piRNA-mediated silencing during *Drosophila* oogenesis. *Nucleic Acids Res* 42: 2512-2524.
- Elgin, S. C., and G. Reuter, 2013 Position-effect variegation, heterochromatin formation, and gene silencing in *Drosophila*. *Cold Spring Harb Perspect Biol* 5: a017780.
- Fanti, L., G. Giovinazzo, M. Berloco and S. Pimpinelli, 1998 The heterochromatin protein 1 prevents telomere fusions in *Drosophila*. *Mol Cell* 2: 527-538.
- Frydrychova, R. C., J. M. Mason and T. K. Archer, 2008 HP1 is distributed within distinct chromatin domains at *Drosophila* telomeres. *Genetics* 180: 121-131.
- Gilboa, L., and R. Lehmann, 2006 Soma-germline interactions coordinate homeostasis and growth in the *Drosophila* gonad. *Nature* 443: 97-100.
- Grewal, S. I., and S. Jia, 2007 Heterochromatin revisited. *Nat Rev Genet* 8: 35-46.
- Gunawardane, L. S., K. Saito, K. M. Nishida, K. Miyoshi, Y. Kawamura *et al.*, 2007 A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315: 1587-1590.
- Hermant, C., A. Boivin, L. Teyssset, V. Delmarre, A. Asif-Laidin *et al.*, 2015 Paramutation in *Drosophila* Requires Both Nuclear and Cytoplasmic Actors of the piRNA Pathway and Induces Cis-spreading of piRNA Production. *Genetics* 201: 1381-1396.
- Hirakata, S., and M. C. Siomi, 2016 piRNA biogenesis in the germline: From transcription of piRNA genomic sources to piRNA maturation. *Biochim Biophys Acta* 1859: 82-92.
- Huynh, J. R., and D. St Johnston, 2004 The origin of asymmetry: early polarisation of the *Drosophila* germline cyst and oocyte. *Curr Biol* 14: R438-449.
- Iwasaki, Y. W., M. C. Siomi and H. Siomi, 2015 PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu Rev Biochem* 84: 405-433.
- Josse, T., C. Maurel-Zaffran, A. de Vanssay, L. Teyssset, A. L. Todeschini *et al.*, 2008 Telomeric trans-silencing in *Drosophila melanogaster*: tissue specificity, development and functional interactions between non-homologous telomeres. *PLoS One* 3: e3249.
- Josse, T., L. Teyssset, A. L. Todeschini, C. M. Sidor, D. Anxolabehere *et al.*, 2007 Telomeric trans-silencing: an epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genet* 3: 1633-1643.
- Le Thomas, A., A. K. Rogers, A. Webster, G. K. Marinov, S. E. Liao *et al.*, 2013 Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* 27: 390-399.
- Malone, C. D., J. Brennecke, M. Dus, A. Stark, W. R. McCombie *et al.*, 2009 Specialized

- piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137: 522-535.
- Margolis, J., and A. Spradling, 1995 Identification and behavior of epithelial stem cells in the *Drosophila* ovary. *Development* 121: 3797-3807.
- Mohn, F., G. Sienski, D. Handler and J. Brennecke, 2014 The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* 157: 1364-1379.
- Muerdter, F., I. Olovnikov, A. Molaro, N. V. Rozhkov, B. Czech *et al.*, 2012 Production of artificial piRNAs in flies and mice. *RNA* 18: 42-52.
- Ni, J. Q., R. Zhou, B. Czech, L. P. Liu, L. Holderbaum *et al.*, 2011 A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat Methods* 8: 405-407.
- Pezic, D., S. A. Manakov, R. Sachidanandam and A. A. Aravin, 2014 piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev* 28: 1410-1428.
- Roche, S. E., and D. C. Rio, 1998 Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the *Drosophila* Polycomb group gene, Enhancer of zeste. *Genetics* 149: 1839-1855.
- Rong, Y. S., 2008 Loss of the histone variant H2A.Z restores capping to checkpoint-defective telomeres in *Drosophila*. *Genetics* 180: 1869-1875.
- Ronsseray, S., T. Josse, A. Boivin and D. Anxolabehere, 2003 Telomeric transgenes and trans-silencing in *Drosophila*. *Genetica* 117: 327-335.
- Ronsseray, S., M. Lehmann, D. Nouaud and D. Anxolabehere, 1996 The regulatory properties of autonomous subtelomeric P elements are sensitive to a Suppressor of variegation in *Drosophila melanogaster*. *Genetics* 143: 1663-1674.
- Rouget, C., C. Papin, A. Boureux, A. C. Meunier, B. Franco *et al.*, 2010 Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* 467: 1128-1132.
- Rozhkov, N. V., M. Hammell and G. J. Hannon, 2013 Multiple roles for Piwi in silencing *Drosophila* transposons. *Genes Dev* 27: 400-412.
- Senti, K. A., and J. Brennecke, 2010 The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet* 26: 499-509.
- Sienski, G., J. Batki, K. A. Senti, D. Donertas, L. Tirian *et al.*, 2015 Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. *Genes Dev* 29: 2258-2271.
- Sienski, G., D. Donertas and J. Brennecke, 2012 Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* 151: 964-980.
- Swaminathan, J., E. M. Baxter and V. G. Corces, 2005 The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of *Drosophila* heterochromatin. *Genes Dev* 19: 65-76.
- Todeschini, A. L., L. Teyssset, V. Delmarre and S. 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: e11032.
- van Daal, A., and S. C. Elgin, 1992 A histone variant, H2AvD, is essential in *Drosophila melanogaster*. *Mol Biol Cell* 3: 593-602.
- Wang, S. H., and S. C. Elgin, 2011 *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* 108: 21164-21169.
- Wieschaus, E., and J. Szabad, 1979 The development and function of the female germ line in *Drosophila melanogaster*: a cell lineage study. *Dev Biol* 68: 29-46.

- Yu, Y., J. Gu, Y. Jin, Y. Luo, J. B. Preall *et al.*, 2015 Panoramix enforces piRNA-dependent cotranscriptional silencing. *Science* 350: 339-342.
- Zhang, Z., J. Wang, N. Schultz, F. Zhang, S. S. Parhad *et al.*, 2014 The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* 157: 1353-1363.



## TABLES

Table 1. Quantification of repression in the GLKD screen.

Gene	TRiP line number	Fertile (F) Unfertile (U)	[Gonads] PGCs number	% Repression
<i>white (control)</i>	33644	F	[36] 4387	91,4
<b>GLKD against piRNA pathway</b>				
<i>Argonaute 3</i>	34815	F	[4] 470	99,7
<i>Argonaute 3</i>	44543	U	[7] 1056	34,2
<i>Argonaute 3</i>	35232	F	[3] 311	64,5
<i>Armitage</i>	34789	U	[6] 631	15,7
<i>Armitage</i>	35343	F	[6] 762	44,2
<i>Aubergine</i>	39026	U	[7] 924	66,3
<i>Aubergine</i>	33728	U	[11] 1847	50,2
<i>Aubergine</i>	35201	U	[4] 746	13,8
<i>Cutoff</i>	35182	F	[4] 286	51,8
<i>Cutoff</i>	35318	F	[3] 250	15,1
<i>Eggless</i>	32445	U	[11] 3502	91,5
<i>Eggless</i>	34803	F	[5] 734	94,4
<i>Eggless</i>	36797	F	[6] 1104	70,1
<i>His2A var</i>	34844	F	[5] 493	38,6
<i>His2A var</i>	44056	F	[7] 1094	35,0
<i>Kumo/Qin</i>	37475	F	[13] 1275	47,0
<i>Maelstrom</i>	34793	U	[6] 818	18,3
<i>Maelstrom</i>	35202	U	[6] 454	27,7
<i>Piwi</i>	33724	U	[5] 484	36,6
<i>Piwi</i>	37483	F	[5] 662	87,8
<i>Rhino</i>	34071	F	[6] 662	57,6
<i>Rhino</i>	35171	F	[4] 247	44,0
<i>Su(var)205</i>	36792	U	[7] 1333	48,8
<i>Su(var)205</i>	33400	U	[10] 893	52,3
<i>Tejas</i>	41929	U	[10] 1098	55,2
<i>Tejas</i>	36879	U	[12] 1595	46,5
<i>Tsunagi</i>	36585	U	[5] 232	15,1
<i>Vasa</i>	38924	U	[4] 644	79,7
<i>Vasa</i>	34950	U	[5] 887	48,0
<i>Vasa</i>	32434	F	[11] 1210	52,4
<i>Zucchini</i>	35227	U	[8] 1023	67,2
<i>Zucchini</i>	35228	F	[6] 650	52,2
<i>Zucchini</i>	36742	U	[12] 1895	57,0
<b>GLKD against mi- or siRNA pathways</b>				
<i>Ago1</i>	33727	NA	[5] 708	96,2
<i>Ago2</i>	34799	F	[4] 398	99,8
<i>Dicer2</i>	33656	F	[2] 202	98,2

TABLE 1. Quantification of target repression in the shRNA screen. Telomeric silencing is sensitive to germinal knock down (GLKD) of multiple genes involved in the piRNA pathway. We list the different shRNA lines tested for each gene, their fertility status, and the number of gonads

[in brackets] and PGC counted per genotype. Immunostainings of third instar larval gonads with anti-VASA antibody allowed PGC counting and anti-GFP allowed repression quantification. Control *white* GLKD (first row) shows strong *RS3*-induced repression of GFP expression (91.4%). By contrast, GLKD of many core piRNA genes results in strong decrease in GFP repression, revealing their requirement for repression. Genes from the miRNA and siRNA pathways show the specificity of the screen: *Ago1*, *Ago2* and *Dicer2* GLKD show no derepression of the GFP target.

## FIGURE LEGENDS

Figure 1. Homology dependent repression is active during female germline development. (A) piRNA-producing transgenes (*RS3* and *P-1152* silencers) are inserted into *3R* and *X* subtelomeric regions, respectively, and share 500 base pair identity with the *UASpGFP* transcript (drawn to scale). (B) Experimental mating scheme: progeny inherits maternally the piRNA-producing and the *nosGal4* transgenes and paternally the *UASpGFP* target transgene. (C) Immunostainings of late embryos (18-22h) allow to count Primordial Germ Cells (PGCs, anti-VASA antibodies, first column, red) and to visualize repression (anti-GFP, second column, green). In controls (*nos>pGFP*, b) almost all PGCs express GFP while the presence of telomeric transgenes leads to partial (with *P-1152*, f) or total (with *RS3*, j) GFP repression. (D, E, F, G) Similar observations were made at the L1, L2, L3 and pupal stages. (H) Quantitative analysis summarizing results. GFP repression is active from the late embryo and all along gonad development, partially ( $\approx 50\%$ ) with the *P-1152* silencer and completely with the *RS3* silencer. Mean number of PGCs  $\pm$  standard error and number of analyzed gonads (n) are given for each stage. Scale bar is 10  $\mu$ m.

Figure 2. Developmental silencing is sensitive to germline knock down (GLKD) of core partners of piRNA-mediated silencing. Immunostainings of third instar larval gonads with anti-VASA antibodies (first column, blue), with anti-GFP antibodies (third column, green) and with antibodies against the gene of interest (second column, red) are shown. Control *white* GLKD (first lane of each panel) shows strong *RS3*-induced repression of GFP expression (c, h, m, r, w). By contrast, GLKD of core piRNA genes results in non-detectable level of the corresponding protein (b', g', l', q', v') as well as presence of GFP (c', h', m', r', w'), revealing their requirement for repression. *piwi* is also expressed in somatic-origin intermingled cells that surround PGCs (q, q'). In a *piwi* GLKD context the protein disappears specifically in the PGCs, as revealed by VASA staining (s', t'). Similar results were obtained with GLKD of HP1a-encoding *Su(var)205* (v', w', x', y'). DAPI is shown in grey in the merge column. Scale bar is 10  $\mu$ m.

Figure 3. Detection of piRNAs in gonads of third instar larvae. SmRNA deep sequencing of gonad-containing fat bodies or carcasses of third instar larvae. Genotype and tissue fraction are indicated on top. Plots show the abundance of 20- to 29-nt smRNAs matching the 42AB locus (A), the 3R telomere (B), and the RS3 transgene (C). Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. Last lane (ping pong) show the number of overlapping sense-antisense small RNA pairs (Y-axis) in the subset of 23-28nt smRNAs (ping-pong signature, X-axis shows the overlap in nucleotides). Proportion of 23-29nt smRNAs beginning with a 5'-uridine (1U bias) is given (lower left corner).

Figure 4. Larval piRNAs are sensitive to *rhino* GLKD and telomeric piRNAs are sensitive to *Su(var)205* GLKD. Genotypes are indicated on top. Plots show the abundance of 20- to 29-nt smRNAs matching the 42AB locus (A), the 3R telomere (B), and the RS3 transgene (C). Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. Last lane (ping pong) shows the number of overlapping sense-antisense small RNA pairs (Y-axis) in the subset of 23-28nt smRNAs (ping-pong signature, X-axis shows the overlap in nucleotides). Proportion of 23-29nt smRNAs beginning with a 5'-uridine (1U bias) is given (lower left corner). Note that the two peaks (one sense, one antisense) that match the RS3 sequence near 3000 in the *w* GLKD background correspond to *w* modified miRNAs produced by the TRiP transgene. (D, E) Third larval instar small RNAs corresponding to natural TEs. Scatter plot depicts normalized 23-29nt antisense smRNAs mapping to annotated TEs in *rhino* GLKD versus control (*w*) GLKD log2 ratio (Y-axis) compared to control GLKD log2 amount (X-axis) (D). Similar plot depicts normalized 23-29nt antisense smRNAs mapping to annotated TEs in *Su(var)205* GLKD versus control (*w*) GLKD log2 ratio (Y-axis) compared to control GLKD log2 amount (X-axis) (E). Colors refer to classes of elements as defined in (MALONE *et al.* 2009): red, strong maternal deposition (germline source); yellow, intermediate and green, weak maternal deposition (predominantly somatic source).

Figure 5. Incomplete repression with cellular lineage reveals a cellular memory of the repression state. (A)  $\beta$ -Galactosidase overnight staining of adult ovarioles bearing maternally inherited telomeric silencer *P-1152* and paternally inherited euchromatic *P-lacZ* enhancer trap *BC69* reveals incomplete silencing: egg chambers show homogeneous ON or OFF staining. (B) Experimental mating scheme to produce progeny containing a telomeric silencer (*P-1152*), its euchromatic target (*P-lacZ*, *BC69*), a heat-shock-driven flipase (*hs-flp*) and a ubiquitous GFP whose expression depends on flip out of the (*FRT*)*STOP*(*FRT*) cassette. (C) In experiments (#1), GFP positive cells

were induced in 18-22-hour embryos and cell progeny was observed in L3 gonads; in experiments (#2), GFP positive cells were induced in late L3 and cell progeny was observed in adult ovaries. (D) Five z (1  $\mu\text{m}$ ) confocal plans of an L3 gonad reveal that all GFP positive PGCs are also  $\beta\text{GAL}$  positive (arrowheads). The upper right corner number indicates focal plane number (out of 31). Most somatic cells of the gonad are GFP positive but anti-VASA staining (in blue) specifically labels PGCs. In this gonad, *P-lacZ* repression by *P-1152* was estimated as 41.1% (n=56). Scale bar is 10  $\mu\text{m}$ . (E) Examples of patterned and mixed ovarioles with heterogeneous staining for both GFP and  $\beta\text{GAL}$ . Scale bar is 100  $\mu\text{m}$ .

## SUPPLEMENTARY DATA

Supplementary data includes a supplementary Table and two supplementary Figures.

TABLE S1. Annotation of smRNA libraries.

Library	GRH113	GRH114	GRH115	GRH116	GRH141	GRH142	GRH154
Parental origin	<i>w<sup>1118</sup> x RS3</i>	<i>RS3 x w<sup>1118</sup></i>	<i>w<sup>1118</sup> x RS3</i>	<i>RS3 x w<sup>1118</sup></i>	<i>RS3, nosgal4 x rhi<sup>GLKD</sup></i>	<i>RS3, nosgal4 x w<sup>GLKD</sup></i>	<i>RS3, nosgal4 x Su(var)205<sup>GLKD</sup></i>
Code Figure 3	<i>RS3</i> paternal	<i>RS3</i> maternal	<i>RS3</i> paternal	<i>RS3</i> maternal	<i>rhino</i> <i>GLKD</i>	<i>w</i> <i>GLKD</i>	<i>Su(var)205</i> <i>GLKD</i>
Tissues	fat bodies + gonads	fat bodies + gonads	carcasses	carcasses	fat bodies + gonads	fat bodies + gonads	fat bodies + gonads
Dmel_matched (Flybase r5.49)	<b>25576964</b>	<b>27619031</b>	<b>19072916</b>	<b>17531963</b>	<b>43157881</b>	<b>45540444</b>	<b>29067472</b>
miRNAs	5119292	5259136	195788	190249	11081050	8758989	8374713
tRNAs	2022883	1009886	401397	510053	2585604	2780667	2908870
ncRNAs	810342	1143553	1635562	1537058	1045613	1063459	303859
miscRNAs	14418738	16299122	13307916	11825247	24553017	26795350	14546798
Transposons	<b>270387</b>	<b>194371</b>	<b>61404</b>	<b>78255</b>	<b>356280</b>	<b>404981</b>	<b>217087</b>
Introns	527792	807826	746725	864245	882197	1278978	267667
Intergenic	564132	849103	336576	661979	821685	639535	383788
Transcripts	1600415	1908923	2361533	1835227	1316310	3352444	1625509
Remaining	242983	147111	26015	2965	516125	466041	439181
Effective depth	<b>9135343</b>	<b>10310023</b>	<b>5363603</b>	<b>5196663</b>	<b>16019260</b>	<b>15964427</b>	<b>11611804</b>
Normalization factor	<b>0,569</b>	<b>0,504</b>	<b>0,969</b>	<b>1,000</b>	<b>0,324</b>	<b>0,326</b>	<b>0,448</b>

TABLE S1. Annotation of smRNA libraries. SmRNAs were prepared from either gonad-containing fat bodies or carcasses of female L3 larvae, as indicated (Tissues row). Genotype is given in the Parental origin row (females x males). Each value for the different categories of sequences corresponds to the total number of sequence reads that matched either the *D. melanogaster* genome release 5.49 in each library (Dmel Matched row), or other annotated group of sequences. For comparisons, libraries were normalized (Normalization factor) relative to the number of sequence reads aligning with the *D. melanogaster* genome but not with tRNAs or miscRNAs (Effective depth). Library GRH116 has the lowest effective depth and was taken as reference to normalize the other libraries.

## SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Overview of piRNA genes required during development of germ cells. Scheme of the germline development: the rough number of primordial germ cells (PGCs, in blue) or germline stem cells (GSCs, in yellow) is indicated per gonad. In the center, a summary of the piRNA genes that we showed required during germline development: in yellow, those for which the germline

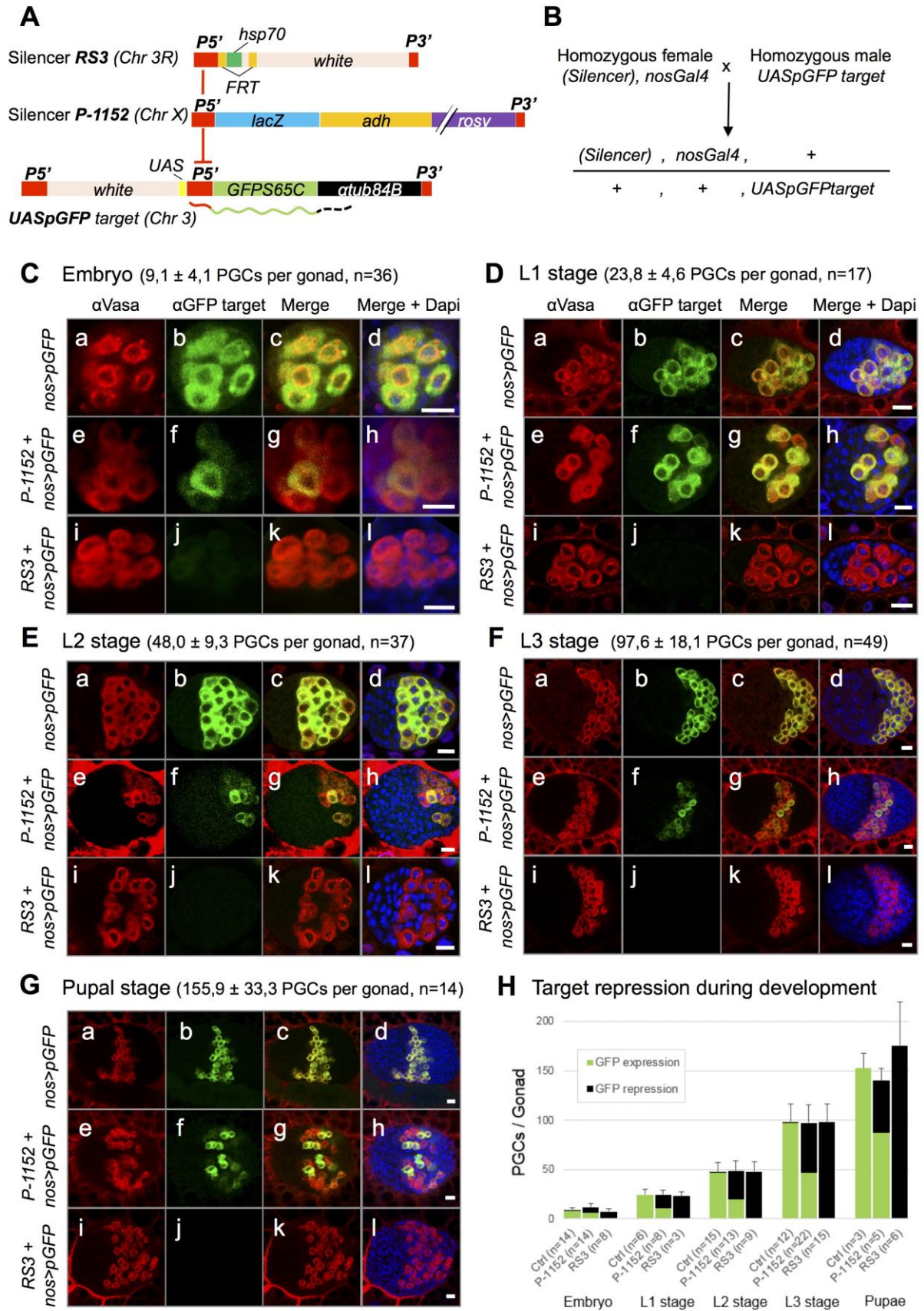
knockdown (GLKD) has been confirmed by immunostaining and in purple, those for which GLKDs lead to a derepression of the target reporter but without confirmation of the GLKD effect by immunostaining.

Figure S2. *rhino* GLKD and *Su(var)205* GLKD effects on third instar larvae gonads. Gonads are stained with anti-Vasa (blue), anti-Rhino (red) and anti-HP1a (green) antibodies. Upper panel, control (*w*) GLKD larval gonads express nuclear Rhino and HP1a proteins, Rhino is specific for PGC (specifically labeled with anti-Vasa) while HP1a is expressed in the nucleus of all gonadal cells. Enlargement of boxed region are shown: Rhino and HP1a colocalize at several foci in the PGC nucleus. Second and third panels show that *rhino* and *Su(var)205* GLKD are efficient and that they affect neither the quantity nor the localization of the other protein.

#### ARTICLE SUMMARY

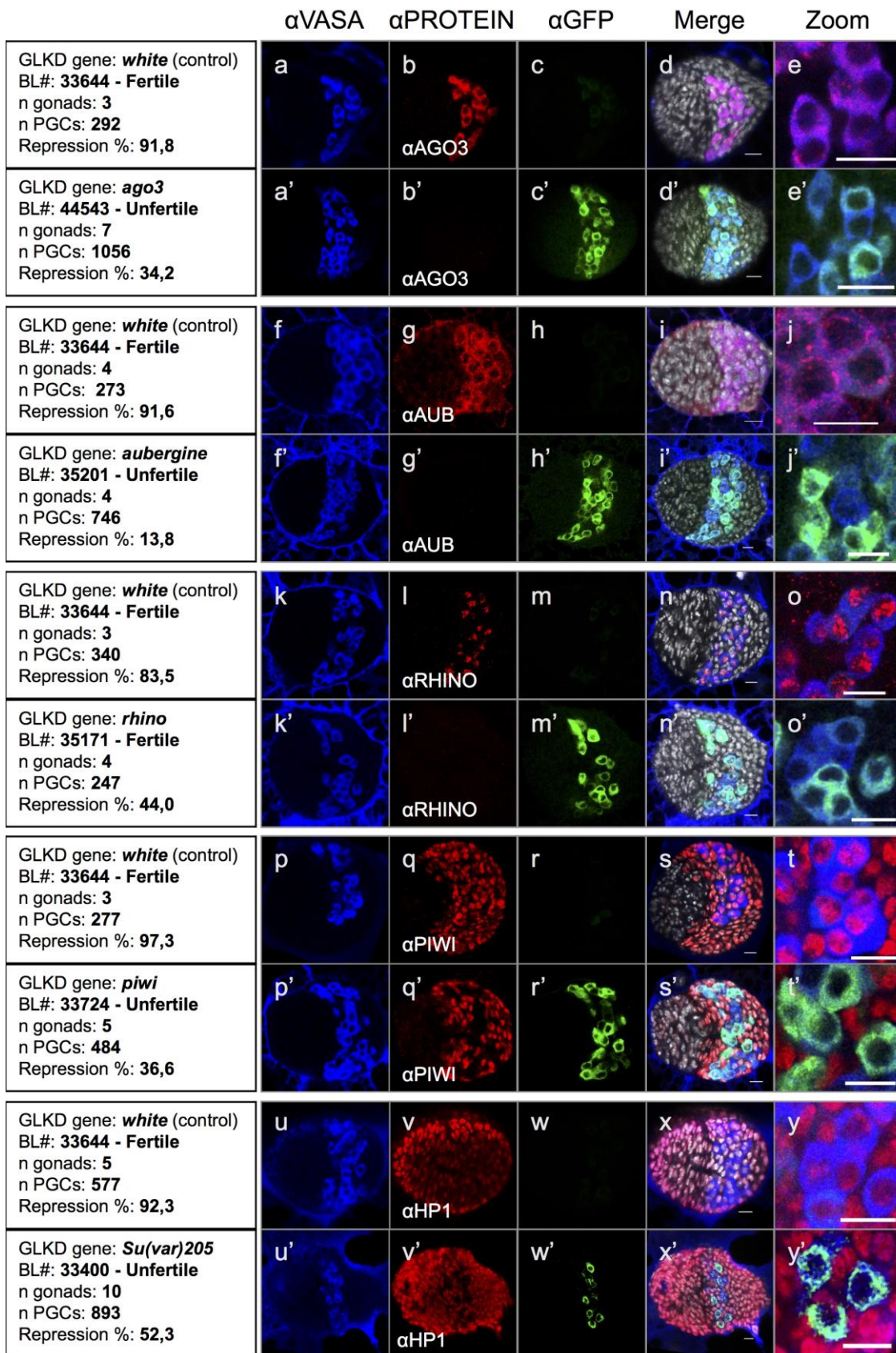
We show here that piRNA-mediated repression takes place in female gonads of *D. melanogaster* from late embryo to pupal stage. We present data demonstrating that *rhino*-dependent piRNAs are present in third instar larvae gonads and that HP1a is required for the production of telomeric piRNAs, specifically. Finally, incomplete repression (variegation) reveals an epigenetic cellular memory maintained from embryonic PGCs through ovaries.

**Figure 1.**



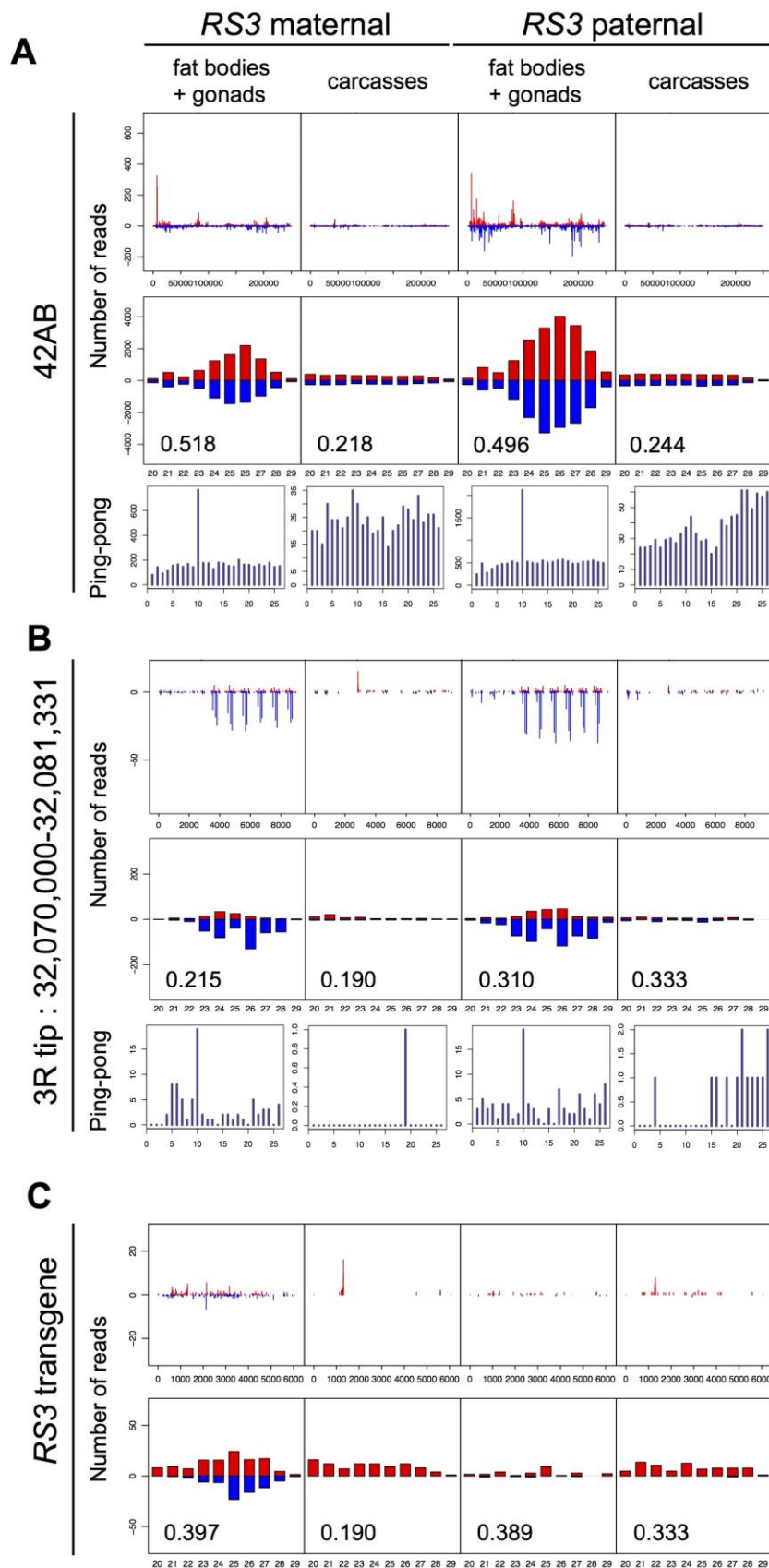


**Figure 2.**

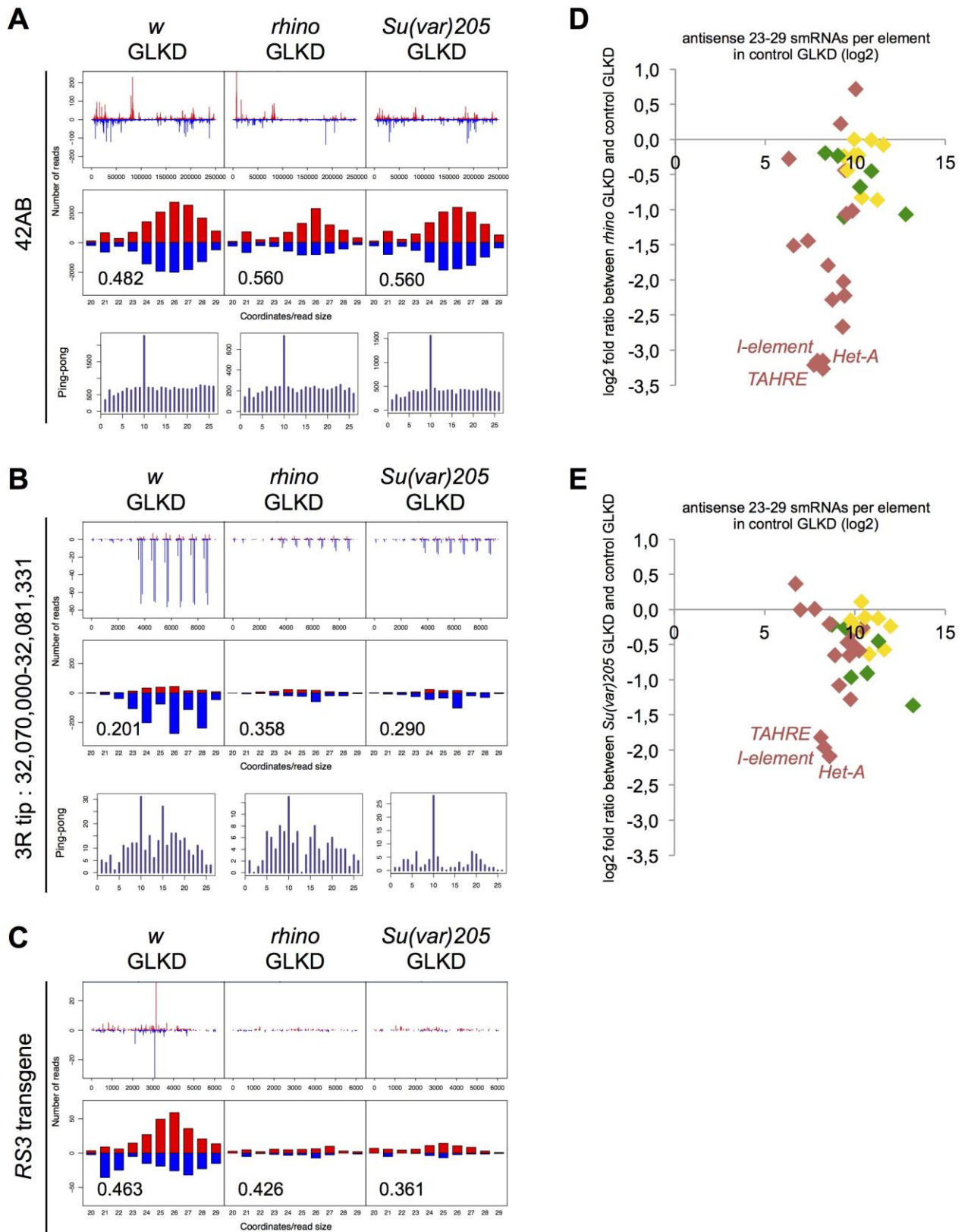




**Figure 3.**



**Figure 4.**



**Figure 5.**

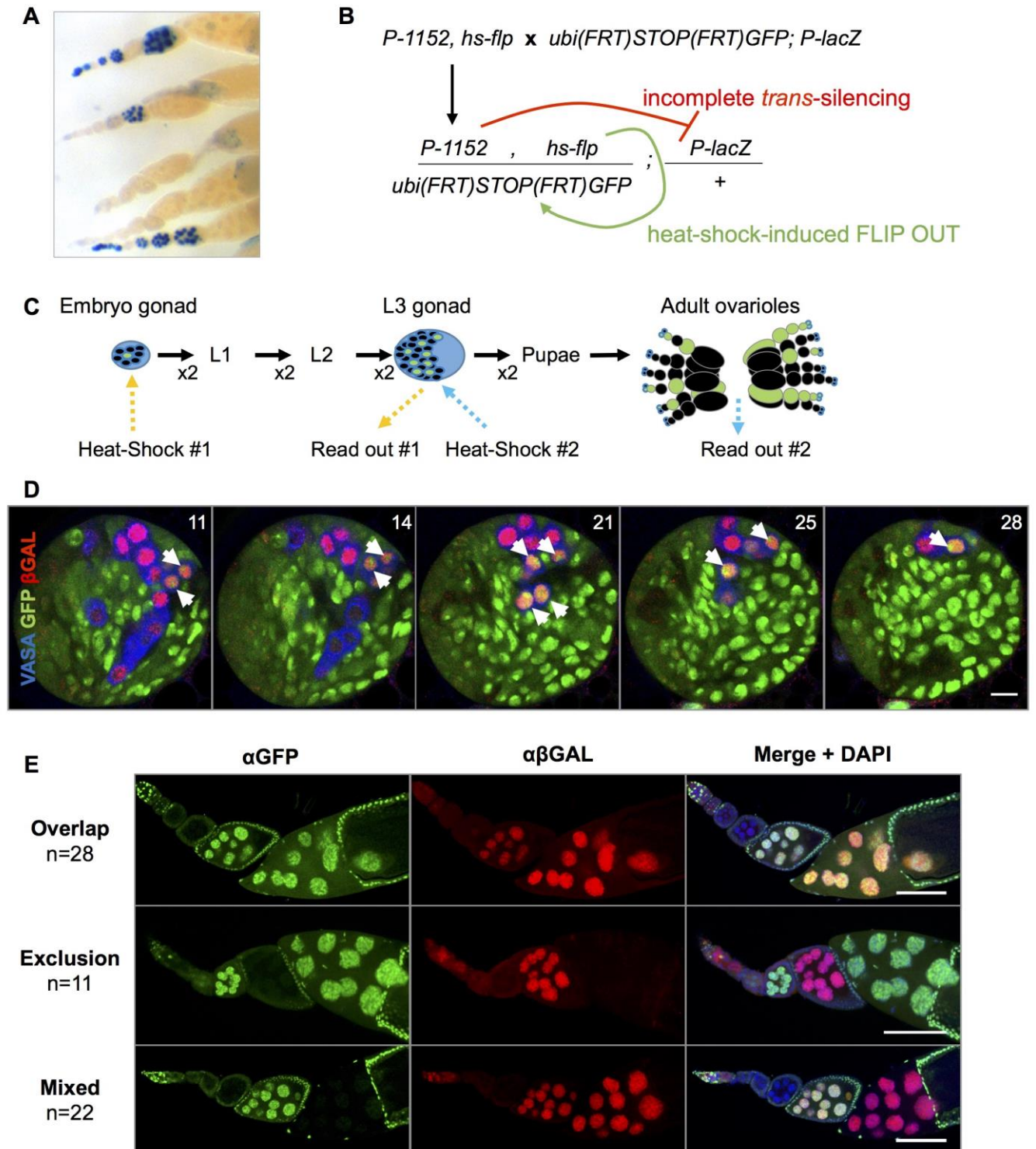


Figure S1.

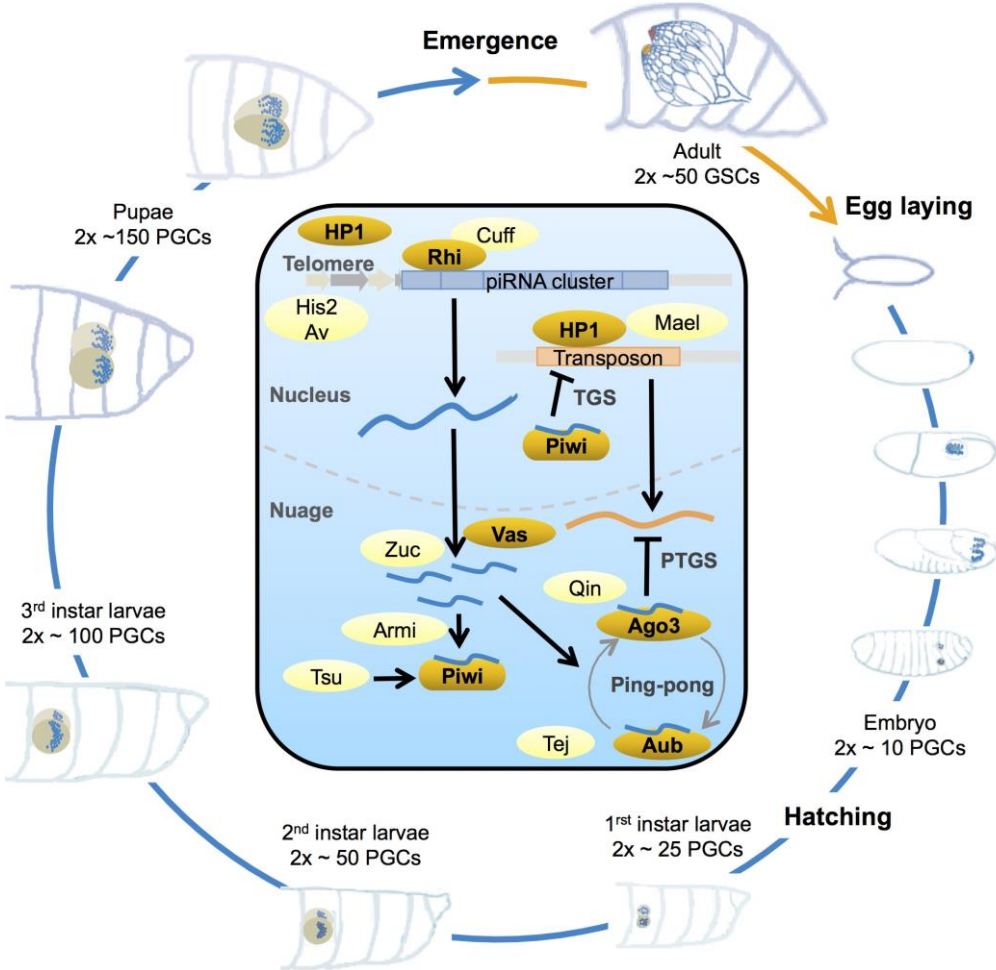
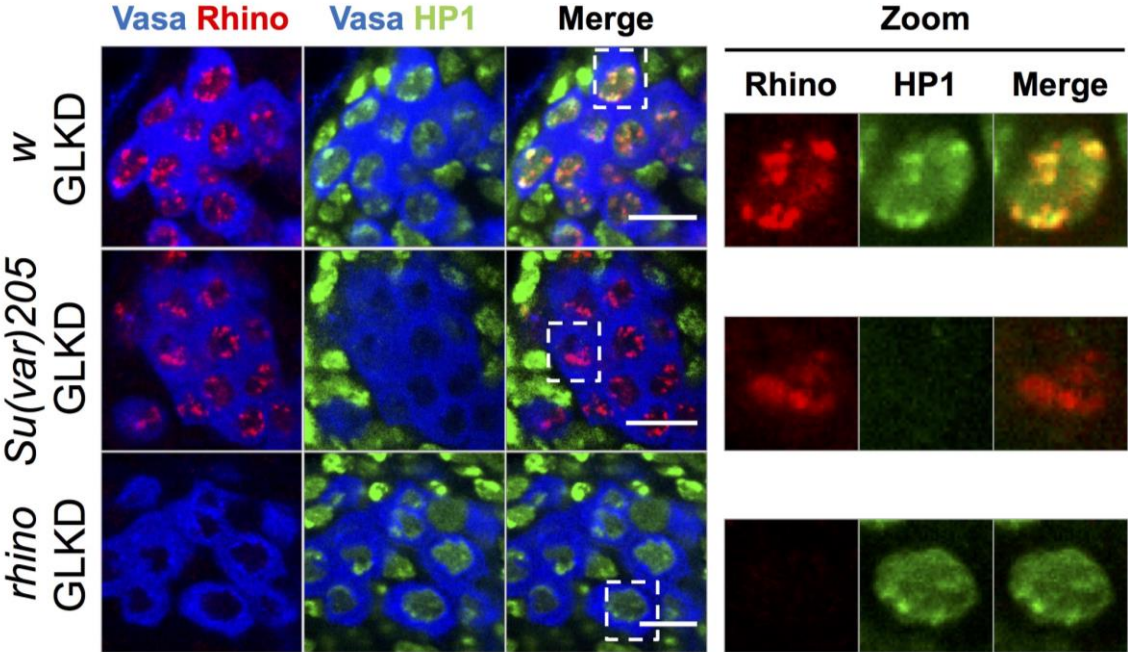




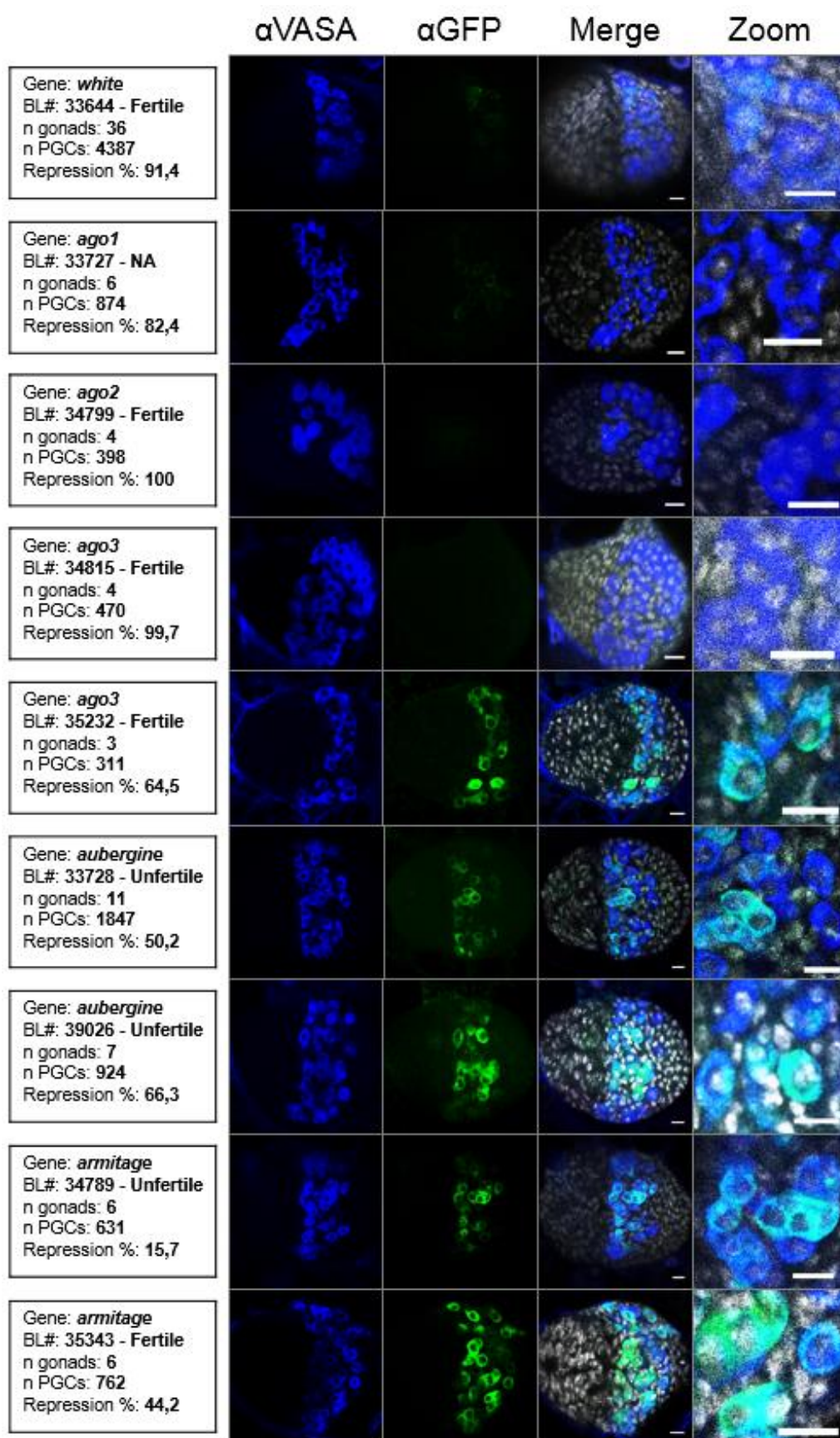
Figure S2.



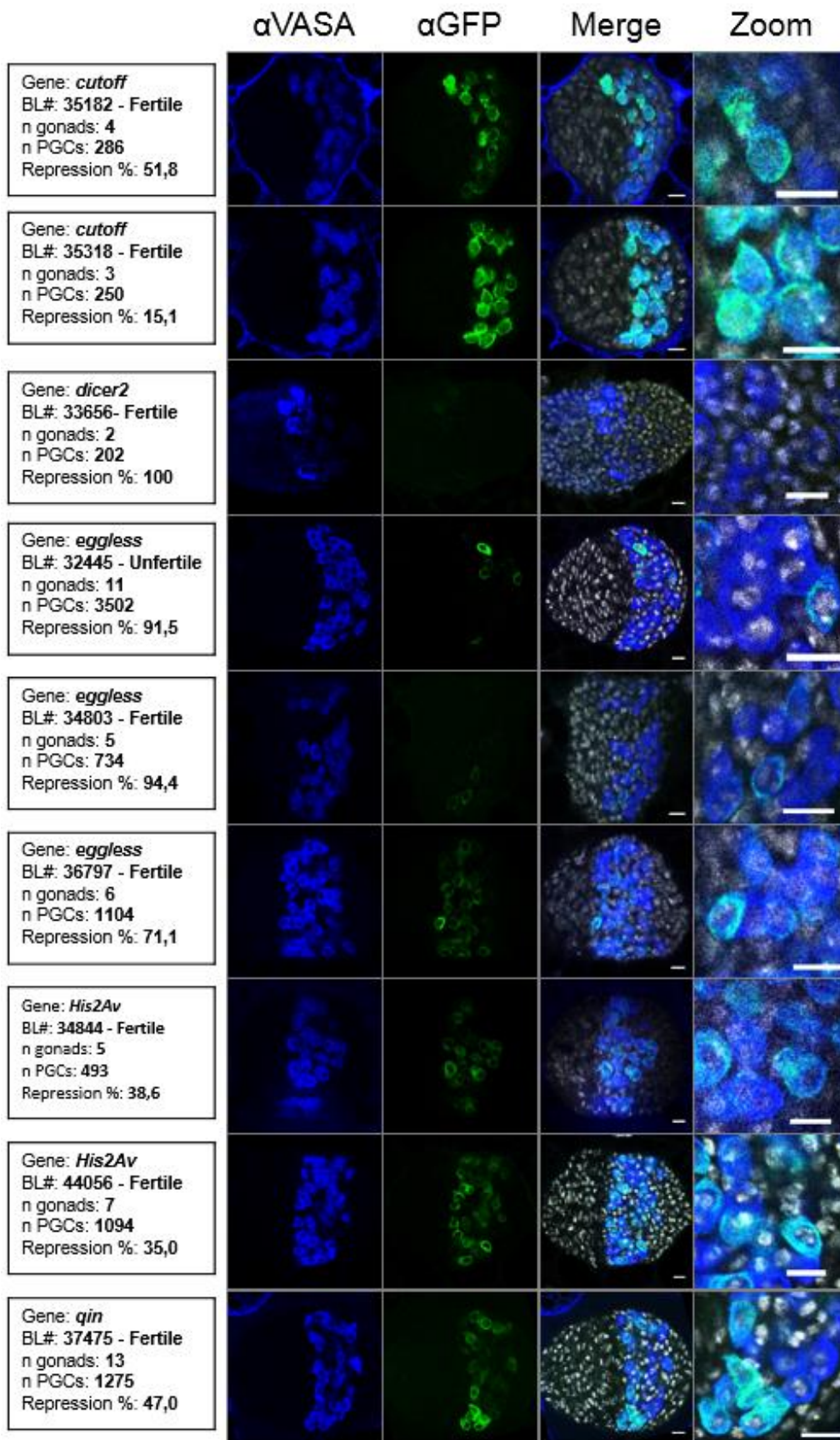
## Résultats additionnels

### Effet de la mutation de gènes impliqués dans la voie des piARN sur la répression au cours du développement.

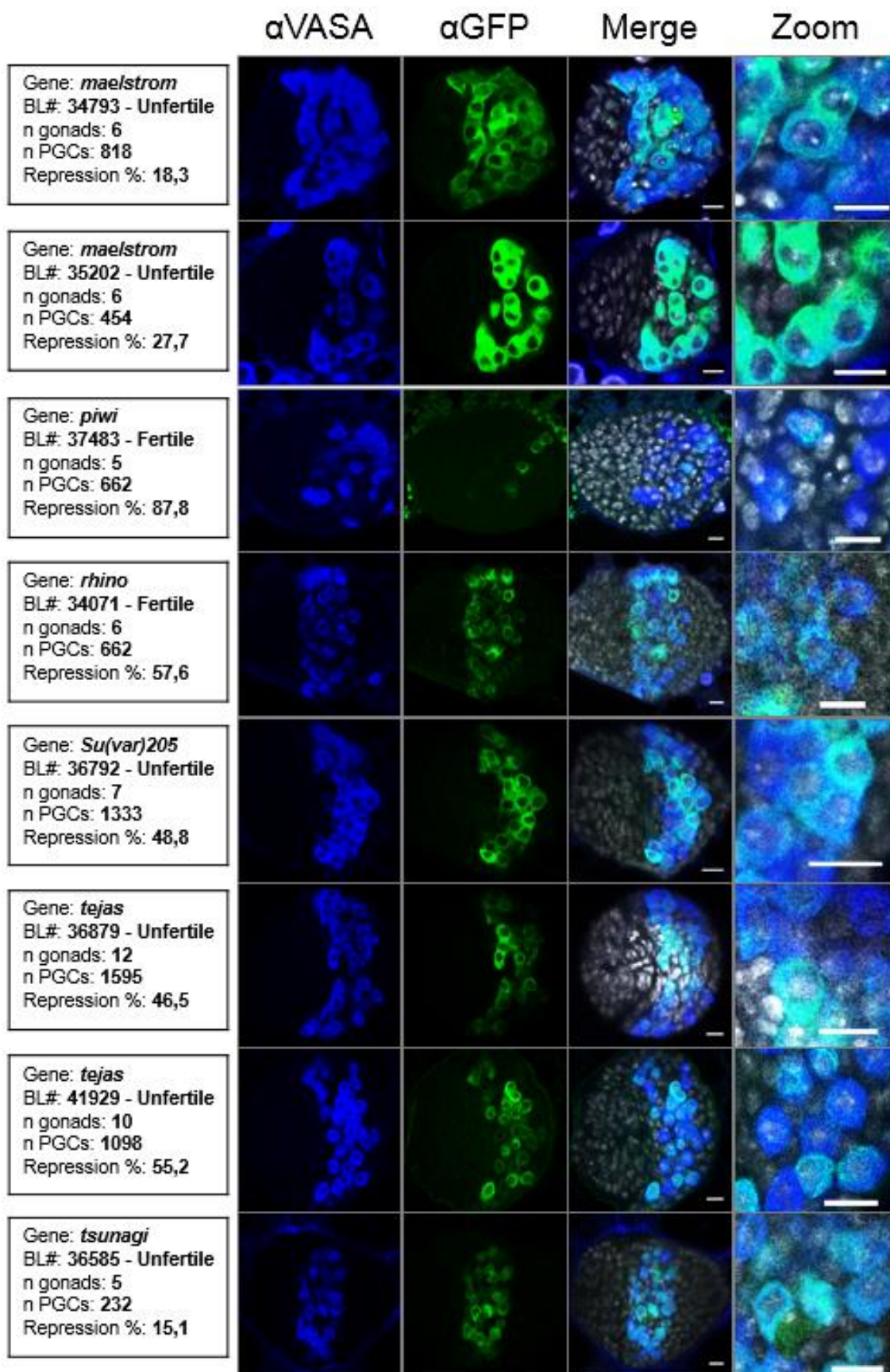
Durant ma thèse, j'ai testé l'impact de la mutation de 15 gènes appartenant à la voie des piARN sur la répression induite par le transgène télomérique *RS3*. Des femelles portant le transgène silencer *RS3*, un pilote *nosGAL4<sup>32180</sup>* et une cible *UASpGFP<sup>7373</sup>* ont été croisées par des mâles « TRiP line ». Les « TRiP line » possèdent des transgènes shARN basés sur le miARN miR-1 (Ni et al. 2011). Ils sont capables de produire des miARN interférents contre n'importe quelle séquence en fonction des brins guides présents dans l'insertion shARN. Il existe une banque regroupant les lignées TRiP disponibles actuellement (<http://www.flyrnai.org/TRiP-ACC.html>). Les constructions shARN sont sous le contrôle de séquences régulatrices UAS, ce qui m'a permis d'induire la production de l'ARN interférant spécifiquement dans la lignée germinale grâce à l'utilisation du pilote *nosGal4*, j'appellerai ces croisements « GermLine Knock Down » (GLKD) dans la suite de ce manuscrit (Fig. 40). J'ai donc pu également conclure que l'héritage maternel de ces protéines n'est pas suffisant pour maintenir la production de piARN au stade L3. En regardant l'effet de ces GLKD au stade L3, j'ai pu ainsi observer que 15 des 16 protéines nécessaires au silencing médié par les piARN au stade adulte que j'ai testé sont également indispensables au cours du développement. L'illustration des résultats regroupés dans la table 1 de l'article est présentée dans la figure suivante.

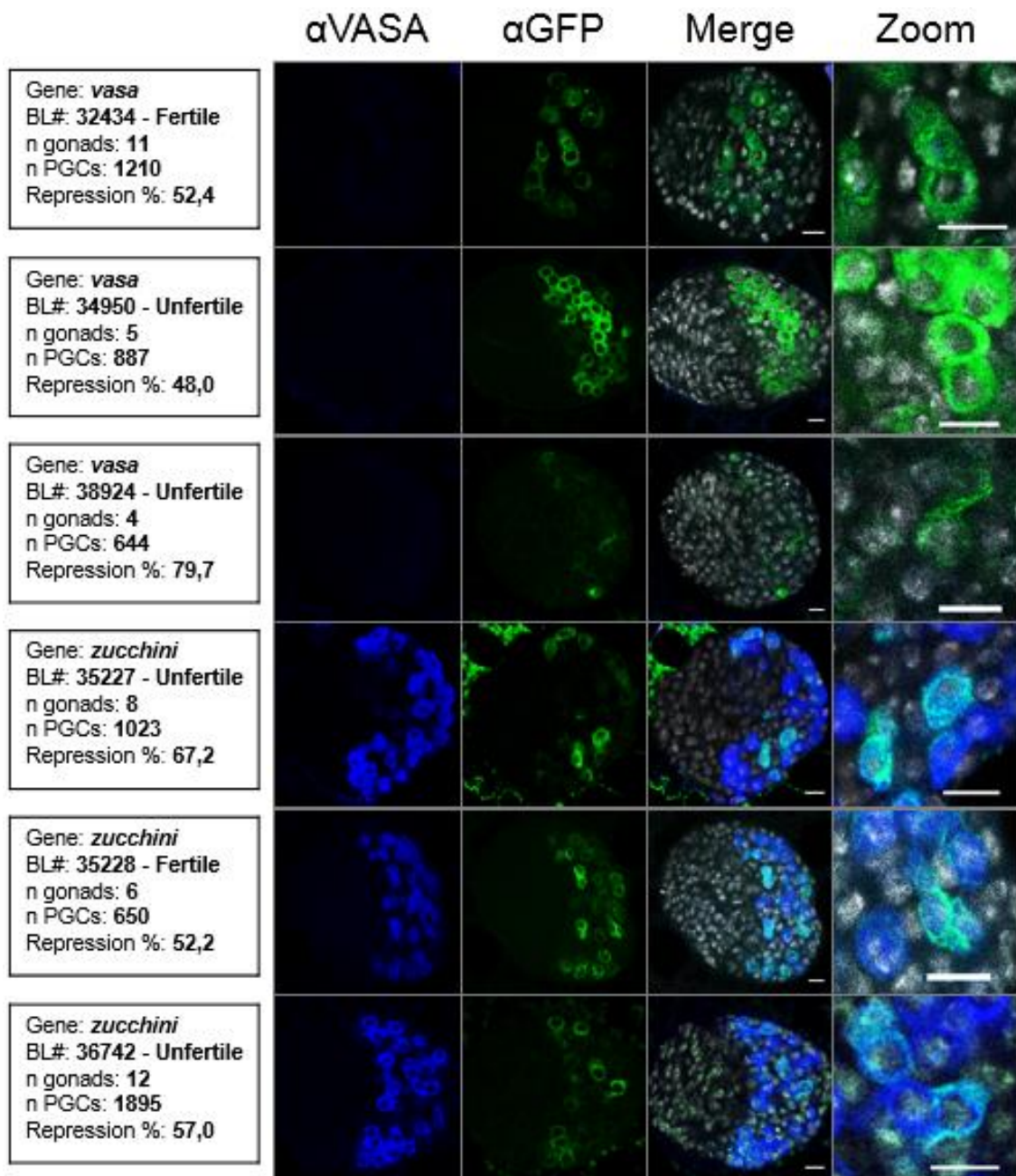












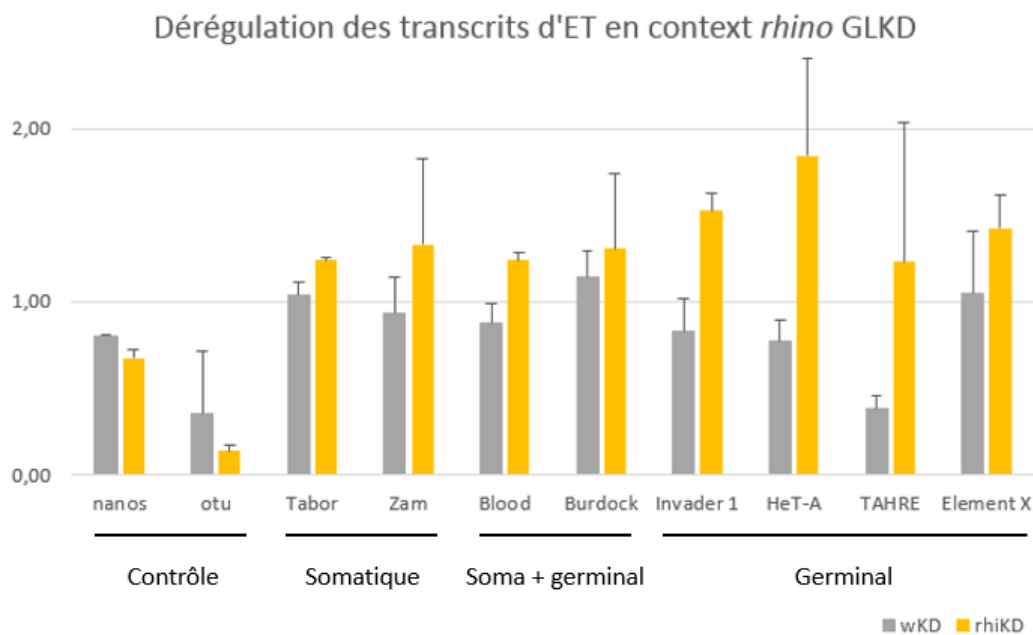
**Figure 40 - Illustration des résultats du crible gène candidat sur l'impact de la mutation des protéines impliquées dans la voie des piARN sur la répression au cours du développement.** 33 lignées ciblant 19 gènes sont présentées. La première colonne indique le gène ciblé par l'ARN interférent, la lignée shARN testée, la fertilité des descendants obtenus, le nombre de gonades et de cellules germinales comptées et la quantification de la répression. La seconde colonne montre le marquage de la protéine Vasa qui marque les cellules germinales (en bleu), la troisième colonne montre l'expression de la cible GFP (en vert) et les quatrième et dernière colonnes montrent la combinaison de ces deux marquages accompagnée du marquage DAPI qui colore l'ADN (en gris).

Les gènes sont classés par ordre alphabétique, on notera que la mutation des gènes AGO1, AGO2 et Dicer-2 qui ne sont pas impliqués dans la voie des piARN n'induit pas de dérégulation de la cible. En revanche, la mutation des gènes *aubergine*, *ago3*, *aubergine*,

*piwi*, *rhino*, *Su(var)205*, *armitage*, *cutoff*, *H2Av*, *qin*, *maelstrom*, *tejas*, *tsunagi*, *vasa* et *zucchini* induit une baisse de la répression au stade larvaire. Les résultats pour le gène *eggless* ne sont pas conclusifs car sur trois lignées RNAi pour ce gène, on obtient une descendance stérile dans une lignée sur trois (qui ne montre pas de dérégulation de la GFP) et de la dérégulation de la GFP dans un croisement sur trois. On ne peut donc pas conclure sur l'efficacité des ARN interférents sur l'expression du gène *eggless* ni sur son implication dans la répression par les piARN.

## Augmentation de la quantité d'ARN de différents éléments transposables en contexte mutant pour Rhino

La Figure 4 de l'article montre la quantité de piARN correspondants à un certain nombre d'éléments transposables germinaux et somatiques dans les contextes *rhino*<sup>GLKD</sup> et *Su(var)205*<sup>GLKD</sup>. Dans le cas du mutant Rhino, une diminution des piARN correspondants aux éléments transposables germinaux est constatée alors que la quantité de piARN correspondants aux éléments transposables somatiques ne varie pas significativement. Ce résultat est cohérent avec le fait que le RNAi contre *rhino* est cantonnée aux cellules germinales (croisement: femelles *nosGal4* par mâles *UASshARN rhino*). J'ai voulu confirmer ces résultats en visualisant si la baisse d'expression de ces piARN correspondants aux éléments transposables est corrélée avec leur dérégulation (plus de transcription de ces ET en contexte *rhino*<sup>GLKD</sup>). J'ai donc effectué des expériences de RTqPCR, sur des ARN extraits de corps gras disséqués et contenant les gonades de larves de troisième stade (Fig. 41).



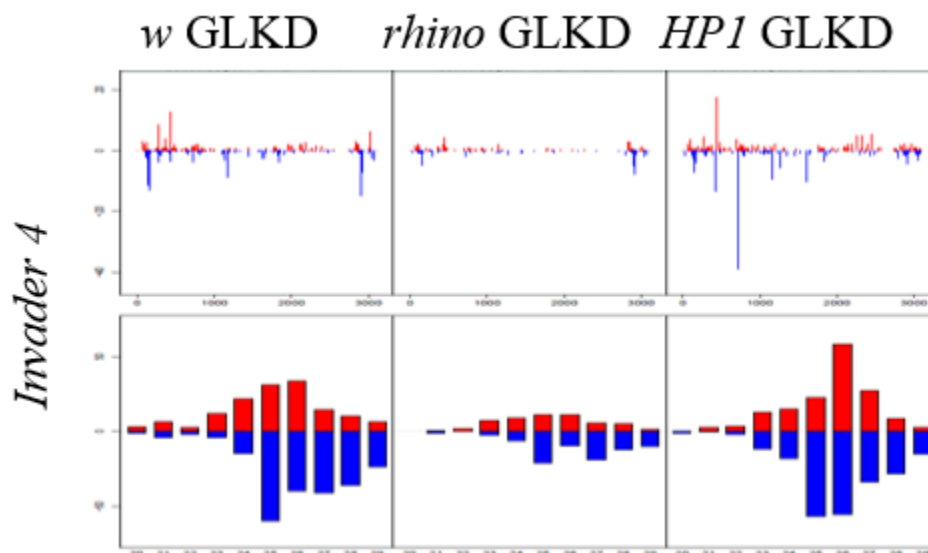
**Figure 41 – RTqPCR sur des ARN de corps gras et de gonades de larves de troisième stade.** Visualisation de l'expression de différents éléments transposables en contexte *rhino*<sup>GLKD</sup> (jaune) comparé au contrôle *w*<sup>GLKD</sup> (gris). De gauche à droite, les gènes utilisés comme contrôle, des ET connus pour être exprimés dans les tissus somatiques, des ET exprimés dans les tissus germinaux et somatiques et des ET spécifiques de la lignée germinale. Ces catégories ont été établies par la comparaison entre des piARN provenant d'ovaires de drosophile et de jeunes embryons (Malone et al. 2009). Les piARN présents dans les ovaires uniquement ne sont pas transmis à la descendance et doivent donc être présents uniquement dans les cellules folliculaires somatiques de la gonade femelles. Normalisation sur le gène *Rp49*.



Les cellules d'intérêt exprimant les ARN testés sont très peu nombreuses dans le tissu disséqué, environ 200 cellules par gonade. Les résultats que j'ai obtenus ne permettent pas de conclure quant à la dérégulation spécifique des éléments transposables germinaux dans *rhino<sup>GLKD</sup>*. On peut tout au plus déceler une tendance dans ce sens. Le tri des cellules germinales larvaires permettrait d'améliorer ces résultats et d'observer si les piARN produits durant le développement régulent effectivement les transcrits d'éléments transposables.

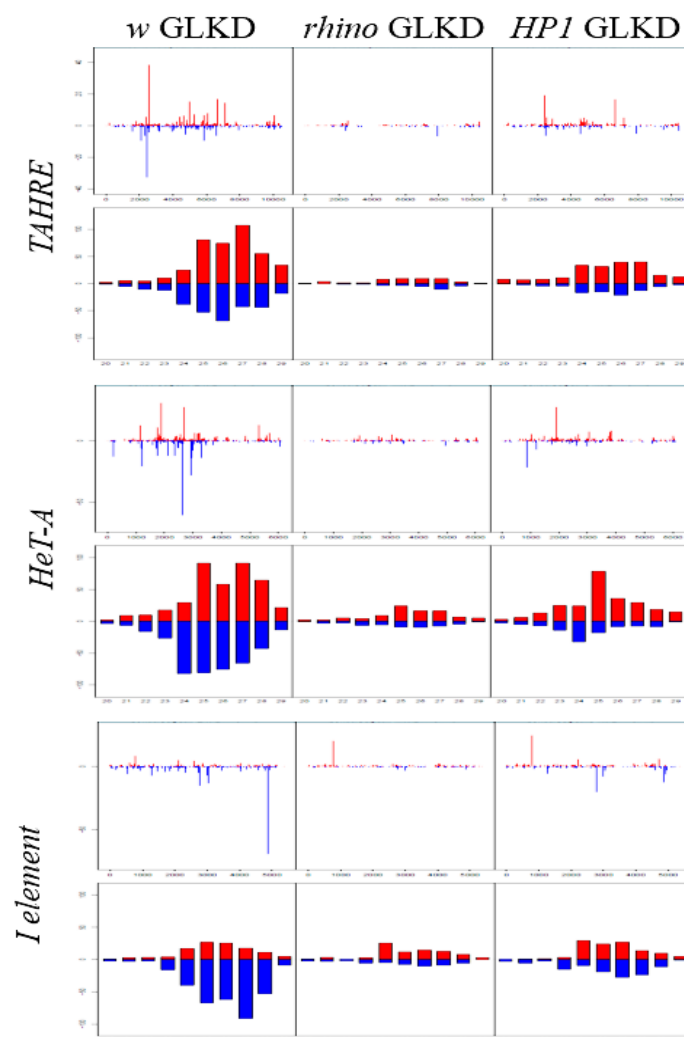
### Impact de la mutation germinale de Rhino et de HP1a sur les piARN de l'élément I, Invader 4, Het-A et TAHRE

Les analyses de séquençage haut débit des petits ARN provenant des descendants *rhino<sup>GLKD</sup>* et *Su(var)205<sup>GLKD</sup>* permettent néanmoins d'observer l'impact de la perte de ces protéines sur la quantité de piARN correspondant aux transposons. De façon étonnante, la quantité de piARN de l'élément transposable *Invader 4* ne semble pas diminuer en condition mutante pour HP1a (Fig. 42). Pourtant la production de piARN par les TAS baisse suite à la perte de HP1a (Fig. 4B Article) or les séquences TAS contiennent des répétitions d'*Invader 4*. Ce sont ces répétitions qui produisent les pics de piARN antisens que l'on voit sur l'encart correspondant au cluster *3R TIP* dans les Figures 3 et 4 de l'article. Ici, les piARN correspondant à *Invader 4* sont alignés sur la séquence complète de l'élément.



**Figure 42 – Comparaison des piARN totaux correspondants à *Invader 4* provenant de gonades de L3 *rhino<sup>GLKD</sup>* et *Su(var)205<sup>GLKD</sup>*.** Sur la première ligne les piARN totaux d'*Invader 4* sont alignés en fonction de leur position sur la séquence. Sur la seconde ligne, les piARN sont distribués en fonction de leur taille et de leur quantité. Rouge, piARN correspondants au sens de transcription de l'élément transposable. Bleu, piARN dont la séquence est complémentaire à l'élément transposable (antisens).

L'effet que nous voyons sur le cluster *3R TIP* n'est pas visible à cette échelle. En effet, le *3R TIP* ne contient qu'une petite partie d'*Invader 4* (un fragment des séquences LTR), les piARN observés sur la Fig. 42 proviennent d'autres sources comme le cluster *42AB*. Dans le contexte *Su(var)205<sup>GLKD</sup>*, seuls les piARN des ET *élément I*, *Het-A* et *TAHRE* diminuent par rapport au contrôle et dans une proportion moins dramatique que pour le *rhino<sup>GLKD</sup>* (Article Fig. 4D-E). L'observation des piARN extraits des 3 croisements portant les shARN (*w<sup>GLKD</sup>*, *rhino<sup>GLKD</sup>* et *Su(var)205<sup>GLKD</sup>*) sur les séquences correspondantes à ces éléments transposables montre également une diminution dans le cas des mutants Rhino et HP1a (Fig. 43).



**Figure 43 - Comparaison des petits ARN de 20 à 29 nucléotides correspondant aux ET *TAHRE*, *Het-A* et *élément I* extraits de gonades de L3 *w<sup>GLKD</sup>*, *rhino<sup>GLKD</sup>* et *Su(var)205<sup>GLKD</sup>*. Sur la première ligne de chaque encart, les petits ARN (20-29nt) correspondants aux trois transposons sont alignés en fonction de leur position sur la séquence de l'élément transposable. Sur la seconde ligne, distribution des piARN en fonction de leur taille et de leur quantité. En rouge, piARN sens, en bleu, piARN antisens.**



## Conclusions

L'utilisation de transgènes insérés dans les clusters producteurs de piARN nous a permis de visualiser la répression médiée par les piARN tout au long du développement de la lignée germinale femelle, des PGC de l'embryon à la puppe. Les gènes typiques de la voie des piARN sont nécessaires pour la répression du transgène rapporteur au stade L3 alors que les gènes impliqués dans la voie des si/miARN ne le sont pas. En outre, nous avons pu montrer la présence de petits ARN correspondant à des piARN dans les gonades de larves. Dans la dernière partie de cet article, nous avons effectué des analyses clonales révélant que le silencing incomplet nommé variégation est probablement soumis à un verrouillage épigénétique.

Le crible shARN a permis l'expression d'ARN interférants spécifiquement dans la lignée germinale et actifs uniquement à partir de la transcription zygotique (GLKD) montre que ces piARN sont fonctionnels (réapparition de la cible GFP). Les analyses des petits ARN pour les contextes *Rhino*<sup>GLKD</sup> et d'*HP1*<sup>GLKD</sup> suggère fortement qu'une partie des piARN détectés dans les gonades de L3 sont produits *de novo* par le zygote. En revanche, les piARN insensibles à l'absence de Rhino sont probablement d'origine maternelle ou somatique (on sait qu'il y a des piARN somatiques fonctionnels dans les corps gras chez l'adulte : Helfand, Oral communication TAGC 2016). Un point intéressant à noter est l'impact de l'absence de HP1a, codée par *Su(var)205*, sur la répression médiée par les piARN. On aurait pu penser que le rôle de HP1a est surtout lié à l'hétérochromatinisation des séquences d'ET euchromatiques mais il semblerait que cette protéine soit également nécessaire pour produire spécifiquement les piARN télomériques. Les expériences de A-L. Todeschini en 2010 avaient déjà montrées une baisse générale de la quantité de piARN dans les ovaires d'un adulte hétérozygote *Su(var)205/+* (Todeschini et al. 2010). Dans cet article, nous notons une baisse de quantité des piARN correspondants spécifiquement au *3R TIP* et au transgène silencer *RS3* en contexte *Su(var)205*<sup>GLKD</sup> alors que la quantité de piARN correspondant au cluster *42AB* reste inchangée. HP1a participerait donc à l'activation de la transcription des clusters de piARN télomériques.





## Persistance des piARN maternels

### 1. Stabilité des piARN hérités maternellement

Une question inhérente à la problématique des piARN au cours du développement est la stabilité des piARN hérités maternellement. J'ai montré que l'héritage maternel n'est pas suffisant pour maintenir une répression complète au stade L3 avec le contexte *rhino*<sup>GLKD</sup> et les autres gènes testés dans le crible shARN mais qu'en est-il dans les stades plus précoces ? En effet, les piARN déposés dans l'embryon par la mère sont une composante essentielle de la répression. Ils permettent vraisemblablement d'initier le mécanisme puisque sans eux (quand le transgène télomérique est apporté par le père) la répression met plusieurs générations à se remettre en place (Josse et al. 2007) (Laidin et al, en préparation). On sait qu'une grande partie des ARNm maternels déposés dans l'œuf sont détruits quelques heures après la ponte, au moment où le génome zygotique prend le relais. C'est la transition MTZ, au cours de laquelle se produit une déadénylation des ARNm médiée par d'énormes complexes protéiques (Simonelig 2012). Mais qu'en est-il des piARN ? La différence de taille entre un ARNm et les petits ARN interférants et l'absence de queue PolyA permet de penser que l'éventuelle dégradation des piARN maternels doit faire intervenir un mécanisme différent.

- Les piARN sont-ils aussi l'objet de la dégradation sélective des ARN hérités maternellement ?
- Servent-ils uniquement à initier la production de piARN zygotiques *via* l'activation des clusters producteurs ou permettent-ils aussi d'éteindre l'expression des éléments transposables (PTGS) durant la vie embryonnaire ?
- Les piARN maternels sont-ils déposés uniquement au pôle postérieur de l'embryon ? Dans le cas d'une distribution homogène dans l'œuf, peuvent-ils avoir un effet somatique dans les premiers temps du développement ?

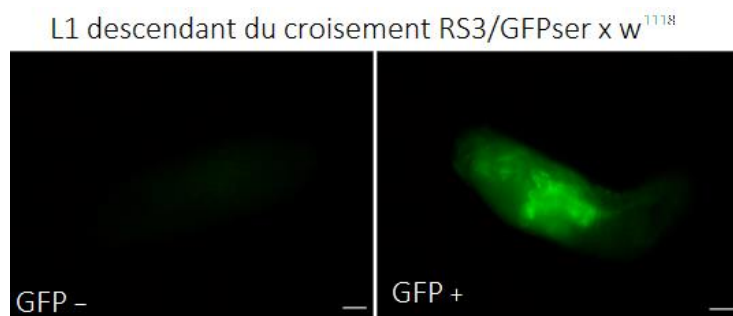
### 2. Effet et demi-vie des piARN maternels

La première étape pour déterminer l'effet des piARN maternellement hérités est de pouvoir comparer des conditions où la transmission du *locus* producteur de piARN et les piARN maternels sont maîtrisées séparément, et ce, dans un fond génétique aussi proche que possible. Nous avons donc utilisé des mères hétérozygotes pour le transgène silencer *RS3*

pour observer si la partie de la descendance qui n'a pas reçu le silencer est capable de réprimer une cible. Et si oui, pendant combien de temps?

### 2.1 Effet des piARN maternels sur une cible germinale

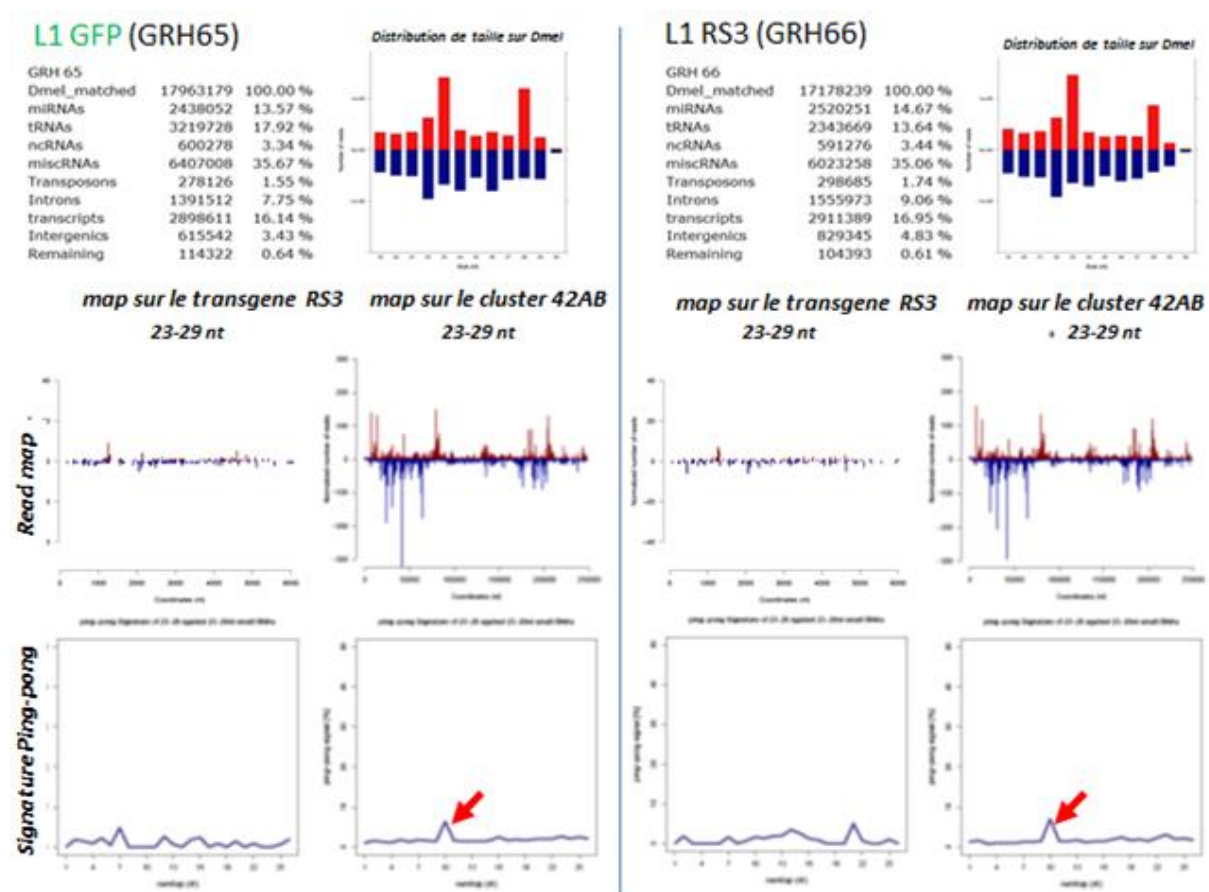
Dans le but de comparer la descendance ayant hérité du silencer télomérique ou seulement des piARN maternels, nous avons construit des souches possédant un transgène silencer sur un balancier fluorescent (chromosome remanié empêchant la recombinaison et portant un marqueur phénotypique dominant). Des femelles *RS3/TM3,actinGFP,Serrate* ont été croisées par des mâles *w<sup>1118</sup>*. Leur descendance devrait être pour moitié composée d'embryons [GFP+] ayant hérité seulement des piARN maternels sans le *locus* chromosomique producteur *RS3* et d'embryons [GFP-] possédant le *locus* silencer (et des piARN maternels). Nous pensions pouvoir discriminer les individus sur leur phénotype GFP, malheureusement, l'héritage maternel de protéine GFP et l'autofluorescence empêchent de différencier les embryons. Nous avons pu séparer les deux populations au stade L1 (Fig. 44).



**Figure 44 – Larves de premier stade issues du croisement RS3/GFPser x *w<sup>1118</sup>*.**

GFP- : RS3/+ et GFP+ : GFPser/+  
Échelle 100µm.

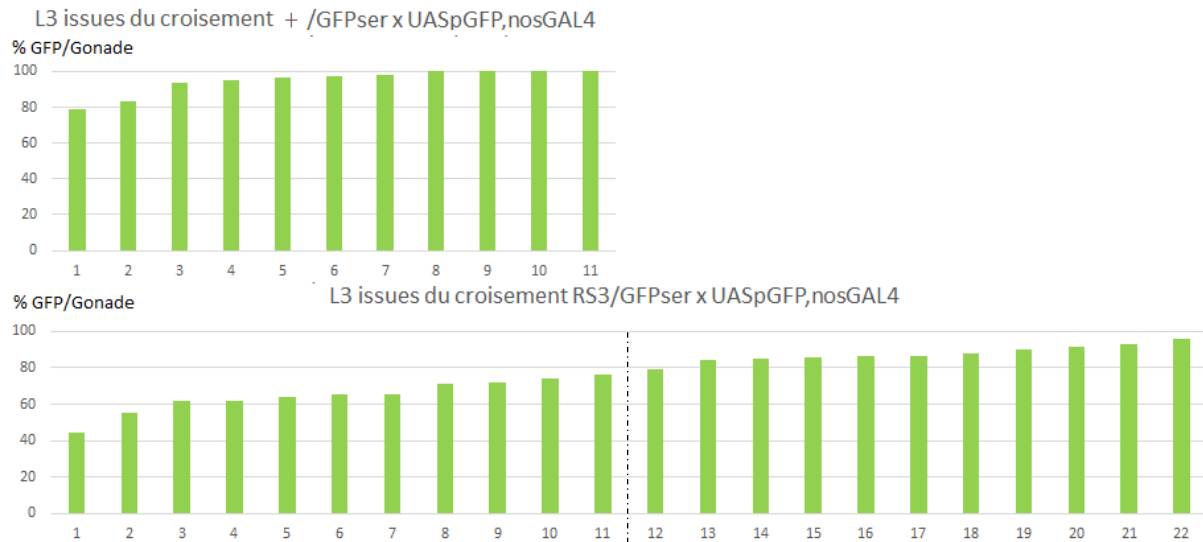
Les individus ayant hérité ou non du balancier GFP sont facilement repérables. Nous avons ensuite voulu voir si des piARN sont présents dans les larves de premiers stades en l'absence du transgène silencer (Fig. 45). Les ARN totaux ont été extraits à partir des deux populations de L1 (GFP+ ou -) et les petits ARN ont été séquencés. Une centaine de descendants de chaque génotype a été utilisée.



**Figure 45 - Résultats du séquençage des petits ARN non-codants des issues de L1 du croisement *RS3/GFP*  $\times$  *w<sup>1118</sup>*.** Première ligne : distribution des petits ARN par catégories : micro ARN (miRNA), ARN de transfert (tRNA), ARN non codants (ncRNA), ARN non catégorisés (miscRNA), transposons, Introns, ARNm (Transcripts), ARN intergéniques. Distribution des petits ARN correspondants au génome de *D. melanogaster* par taille. Seconde ligne, position des petits ARN de 23-29nt sur la séquence du transgène *RS3* et sur la séquence du cluster *42AB*. Troisième ligne, signature Ping-pong (nombre de paires de piARN ayant une séquence homologue de 10nt) correspondante aux séquences de la deuxième ligne.

On détecte effectivement des piARN correspondants au cluster *42AB* (le plus productif de *D. melanogaster*) au stade L1. Ces piARN ont une petite signature ping-pong. En revanche les piARN correspondants au locus *RS3* sont trop peu nombreux pour distinguer une différence entre les deux types de descendants. Soit les piARN maternels sont toujours présents et le relais zygotique n'a pas encore commencé dans les cellules germinales des descendants *RS3*, soit la quantité de piARN séquencés est trop faible pour voir une différence. Le fait de ne pas voir de signature ping-pong penche dans le sens de la seconde hypothèse, le nombre peu élevé de cellules germinales (une dizaine par gonade) comparativement au nombre de cellules somatiques de la larve également.

J'ai répété cette expérience à des stades plus tardifs où la plus grande taille des gonades m'a permis de les disséquer. J'ai quantifié l'expression d'une cible GFP dans les descendants du croisement *RS3/TM3;actineGFP;Serrate* x *UASpGFP<sup>7373</sup>,nosGal4<sup>32180</sup>* par immunomarquage et je l'ai comparée à un croisement contrôle sans transgène silencier.



**Figure 46 – Expression d'une cible GFP** dans les descendants des croisements *RS3/TM3;actineGFP;Serrate* x *UASpGFP<sup>7373</sup>,nosGal4<sup>32180</sup>* et *+TM3;actineGFP;Serrate* x *UASpGFP<sup>7373</sup>,nosGal4<sup>32180</sup>*. Chaque barre représente le pourcentage de PGC exprimant la cible GFP dans une gonade. La GFP est exprimée dans 96% des PGC du contrôle et dans 76% des PGC des descendants du croisement contenant le silencier *RS3*.

Encore une fois, l'expression du balancier GFP n'est pas discriminable de l'autofluorescence dans les descendants du croisement *RS3/TM3;actineGFP;Serrate* x *UASpGFP<sup>7373</sup>,nosGal4<sup>32180</sup>*. Il n'a donc pas été possible de choisir les individus possédants, ou non, le *locus* silencier. On ne voit pas non plus deux populations de gonades ayant une différence marquée de répression de la cible GFP. Les données de séquençage haut débit sur les ARN de L3 ayant hérité du transgène silencier maternellement et paternellement (Fig. 2 de l'article) montrent pourtant une différence nette de quantité des piARN correspondants à *RS3* mais elles proviennent de mouches dont la mère était homozygote pour le silencier *RS3*. La transmission de piARN est probablement moins importante dans le cas de mères hétérozygotes. Je l'ai constaté moi-même lors de la création de la souche utilisée pour le crible shARN : les larves descendantes de mères *RS3,UASpGFP<sup>7373</sup>,nosGal4<sup>32180</sup> /+* ne montraient pas de répression de la GFP dans toutes les PGC en L3. Il a fallu plusieurs générations homozygotes pour établir la répression dans 91% des PGC. En revanche, une seule génération sans une protéine de la voie des piARN suffit à briser la répression (voir Table 1 Article).

On remarque que la moitié des gonades présente plus de PGC réprimées que les gonades contrôles (<80% de PGC exprimant la GFP). On observe jusqu'à 60% de répression du transgène UASpGFP<sup>7373</sup> dans la plus réprimée. Il pourrait donc y avoir deux populations: une similaire au contrôle >80% d'expression (gonade 12 à 22) et une population montrant de la répression de PGC (<80% d'expression gonades 1 à 11). Ce qui correspondrait à une ségrégation 50/50 : les gonades montrant moins de 80% de PGC exprimant la GFP pourraient être celles qui ont hérité du *locus* silencer *RS3*. Pour le moment, si on sépare les gonades en deux groupes (1-11) et (12-22), on peut penser que celles ayant hérité des piARN maternels seuls (12-22) se comportent comme le contrôle et ne répriment pas la GFP, tandis que celles ayant reçu le transgène *RS3* (1-11) répriment mieux la GFP. Les piARN maternels semblent donc ne plus être actifs au stade L3. Les résultats de la partie suivante paraissent corroborer cette conclusion.

Afin d'améliorer cette analyse, une souche *BX2\*/Balanceur RFP* est en cours de construction : le cluster silencer *BX2\** est connu pour être efficace à l'état hétérozygote (Chap. 2 B.1) et le nouveau balanceur RFP semble plus visible. On pourra ainsi discriminer plus aisément les individus RFP+ des RFP- car les larves (en particulier les corps gras) autofluorescent moins en rouge qu'en vert. Ceci facilitera la séparation des individus ayant hérité du transgène silencer de ceux qui n'auront eu que les piARN maternels.

## **2.2 Effet des piARN maternels sur une cible somatique**











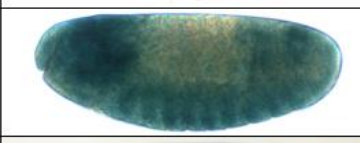
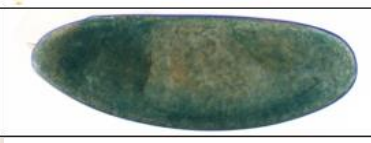



Pour répondre à la question de la distribution des piARN maternels dans l'œuf et de leur éventuel effet somatique pendant le développement, nous avons utilisé la lignée *12179* [*y<sup>1</sup> w<sup>67c23</sup>; P{lacW}mRpL4<sup>k14608</sup>/CyO*] identifiée dans le crible effectué durant mon master à la recherche d'une lignée exprimant un transgène *placZ* précocement (Première partie des résultats). Cette insertion permet l'expression somatique de la  $\beta$ galactosidase relativement tôt au cours du développement embryonnaire (environ 3h AEL) grâce à la présence d'un transgène *placZ* dans le gène *RPL4*, elle est létale homozygote. Elle sera appelée *placZ<sup>12179</sup>* dans la suite de ce chapitre. Nous avons voulu tester si des piARN déposés par la mère pouvaient avoir une influence sur l'expression de ce transgène. Malgré une homologie de séquence plus importante entre le silencer *P-1152* et l'insertion de *12179*, qui contiennent tous les deux *lacZ*, nous avons utilisé *RS3* comme lignée silencer. En effet, la lignée *P-1152* présente une expression somatique de la  $\beta$ galactosidase au cours de l'embryogenèse qui

disparaît par la suite. L'homologie entre le silencer et la cible se fait donc au niveau des séquences de l'élément *P*.

Nous avons croisé des femelles  $w^{1118}$  (contrôle) ou portant le transgène silencer *RS3*, avec des mâles *placZ*<sup>12179</sup>/*CyO*. Nous avons stadé précisément les embryons pour observer une éventuelle répression somatique lors des premières heures du développement. Karine Casier, alors étudiante en L3, s'est chargée de produire des embryons d'âges différents et de les colorer (coloration X-gal). Les résultats sont présentés dans la figure 47.

Le gène *lacZ* inclus dans le transgène s'exprime dans l'embryon à partir de 3-4h AEL lorsque qu'aucun piARN homologue à sa séquence n'est déposé dans l'œuf (Fig. 47, à gauche). De façon intéressante, un retard de l'expression somatique de *placZ*<sup>12179</sup> de 2 à 4 heures est observé dans la descendance des mères portant le silencer *RS3*. Il semblerait donc que les piARN maternels provenant de *RS3* soient capables de retarder l'expression somatique au début du développement et répriment donc un ARN cible. Ils sont probablement complexés aux protéines Piwi (ou Aubergine) lors que la mise en place du plasma germinale, ce qui les protégerait de la dégradation. Ils pourraient cibler les ARNm chimères de *placZ*<sup>12179</sup> via l'homologie de séquence avec *P*. La réexpression du gène rapporteur *placZ*<sup>12179</sup> à partir de 10h suggère que les piARN hérités maternellement sont bien capables d'assurer une répression fonctionnelle d'ARNm homologues pendant le début du développement et ce, dans tout l'embryon. Ensuite, on peut penser que comme le silencer *RS3* n'est pas transcrit dans les cellules somatiques, par conséquent incapable de produire de nouveaux piARN somatiques zygotiques, cet effet va être dilué. À mesure que l'embryon se développe, *placZ*<sup>12179</sup> sera exprimé dans les tissus somatiques. Le temps d'action des piARN maternels serait donc de 8-10h dans les tissus somatiques.

Ces résultats sont en accord avec les expériences récentes du labo de M. Simonelig qui montrent que les piARN liés à Aub interagissent avec des centaines d'ARNm dans les deux premières heures du développement (iCLIP sur embryons de 0-2h) (Barckmann et al. 2015). Ici, nous montrons donc que ces piARN peuvent effectivement retarder l'expression d'un ARNm pour plusieurs heures. Pour l'instant nous ne savons pas si l'ARNm du rapporteur *placZ*<sup>12179</sup> est séquestré ou détruit au début du développement.

Croisement Stade	♀ ♂ w <sup>1118</sup> x 12179/Cy	♀ ♂ RS3 x 12179/Cy
1-3 h stade 5		
3-5 h stade 8		
4-6 h stade 10		
6-8 h stade 12		
8-10 h	NA	NA
10-12 h		
12-14 h		
14-16 h stade 16		NA
17-19 h stade 17		

**Figure 47 - Effet de l'héritage de piARN sur une cible somatique au cours du développement.** Première colonne, cinétique d'expression du transgène *placZ*<sup>12179</sup> apporté paternellement. Seconde colonne, effet de la présence de piARN maternels homologues (P) sur l'expression du transgène *placZ*<sup>12179</sup> apporté paternellement.





## **Article 2 «Knock down of *half pint* gives birth to a meiotically stable piRNA-producing locus in *Drosophila* »**

### **Présentation de l'article**

Les piARN originaires de *loci* hétérochromatiques sont responsables de la répression des éléments transposables et de la bonne expression des gènes du développement. Chez *D. melanogaster*, la plupart des piARN correspondent à environ 140 *loci* génomiques hétérochromatiques, la plupart situés à proximité des centromères ou des télomères (Brennecke et al. 2007). En dépit de leur localisation, leur caractère répétitif et leur statut hétérochromatique, les clusters de piARN sont transcrits en longs précurseurs simple brins par l'ARN Pol II qui seront ensuite maturés en piARN (Mohn et al. 2014; L. Zhang et al. 2014). Les clusters de piARN dual-strands comme *42AB*, sont transcrits de façon non-canonique grâce au Rhino-Deadlock-Cutoff (RDC) complexe (Mohn et al. 2014; L. Zhang et al. 2014). Les conditions nécessaires à l'activation de leur transcription restent une question ouverte.

Dans cet article, l'utilisation d'un cluster correspondant à 7 copies en tandem d'un transgène *placW* (*BX2*) pouvant être activé pour la production de piARN grâce à l'héritage de piARN maternels (de Vanssay et al. 2012) a permis de cribler l'effet de la perte de fonction d'un certain nombre de gènes. L'utilisation de transgènes shARN inductibles dans la lignée germinale a mis en évidence que la mutation du gène *half pint* (*hfp*) démarre la production de piARN par le cluster *BX2* inactif. *hfp* code pour le facteur d'épissage Puf68 qui se lie à des séquences poly-U.

J'ai participé à ce travail en cherchant à savoir à quel moment l'inactivation du gène *hfp* permet la levée de la répression sur *BX2*, le transformant en cluster producteur de piARN actif.

L'article est en préparation.



# **Knock down of *half pint* gives birth to a meiotically stable piRNA-producing locus in *Drosophila*.**

## **Running title**

Loss of *hfp* activates a piRNA cluster

## **Authors**

Boivin Antoine<sup>1\*</sup>, Guegen Nathalie<sup>2,3,4</sup>, Marie P. Pauline<sup>1</sup>, Delmarre Valérie<sup>1</sup>, Vaury Chantal<sup>2,3,4</sup>, Brasset Emilie<sup>2,3,4\$</sup>, Ronsseray Stéphane<sup>1\$\*</sup>

\* co-corresponding authors

\$ co-last authors

## **Affiliations**

<sup>1</sup> Sorbonne Universités, Université Pierre et Marie Curie (UPMC), CNRS, Institut de Biologie Paris-Seine (IBPS), Laboratoire de Biologie du Développement, F-75005, Paris, France

<sup>2</sup> Inserm, UMR1103, F-63001 Clermont-Ferrand, France

<sup>3</sup> CNRS, UMR6293, F-63001 Clermont-Ferrand, France

<sup>4</sup> Clermont Université, Université d'Auvergne, Laboratoire GReD, BP 10448, F-63000 Clermont-Ferrand, France

## **Contacts**

antoine.boivin@upmc.fr or stephane.ronsseray@upmc.fr

Tel: (33) 1 44 27 22 88 / Fax: (33) 1 44 27 34 45

## **Keywords**

piRNA cluster / splicing factor / germline / silencing / *Drosophila*

## **Abstract**

In animal germline, specific loci undergo non-canonical transcription leading to the production of small RNAs that interact with PIWI proteins (piRNAs). piRNAs are responsible for the repression of transposable element expression and participate to the correct expression of developmental genes, thus ensuring both the protection of the genome against mobile invasive elements and an optimal development of the organism. How piRNA-producing loci are specified remains an open question. Using a unique transgenic locus that can be switch into a piRNA-producer locus, we show here that the *half pint (hfp)* gene, which encodes a poly-U binding splicing factor, is required to maintain the locus in a non-piRNA producer state. Indeed, knock down of *hfp* leads to the enrichment of the H3K9me3 mark and the recruitment of Rhino at the locus and to the production of *de novo* piRNAs from the locus. These changes are maintained over generations despite the return to a normal dosage of *hfp* thus corresponding to a genetically-induced epiconversion. Conceptually, our results demonstrate that a non piRNA-producing state is in fact not a "by default" state but rather a cellular choice that is genetically maintained.

## Introduction

Half of the human genome and 15% of the *Drosophila* one is made of transposable elements (TEs). To prevent TE mobilization that may cause deleterious DNA damage, insertional mutagenesis and chromosomal rearrangement, a class of small – 23-29 nt – non-coding RNAs associated with Argonaute protein subfamily has been dedicated specifically to the animal germline (ARAVIN *et al.* 2007; IWASAKI *et al.* 2015). These Piwi-interacting RNAs (piRNAs) repress TE expression using RNA sequence complementarity both through transcriptional gene silencing (TGS) and through post-transcriptional gene silencing (PTGS) (HIRAKATA AND SIOMI 2016). TGS is characterized by an enrichment of transcriptional repressive H3K9 methylation marks on the genomic TE copies, likely mediated through nascent transcript recognition by nuclear Piwi-loaded piRNAs. For TE copies that have escaped TGS, PTGS occurs thanks to recognition and cleavage of TE transcripts by the slicer activity of Aubergine (Aub) and Argonaute 3 (Ago3), two other piRNA-loaded Piwi family proteins. The majority of Aub and Ago3 are found in the nuage, a cytoplasmic dense region around the nucleus (BRENNECKE *et al.* 2007; LI *et al.* 2009; MALONE *et al.* 2009).

In *Drosophila melanogaster*, most piRNA sequences match to 142 genomic loci mostly located in pericentromeric and telomeric heterochromatin (BRENNECKE *et al.* 2007). These loci, called piRNA clusters, are heritable catalogs that constitute the memory of past TE invasions (BRENNECKE *et al.* 2007; SENTI AND BRENNECKE 2010). Despite their localization, their repetitive nature and their heterochromatic status, piRNA clusters are transcribed as long RNA Pol II-dependent single strand transcripts that are processed further into piRNAs (MOHN *et al.* 2014; ZHANG *et al.* 2014). Two types of piRNA clusters coexist in *Drosophila*. On one hand, uni-strand clusters, such as the *flamenco* locus, have a defined promoter that produce transcripts from one genomic strand which are alternatively spliced, giving rise to piRNAs that map only to this strand (BRENNECKE *et al.* 2007; MALONE *et al.* 2009; GORIAUX *et al.* 2014). On the opposite, dual-strand clusters, such as cluster 42AB, must be transcribed in both directions because of the existence of piRNAs matching both genomic strands. Recently, it has been shown that dual-strand piRNA clusters undergo Pol II dependent non-canonical transcription: licensed by the Rhino-Deadlock-Cutoff (RDC) complex (MOHN *et al.* 2014; ZHANG *et al.* 2014; CHEN *et al.* 2016). Current model favors this characteristic as the starting point of specification for a piRNA cluster.

One of the most mysterious question is how the RDC complex is deposited on piRNA cluster? The predominant idea is that piRNAs inherited maternally, loaded by PIWI, target

homologous sequences contained in piRNA clusters and bring RDC complex through protein-protein interactions. Fitting well with this idea, we have shown that maternal inheritance of homologous piRNAs can turn on an inactive cluster (*BX2*) composed of seven *P-lacZ* transgenes repeated in tandem and inserted in euchromatin (DE VANSSAY *et al.* 2012). Indeed, the repetitive nature of this locus is not sufficient *per se* to induce the production of piRNA (*BX2<sup>OFF</sup>* state). However, a maternal inheritance of homologous piRNAs allows switching on the piRNA production from *BX2* (*BX2\** state, (DE VANSSAY *et al.* 2012)). The activated state is then stably maintained through generations (DE VANSSAY *et al.* 2012; DE VANSSAY *et al.* 2013; HERMANT *et al.* 2015). In the subsequent generations, the cluster produces piRNAs and is enriched in H3K9me3 marks and Rhino (LE THOMAS *et al.* 2014).

Interestingly, piRNA production is lost if the *BX2* locus is paternally inherited and cannot be restored until a new maternal inheritance of homologous piRNAs. We took advantage of these characteristics to ask whether there is an active processes to maintain a such locus in a non-piRNA-producing state (*BX2<sup>OFF</sup>*). We then wondered if the loss of function of specific genes could activate the piRNA production from a paternally inherited *BX2* locus without a maternal inheritance of homologous piRNAs. Conceptually, finding such genes will demonstrate that non-piRNA-producing state (*BX2<sup>OFF</sup>*) is in fact not a "by default" state but rather a cellular choice that is actively maintained. We designed a F1 screen to test the ability of RNAi-induced loss of functions to switch on a paternally inherited *BX2* locus.

While the screen is still running, we report here that reducing the dosage of *hfp* lead to turn on the *BX2<sup>OFF</sup>* locus into a piRNA-producing *BX2\**-like. The locus acquires repression capacities, produces *bona fide* piRNAs, is enriched in Rhino and H3K9me3 chromatin marks. Moreover, flies are fertile, allowing to follow the progeny that recover the wild type dosage of *hfp*. Daughters still produce piRNA showing the stable acquirement of piRNA-producer state. These results establish, for the first time in *Drosophila*, that factors may actively restrain specific loci for piRNA production.

## MATERIALS and METHODS

**Fly strains.** All crosses were performed at 25°. Most of the *Drosophila* strains were obtained from the Bloomington Stock Centre, in particular TRiP lines listed in Table S1. *hfp*<sup>9</sup>, *hfp*<sup>13</sup>, *otu*<sup>7</sup>, *otu*<sup>11</sup> mutant alleles were kind gifts from T. Schupbach. P-lacZ fusion enhancer trap transgenes 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 marker (O'KANE AND GEHRING 1987). *BQ16* is located at 64C in euchromatin of the third chromosome and is homozygous, viable and fertile. BC69 is inserted in chromosome 2 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 and is homozygous lethal. *BQ16* and *BC69* are strongly expressed in the nurse cells and in the oocyte and *P-1039* shows strong  $\beta$ -Galactosidase staining in numerous tissues including the follicle cells, the nurse cells and the oocyte. *BX2* carries seven *P-lacZ-white* copies inserted in tandem at cytological site 50C on the second chromosome (DORER AND HENIKOFF 1994). The *P-lacZ-white* construct contains the *P-lacZ* translational fusion and is marked by the *mini-white* gene (*P{lacW}*, FBtp0000204).

### ***trans*-silencing quantification**

*lacZ* expression assays were carried out using X-gal overnight staining (LEMAITRE *et al.* 1993).

### **ChIP analyses**

100 Ovaries were dissected in Schneider medium supplemented with insulin at room temperature. The cross-linking was performed 10 min at room temperature in PBS 1X 1% formaldehyde (Sigma). The cross-linking reaction was stopped by addition of glycine at a final concentration of 0.125 mM in PBS 0,1% triton and incubating 5 min in ice. The cross-linked ovaries were washed with 1 ml of PBS 0,1% triton and crush in a potter with dounce A 20 times. Then a centrifugation at 400g 1min was performed. The pellet was suspended with 1 ml of cell lysis buffer (KCl 0,085M, HEPES 5mM, IGEPAL 0,5%, DTT 0,5mM, Na butyrate 10mM, 1/2 EDTA free protease inhibitor cocktail Roche) and crush in a potter with dounce B 20 times, then 2mL of cell lysis buffer were added. A centrifugation at 2000g 5 min was performed and the pellet was suspended with 0,5 mL nucleus buffer (HEPES 0,05M, EDTA 0,01M, N lauryl Sarcosyl 0,5%, Na butyrate 0,01M ¼ EDTA free protease inhibitor cocktail Roche) and incubated 15min in a cold room on a rotator. Sonication was performed



with Bioruptor (Diagenode) set to high power for 10 cycles (15 sec on and 15 sec off). A centrifugation was performed 15 min at 16K G at 4°C. Each IP is performed on 5 ug of chromatin. A preclear of 4 hours was performed with 25uL of dynabead Protein A. The immunoprecipitation reaction was performed with 50uL of dynabeads coated by 5ug of antibodies ( H3K9me3 polyclonal antibody C1540030 diagenode, Normal Mouse <IgG 12-371 Millipore), or 20uL of serum for the antibody against Rhino (W. Theurfauf) over night in a cold room in a rotator. 3 washes of 10 min in a Hight salt buffer (tris HCl pH7,5 0,05M, NaCl 0,5M triton 0,25%, IGEPAL 0,5%, BSA 0,5%, EDTA 5mM) were performed and the elution of Chromatin was performed 30 min with 500uL of elution buffer (Tris pH7,5 0,05M, NaCl 0,05M, EDTA 5mM, SDS 1%) ;RNase treatment was omitted; qPCR were performed in a light Cyclor Roche.

### **Small RNA deep sequencing**

Experiments were performed as in (HERMANT *et al.* 2015).

## RESULTS

### **A one-generation screen to identify loss of function that induces *trans*-silencing capacities**

The principle of the screen is to bring the *BX2*<sup>OFF</sup> locus in a loss-of-function context for a given gene and to test if *de novo* piRNAs are produced through a functional *trans*-silencing assay. On one hand, mothers bring a transgene from which a modified miRNA (shRNA), which is under a GAL4-dependent promoter, can induce the specific knockdown of a given gene (KD) (NI *et al.* 2011). On the other hand, fathers bring the *BX2*<sup>OFF</sup> cluster, a *nos-gal4* transgene that allows the expression of the shRNA specifically in the germline of the progeny (GLKD) and a euchromatic *P-lacZ* target transgene expressing the  $\beta$ -galactosidase in the germline and serving as a read out for the production of piRNA from the *BX2* locus. If *BX2* switches ON, the produced piRNAs will silence the expression of the *P-lacZ* (white ovaries), if *BX2* remains OFF, the target will be expressed (blue ovaries) (Fig 1A). We made a gene candidate approach and choose first to test the loss-of-function of genes that are known to act in RNA pathways (Table S1).

Currently, 221 TRiP lines, corresponding to 149 genes have been tested. 16,7% of the TRiP lines (n=37) result in severe atrophied ovaries for which no conclusion concerning the *BX2* turn on can be made and 39,8% of the TRiP lines (n=88) lead to sterile progeny (Table S1). Only two TRiP lines, *hfp* #34785 and *Rm62* #34829, lead to repressed egg chambers in ovaries (Fig 1B). Both are fertile. Since the control sisters of *hfp* GLKD that do not contain *BX2* locus are not silenced, we interpreted that the loss of *hfp* expression by itself is not required for the expression of the *P-lacZ* transgene (Fig 1B). On the contrary, *Rm62* appears to be required for the expression of the *P-lacZ* (Fig S1). For *Rm62*, we reproduced the experiment with another *P-lacZ* target (Fig S1) confirming that *Rm62* appears to be required for the expression of *P-lacZ* transgenes. While this is a potentially interesting finding, we do not develop this result further and focus on the *hfp* gene.

### **A paternally inherited *BX2* acquires *trans*-silencing capacity in *hfp* knock down context**

*hfp* (officially named *pUf68*) was first identified by homology to human PUF60, a poly-U binding splicing factor that could stimulate splicing *in vitro* (PAGE-MCCAW *et al.* 1999). Interestingly, *hfp* was found later through a two-hybrid assay against a large domain of the Encore protein that is required for the regulation of the germline mitosis, karyosome formation and establishment of dorsoventral polarity of the *Drosophila* egg and embryo (VAN BUSKIRK AND SCHUPBACH 2002). *hfp* was confirmed as a splicing factor but rather

involved in the specific regulation of the alternative splicing of a subset of genes, especially *ovarian tumor (otu)* in the female germline or *transformer2 (tra2)* in the male germline (VAN BUSKIRK AND SCHUPBACH 2002; WANG *et al.* 2013). Furthermore, Hfp was found to behave as a tumor suppressor through a complex with Haywire (Hay) that negatively regulates *dmyc* expression and cell growth (QUINN *et al.* 2004; MITCHELL *et al.* 2010; LEE *et al.* 2015). Hfp thus appears to act at different level of regulation on a subset of genes tightly regulated through time and space.

To address any potential off-target effects of the *hfp* TRiP line, we first tested whether known OFF targets of *hfp*<sup>KD</sup>, *rab5* and *metro*, are not able to confer the phenotype when directly targeted. GLKD of both genes give negative results (*rab5* TRiP lines #34832 TSE=0% n=1530 and #51847 TSE=0% n=1655 and *metro* TRiP line 35810 TSE=0% n=1150). Second, we tested whether characterized *hfp* mutant alleles also give the ability to a paternally inherited *BX2*<sup>OFF</sup> to silence a euchromatic *P-lacZ* transgene. While amorphic or strong alleles of *hfp* lead to homozygous lethality, *hfp*<sup>9</sup> and *hfp*<sup>13</sup> are two hypomorphic alleles that allow adult survival (VAN BUSKIRK AND SCHUPBACH 2002). They were both obtained from imprecise excisions of a transgene (*EP(3)3058*) located in the 5' UTR of the *hfp* gene (VAN BUSKIRK AND SCHUPBACH 2002). *Drosophila* lines carrying the *BX2*<sup>OFF</sup> locus, a euchromatic *P-lacZ* transgene used as *trans*-silencing read out (*P-1039*) and a *hfp* mutant allele were built. Males from such lines were crossed to females harboring *hfp* mutant allele. In the progeny, *trans*-silencing assays were performed. Table 1 summarizes results and shows that while there is no dominant effect of *hfp* alleles, the *BX2* locus can be switch ON for repression in homo or hetero mutant allelic combinations (Table 1). Interestingly, we observe that the strongest hypomorphic allele (*hfp*<sup>9</sup>) leads to the strongest *BX2* activation (almost 100% *trans* silencing) while the weakest allele (*hfp*<sup>13</sup>) leads to a weak *BX2* activation (only 2% *trans* silencing). Heterozygous combinations lead to intermediate activation with no significant maternal effect detected. These results confirm the role of *hfp* in the maintenance of a non-silencer state for the *BX2* locus.

Parental cross <i>mat x pat</i>	Progeny genotype	Ovariole number	Non repressed egg chambers	Repressed egg chambers	<i>trans</i> silencing (%)
<i>hfp</i> <sup>13/+</sup> x <i>hfp</i> <sup>9/+</sup>	<i>hfp</i> <sup>9/+</sup>	95	475	0	0
	<i>hfp</i> <sup>13/+</sup>	476	2380	0	0

	<i>hfp</i> <sup>13/9</sup>	232	750	257	25,5
<i>hfp</i> <sup>9/+</sup> x <i>hfp</i> <sup>13/+</sup>	<i>hfp</i> <sup>9/+</sup>	401	2005	0	0
	<i>hfp</i> <sup>13/+</sup>	264	1320	0	0
	<i>hfp</i> <sup>9/13</sup>	141	442	248	35,9
<i>hfp</i> <sup>13/+</sup> x <i>hfp</i> <sup>13/+</sup>	<i>hfp</i> <sup>13/+</sup>	1369	6845	0	0
	<i>hfp</i> <sup>13/13</sup>	651	3185	70	2,2
<i>hfp</i> <sup>9/+</sup> x <i>hfp</i> <sup>9/+</sup>	<i>hfp</i> <sup>9/+</sup>	557	2779	6	0,2
	<i>hfp</i> <sup>9/9</sup>	196	16	954	98,4

Table 1. *hfp* allelic mutant combinations confirmed the effect observed with the TRiP line. The parental cross corresponds to females *w*; *hfp*<sup>mat</sup> / *TM6B*, *hu* crossed by males *w*<sup>1118</sup>; *BX2*, *P-1039* / *CyO* ; *hfp*<sup>pat</sup> / *TM2*, *Ubx*. For clarity, only the maternal and paternal *hfp* alleles are shown in the table.

### The genetic epiconversion is stable through generations after the return to a normal dosage of *hfp*

The fertility of *hfp* GLKD flies allows to follow *BX2* silencing capacities in the next generations. We wonder if the cluster will keep its repressive capacity once the *hfp* expression get back to normal. A genetic scheme has been designed to follow the *BX2* cluster through generations (Fig 2). At each generation (except G2), some females were crossed to males bearing a *P-lacZ* transgene (*BQ16*) in order to measure *BX2* silencing capacities in the progeny ovaries thanks to *lacZ* staining. Seven independent lines were made from unique females. All present 100% silencing capacity that is maintained through generations. These results show that even after the loss of the activated GLKD due to the segregation of *nosGAL4* and *hfpKD* transgenes and, thus, to a return to a normal dosage of *hfp*, the *BX2* cluster remains fully active for repression capacities (Fig 2). The genetic epiconversion induces by the loss of *hfp* in G1 is thus converted into a stable epigenetic modification.

We then wondered if the *BQ16* (*P-lacZ*) could influence the ability of *hfp* GLKD to activate the *BX2* cluster. We reproduce the experiment without the *BQ16* transgene and show that *hfp* GLKD was still able to convert a naive *BX2* into a strong silencer (Fig S2).

### ***hfp* GLKD modifies the chromatin of *BX2* and allows the production of *de novo* piRNAs from *BX2***

It has been showed that an activated *BX2* locus that produces piRNA (*BX2*<sup>ON</sup>) presents an enrichment of H3K9me3 marks when compared to *BX2*<sup>OFF</sup> (LE THOMAS *et al.* 2014). We wonder if *BX2* activated by *hfp* GLKD could present such an enrichment. CHiP analyses on *w* sequences that are specific of *BX2* locus revealed that H3K9me3 marks are significantly enriched on the locus in *hfp* GLKD compared to their control sisters (called S flies, Fig 3B). Furthermore, Rhino also appear to be enriched in *BX2* activated by *hfp* GLKD compared to *BX2*<sup>OFF</sup> (Fig 3C). These results show that *hfp* GLKD induces chromatin modifications at the *BX2* locus, as a *BX2* activated by maternally inherited homologous piRNAs (LE THOMAS *et al.* 2014).

We next ask if *de novo* piRNAs could be produced from *BX2*, correlatively to silencing capacities and chromatin modifications. Deep sequencing of small RNAs coming from *hfp* GLKD and their control sisters (S flies) but also from daughters of *hfp* GLKD (called D flies) and daughters of control sisters (corresponding to the nieces of *hfp* GLKD flies and therefore called N flies, see Fig 4A) were performed. Small RNAs homologous to *BX2*-contained *Placw* transgene were detected in *hfp* GLKD and their progeny (D flies) while almost none were observed in S and N flies (Fig 4B). 23-29-nt small RNAs in *hfp* GLKD and D flies were enriched in 1U (11442/25329 (45,2%) and 13565/31342 (43,3%), respectively) but they appear to be not enriched in A10 (6443/25329 (25,4%) and 8205/31342 (26,2%), respectively). Correlatively, they do not present a strong ping pong signature (Fig S3B). Taken together, these data suggest that they are primary piRNAs. Thus, these piRNAs differs from those produced by *BX2* activated by maternal inheritance of homologous piRNAs since in the latter case, 23-29 RNAs present strong ping-pong signature (DE VANSSAY *et al.* 2012; HERMANT *et al.* 2015).

In parallel, small RNA analyzes allow to detect the shRNA produced in the *hfp* GLKD females as well as its absence in all others genotypes, confirming the good segregation of *nosgal4* and *hfp*<sup>KD</sup> transgenes in D females (Fig S3C).

Taken together, the results presented here show that *hfp* GLKD induces chromatin modifications, piRNA production and *trans*-silencing capacities of the *BX2* locus. Here, we demonstrate for the first time that a potential piRNA cluster can be genetically maintained in a non-piRNA producer state revealing an interesting world of competition for the fate of such sequences in the genome.

### ***hfp* GLKD does not induce global changes in piRNA production**

Therefore, we seek through deep sequencing data mining if piRNA production from known piRNA clusters could be affected and, above all, if other *de novo* piRNAs could be produced, i.e. if *hfp* is required to prevent the formation of piRNAs from other genome region than the *BX2* cluster. Comparing 23-29-nt smRNAs that match known piRNA clusters between *hfp* *GLKD*, S, D and N females revealed no major change in the piRNA production (Fig 5A). We also searched for other regions yet no reported to be piRNA producer loci. As positive control, we detect *de novo* piRNAs corresponding to the *white* sequences and matching the *white* gene locus on the X chromosome (Fig 5B). These piRNAs were produced by the *BX2* cluster on the chromosome 2, activated in *hfp* *GLKD* and D females. Despite fluctuations in telomeric and pericentromeric regions, no other euchromatic regions appear to produce new piRNAs in *hfp* *GLKD* context (Fig 5 and data not shown yet). In heterochromatic regions, fluctuations do not appear to be dependent on the genotype (new positions can be observed for *hfp* *GLKD*, D, S or N females) and likely reflect the high diversity of piRNAs that can be produced. Not all positions always give rise to piRNAs.

These results suggest that *hfp* *GLKD* acts only on the fate of *BX2* cluster, probably because no others sequences are such in an equilibrium state as *BX2*. In the genome there is no other highly tandem repetitive sequences that are not already piRNA-producing locus.

### ***otu* mutants do not activate *BX2* for trans silencing capacity**

*hfp* could act indirectly on *BX2*. Indeed, Hfp has been reported to participate to the splicing of numerous genes (VAN BUSKIRK AND SCHUPBACH 2002; WANG *et al.* 2013). Among them, *otu* appears to be a good candidate since the alternative 6a exon of *otu* that is spliced by *hfp* (VAN BUSKIRK AND SCHUPBACH 2002) contains a TUDOR domain that is shared by several piRNA-processing proteins such as *eggless*, *krimper*, *qin* or *spindle E* (HIRAKATA AND SIOMI 2016). In wild-type, two isoforms are produced : a 98 kDa protein and a less abundant 104 kDa TUDOR domain-containing protein. Lack of *hfp* decreases the amount of the 104 kDa isoform (VAN BUSKIRK AND SCHUPBACH 2002). *otu* encodes a deubiquitinilase (DUB) protein implied in numerous processes such as chromosome organization, oogenesis, female germ-line sex determination, germ cell development, germ-line sex determination, positive regulation of DNA endoreduplication, protein deubiquitination and germ-line cyst formation (Flybase) but the specific roles of each isoform

are unknown. We reasoned that if *hfp*'s effect in the activation of *BX2* requires the loss of Otu 104 kDa, then we could reproduce the activation of *BX2* with Otu 104 kDa mutant. *otu<sup>11</sup>* allele mutates specifically the TUDOR domain (a C343Y substitution in the TUDOR domain, (STEINHAUER AND KALFAYAN 1992)) and could be combined with the nonsense *otu<sup>7</sup>* allele that leads to truncated both isoforms of OTU.

Lines containing *otu* alleles, *BX2* and *BQ16* were built and crossed in order to put a paternally inherited *BX2* (*BX2<sup>OFF</sup>* state) in different *otu* mutant contexts (Fig S4). Unambiguous results show that *otu* mutant contexts do not reproduce the *hfp* mutant effect on *BX2* (Table 2). From these results, we conclude that *hfp* GLKD activates *BX2* for piRNA production through another pathway than *otu*.

Parental cross <i>mat x pat</i>	Progeny genotype	Ovariole number	Non repressed egg chambers	Repressed egg chambers	<i>trans</i> silencing (%)
<i>otu<sup>11/+</sup> x otu<sup>11</sup></i>	<i>otu<sup>11/11</sup></i>	113	565	0	0
	<i>otu<sup>11/+</sup></i>	344	1720	0	0
<i>otu<sup>11/+</sup> x otu<sup>7</sup></i>	<i>otu<sup>11/7</sup></i>	307	1535	0	0
	<i>otu<sup>+/7</sup></i>	686	3430	0	0

Table 2. *otu* allelic mutant combinations do not reproduce the effect observed with *hfp* GLKD on *BX2*. The parental cross corresponds to females *otu<sup>11</sup> / M5 ; BQ16* crossed by males *otu<sup>11</sup>* or *otu<sup>7</sup>*; *BX2 / CyO*. For clarity, only the maternal and paternal *otu* alleles are shown in the table.

## DISCUSSION

### Is *hfp* required to prevent loci made of repetitive copies to produce piRNAs?

Data presented here clearly demonstrate that the partial loss of function of *hfp* induces the conversion of an inactive *BX2* to a fully characteristic piRNA-producer locus. Changes are equivalent to those observed in case of maternally inherited homologous piRNAs: *trans* silencing capacities, chromatin modifications, piRNA production and stability through generations despite a return to a wild type dosage of *hfp*. Our results demonstrate that a non piRNA-producing state is in fact not a "by default" state, resulting from the lack of inherited homologous piRNAs but rather a choice that is genetically maintained. A mechanism appears thus to exist in order to prevent piRNA machinery to transform all repetitive transcripts into

piRNAs. Is *hfp* role limited to *BX2*? By bioinformatic analyses, we failed to find new other piRNA producing region in a *hfp* GLKD context. However, we have to look more precisely to known tandem repetitive regions such as tRNA or histone-coding loci that could be less accurately annotated, to see if *hfp* GLKD could not change the piRNA production. On a functional point of view, the fact that *hfp* GLKD leads to fertile flies leads to think that it does not induce major changes and does not alter important genes.

### **Direct or indirect mechanism?**

Through which mechanism(s), *hfp* could maintain *BX2* in a non-piRNA producer state? As *hfp* is known to behave as splicing factor but also as a transcriptional regulator of several genes (VAN BUSKIRK AND SCHUPBACH 2002; QUINN *et al.* 2004; MITCHELL *et al.* 2010; WANG *et al.* 2013; LEE *et al.* 2015), its role could be indirect. Here we first tested *otu* because of the solid description of the *hfp* effect (VAN BUSKIRK AND SCHUPBACH 2002) as well as the availability of specific mutants. Others *hfp* target genes should be tested through true mutant if possible or through the GLKD screen we designed. RNA seq comparisons between *hfp* GLKD and wild type would certainly help to restrain the list of candidate genes.

Alternatively, the role of *hfp* on *BX2* could be direct. Previously we failed to find any difference in steady-state amount of *P-lacZ* transcripts between non piRNA producer locus *BX2<sup>OFF</sup>* and piRNA producer locus *BX2\** using specific strand qRT-PCR (DE VANSSAY *et al.* 2012). While we cannot exclude that *BX2\** is in fact more transcribed and that the excess of transcripts is targeted to the piRNA pathway, this result shows at least that the *BX2<sup>OFF</sup>* locus is transcribed and suggests that transcripts likely follow a "degradome" pathway since no expression of the  $\beta$ -galactosidase is observed in ovaries. The starting point of the loss of function screen was that we reasoned that if we reduce the "degradome" pathway activity, we could change the fate of the *BX2<sup>OFF</sup>* transcripts, thus hoping to force the piRNA pathway to take in charge the *BX2<sup>OFF</sup>* transcripts. *hfp* could be required in such a "degradome" pathway and its loss of function could induce an accumulation of *BX2* transcripts that could alert the piRNA machinery.

A even more direct role of *hfp* could be proposed. A hallmark of piRNA-producer bi-directional cluster is that the Rhino-Deadlock-Cutoff (RDC) complex induces a non-canonical transcription including the loss of splicing and transcription termination signals (MOHN *et al.* 2014; ZHANG *et al.* 2014; CHEN *et al.* 2016). *rhino* and *cutoff* has been shown to be required for piRNA production from *BX2* (HERMANT *et al.* 2015) and we and others show



that Rhino is enriched in  $BX2^{ON}$  (this study and (LE THOMAS *et al.* 2014). It appears thus that  $BX2$  behaves as a bi-directional cluster. While this had to be demonstrated, long RNAs, both sense and antisense, and covering the entire length of the cluster, should be produced. We can then imagine that Hfp could compete Cutoff at the splicing level then determining the fate of  $BX2$  transcripts. In a *hfp* loss of function context, Cutoff send  $BX2$  transcripts into the piRNA machinery. Once sufficient amounts of Piwi-loaded piRNAs homologous to  $BX2$  are produced, the cluster remains activated for generations. This hypothesis implies that it could be possible to change the fate of transcripts and to produce piRNAs while the loci was not specify as a piRNA cluster, a crucial step that remains to be demonstrated.

## ACNOWLEDGEMENTS

## REFERENCES

- Aravin, A. A., G. J. Hannon and J. Brennecke, 2007 The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 318: 761-764.
- Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis *et al.*, 2007 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128: 1089-1103.
- Chen, Y. A., E. Stuwe, Y. Luo, M. Ninova, A. Le Thomas *et al.*, 2016 Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors. *Mol Cell*.
- de Vanssay, A., A. L. Bouge, A. Boivin, C. Hermant, L. Teyssset *et al.*, 2012 Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature* 490: 112-115.
- de Vanssay, A., A. L. Bouge, A. Boivin, C. Hermant, L. Teyssset *et al.*, 2013 piRNAs and epigenetic conversion in *Drosophila*. *Fly (Austin)* 7: 237-241.
- Dorer, D. R., and S. Henikoff, 1994 Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. *Cell* 77: 993-1002.
- Goriaux, C., S. Desset, Y. Renaud, C. Vaury and E. Brasset, 2014 Transcriptional properties and splicing of the flamenco piRNA cluster. *EMBO Rep* 15: 411-418.
- Hermant, C., A. Boivin, L. Teyssset, V. Delmarre, A. Asif-Laidin *et al.*, 2015 Paramutation in *Drosophila* Requires Both Nuclear and Cytoplasmic Actors of the piRNA Pathway and Induces Cis-spreading of piRNA Production. *Genetics* 201: 1381-1396.
- Hirakata, S., and M. C. Siomi, 2016 piRNA biogenesis in the germline: From transcription of piRNA genomic sources to piRNA maturation. *Biochim Biophys Acta* 1859: 82-92.
- Iwasaki, Y. W., M. C. Siomi and H. Siomi, 2015 PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu Rev Biochem* 84: 405-433.
- Le Thomas, A., E. Stuwe, S. Li, J. Du, G. Marinov *et al.*, 2014 Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev* 28: 1667-1680.
- Lee, J. E., N. C. Mitchell, O. Zaytseva, A. Chahal, P. Mendis *et al.*, 2015 Defective Hfp-dependent transcriptional repression of dMYC is fundamental to tissue overgrowth in *Drosophila* XPB models. *Nat Commun* 6: 7404.
- Lemaitre, B., S. Ronsseray and D. Coen, 1993 Maternal repression of the P element promoter in the germline of *Drosophila melanogaster*: a model for the P cytotype. *Genetics* 135:

- 149-160.
- Li, C., V. V. Vagin, S. Lee, J. Xu, S. Ma *et al.*, 2009 Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell* 137: 509-521.
- Malone, C. D., J. Brennecke, M. Dus, A. Stark, W. R. McCombie *et al.*, 2009 Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* 137: 522-535.
- Mitchell, N. C., T. M. Johanson, N. J. Cranna, A. L. Er, H. E. Richardson *et al.*, 2010 Hfp inhibits *Drosophila* myc transcription and cell growth in a TFIID/Hay-dependent manner. *Development* 137: 2875-2884.
- Mohn, F., G. Sienski, D. Handler and J. Brennecke, 2014 The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* 157: 1364-1379.
- Ni, J. Q., R. Zhou, B. Czech, L. P. Liu, L. Holderbaum *et al.*, 2011 A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat Methods* 8: 405-407.
- O'Kane, C. J., and W. J. Gehring, 1987 Detection in situ of genomic regulatory elements in *Drosophila*. *Proc Natl Acad Sci U S A* 84: 9123-9127.
- Page-McCaw, P. S., K. Amonlirdviman and P. A. Sharp, 1999 PUF60: a novel U2AF65-related splicing activity. *RNA* 5: 1548-1560.
- Quinn, L. M., R. A. Dickins, M. Coombe, G. R. Hime, D. D. Bowtell *et al.*, 2004 *Drosophila* Hfp negatively regulates *dmyc* and *stg* to inhibit cell proliferation. *Development* 131: 1411-1423.
- Senti, K. A., and J. Brennecke, 2010 The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet* 26: 499-509.
- Steinhauer, W. R., and L. J. Kalfayan, 1992 A specific ovarian tumor protein isoform is required for efficient differentiation of germ cells in *Drosophila* oogenesis. *Genes Dev* 6: 233-243.
- Van Buskirk, C., and T. Schupbach, 2002 Half pint regulates alternative splice site selection in *Drosophila*. *Dev Cell* 2: 343-353.
- Wang, S., E. J. Wagner and W. Mattox, 2013 Half pint/Puf68 is required for negative regulation of splicing by the SR splicing factor Transformer2. *RNA Biol* 10: 1396-1406.
- Zhang, Z., J. Wang, N. Schultz, F. Zhang, S. S. Parhad *et al.*, 2014 The HP1 homolog rhino anchors a nuclear complex that suppresses piRNA precursor splicing. *Cell* 157: 1353-1363.

## FIGURE LEGENDS

FIGURE 1. A F1 screen to find genes required to repress piRNA production. A. Genetic scheme used: females carrying modified miRNA that target the tested gene and that is under the transcriptional control of *UAS* were crossed with males bearing inactive *BX2* locus, *nosgal4* transgene that allows the expression of the modified miRNA specifically in the germline and a *P-lacZ* transgene (*BQ16*) that allows to test in the progeny if *BX2* become able to silence homologous sequences (read out). B. hfp GLKD does not modify the expression of *P-lacZ* without *BX2* locus (control on the left) while in presence of *BX2*, the expression of *P-lacZ* is repressed (84% n=1785).

FIGURE 2. *BX2* locus is stably activated for repression capacities through generations. The scheme shows crosses performed at each generation in order to induce the epigenetic conversion (G1), to establish 7 independent lines from seven females (G2), to establish lines that show complete repression capacities over time, as tested by crosses with a line carrying a *P-lacZ* transgene.

FIGURE 3. *BX2* locus is enriched in H3K9me3 marks and Rhino in a *hfp* GLKD context. A. Genetic scheme used to produced *hfp* GLKD flies and their control sisters (S flies). B. H3K9me3 signal was measured by ChIP-qPCR on one region of *w* (n = 3). Signal was normalized to 42AB cluster (mean of two regions). Error bars represent the standard error of the mean between four independent biological replicates. A Mann Whitney test was used to calculate significance (P = 0.0286). C. Rhino signal was measured by ChIP-qPCR on one region of *w* (n = 3). Signal was normalized to 42AB cluster (mean of two regions). Error bars represent the standard error of the mean between four independent biological replicates. A Mann Whitney test was used to calculate significance (P = 0.0286).

FIGURE 4. Stable *de novo* piRNAs are produced in a *hfp* GLKD context. A. Genetic scheme used to produced *hfp* GLKD flies and their control sisters (S flies). By crossing these females by *w* males, G2 (daughter (D) and niece (N)) flies were generated. B. SmRNA deep sequencing of ovaries from *hfp* GLKD, D, S and N flies. Plots show the abundance of 23- to 29-nt smRNAs matching the *P-lacZ-white* transgene and histograms show the distribution of 19- to 28-nt smRNAs matching the *P-lacZ-white* transgene. Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively.

FIGURE 5. *hfp* GLKD does not affect global piRNA production. A. The 60 most expressed piRNA cluster in *hfp* GLKD females do not show dramatic changes in their piRNA production compared to D, S and N females. Name of the cluster is given in the first column, number of 23-29-nt smRNAs in *hfp* GLKD females is given in the second column. In D, S and N females, the log2 ratio compared to *hfp* GLKD females is given as a heat map. B. Comparative analysis of specific piRNAs from *hfp* GLKD flies versus control S flies. Empty circles represent specific piRNAs from *hfp* GLKD flies while full circles represent specific piRNAs from control S flies. Positive and negative values correspond to sense (blue) and antisense (red) reads, respectively. X-axis corresponds to X chromosome coordinates, Y-axis represents the number of specific smRNA found at one position.

## SUPPLEMENTAL FILES

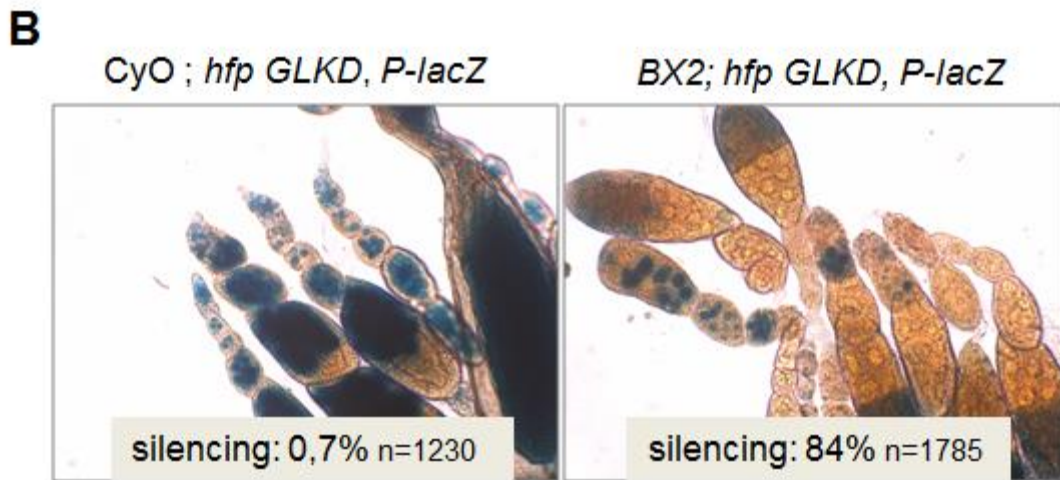
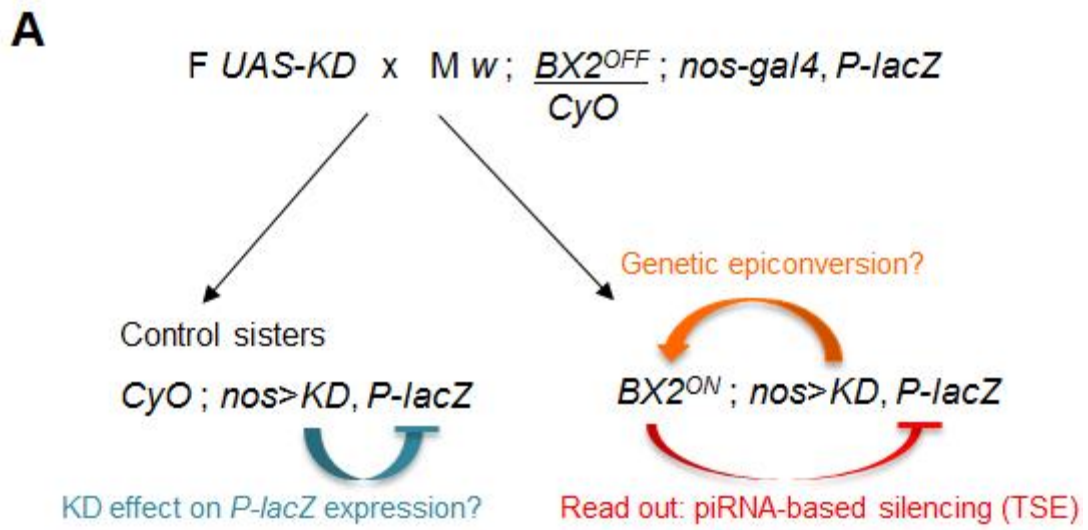
FIGURE S1. *Rm62* is required for the expression in *P-lacZ* in the germline. A. Loss of *Rm62* induces repression of the expression of *BQ16* with (right) and without (left) the *BX2* cluster. B. Confirmation that the loss of *Rm62* alters the expression of the *BQ16 P-lacZ*, independently of the *BX2* cluster. C. Loss of *Rm62* also alters the expression of another *P-lacZ*, *BC69*, suggesting that *Rm62* encodes or regulates an essential factor for *P*-dependent transcription.

FIGURE S2. A euchromatic *P-lacZ* sequence is not required for the epigenetic conversion of *BX2* in *hfp* GLKD context. G1 females were crossed by males carrying *BQ16* in order to test in the progeny if the genetic epiconversion of *BX2* by *hfp* GLKD has occurred.

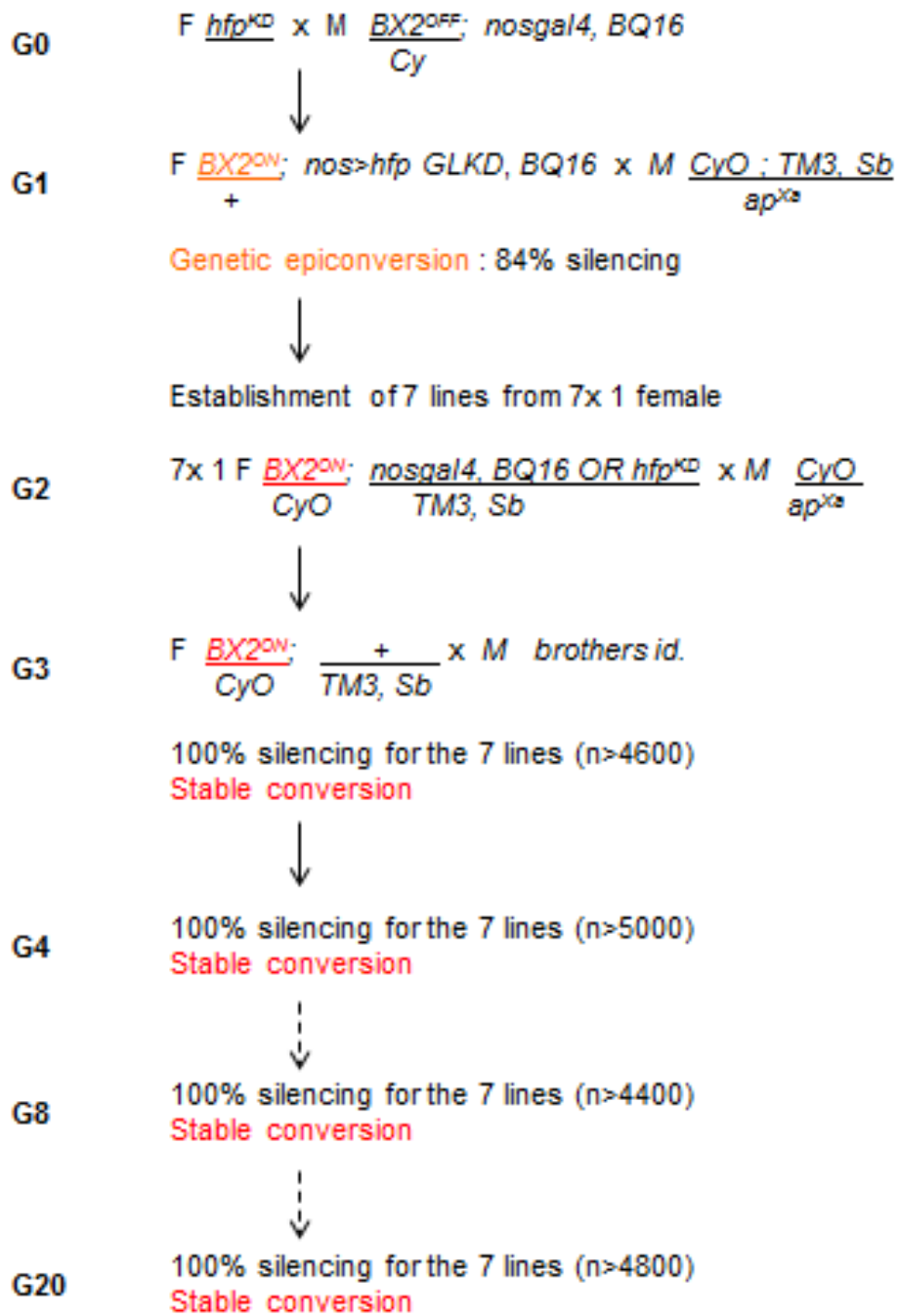
FIGURE S3. Only *hfp* GLKD flies produce the modified miRNA against *hfp*. A. Genetic scheme used to produce *hfp* GLKD flies and their control sisters (S flies). By crossing these females by w males, G2 (daughter (D) and niece (N)) flies were generated. B. Number of overlapping sense-antisense small RNA pairs (Y-axis) in the subset of 23-28nt smRNAs (ping-pong signature, X-axis shows the overlap in nucleotides). C. SmRNA deep sequencing of ovaries from *hfp* GLKD, D, S and N flies. Plots show the abundance of 23- to 29-nt smRNAs matching the *hfp* locus and histograms show the distribution of 19- to 28-nt smRNAs matching the *hfp* locus. Positive and negative values correspond to sense (red) and antisense (blue) reads, respectively. This is a control to show that in D flies, transgenes *nosgal4* and *hfp<sup>KD</sup>* have segregated.

FIGURE S4. *otu* mutant alleles do not mimic *hfp* GLKD effect on *BX2* repression capacities. Homozygous *otu<sup>11/11</sup>* context is not able to initiate *BX2* repression capacities. Inserts show the non decondensation of chromosomes in latter stage of oogenesis in *otu<sup>11/11</sup>* ovaries compared to heterozygous control (DAPI staining).

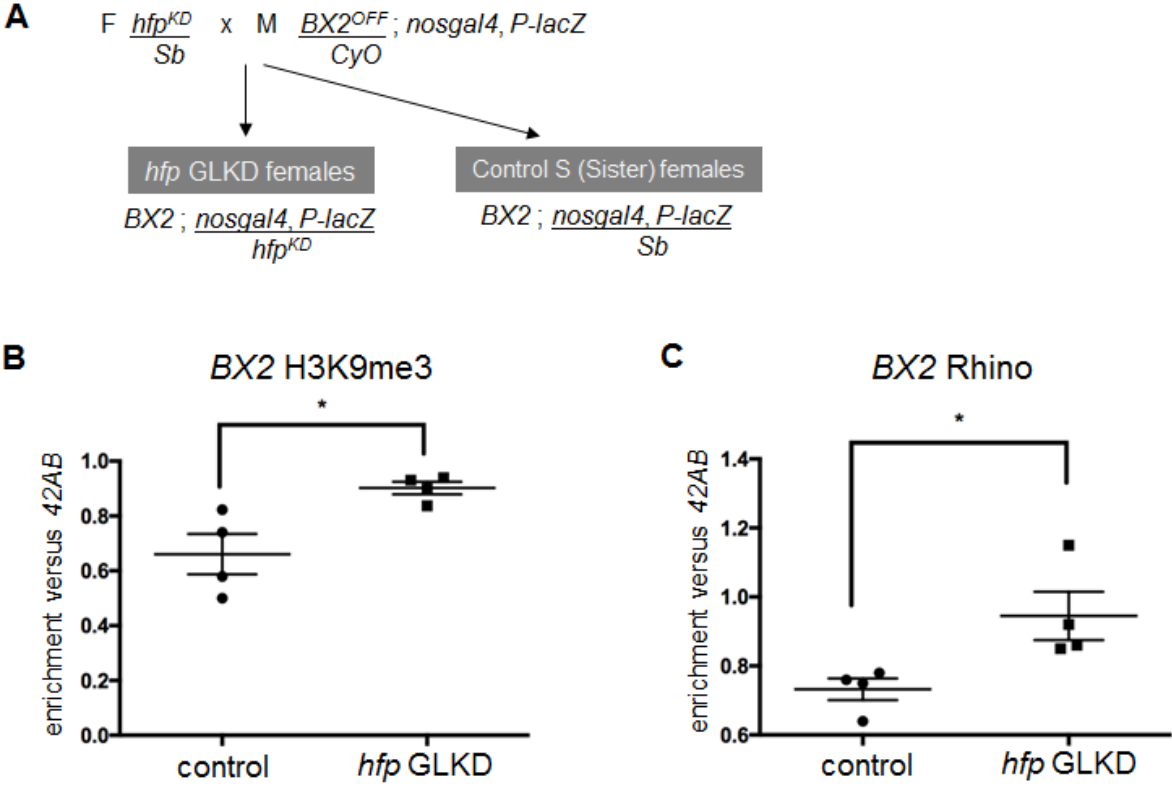
**Figure 1.**



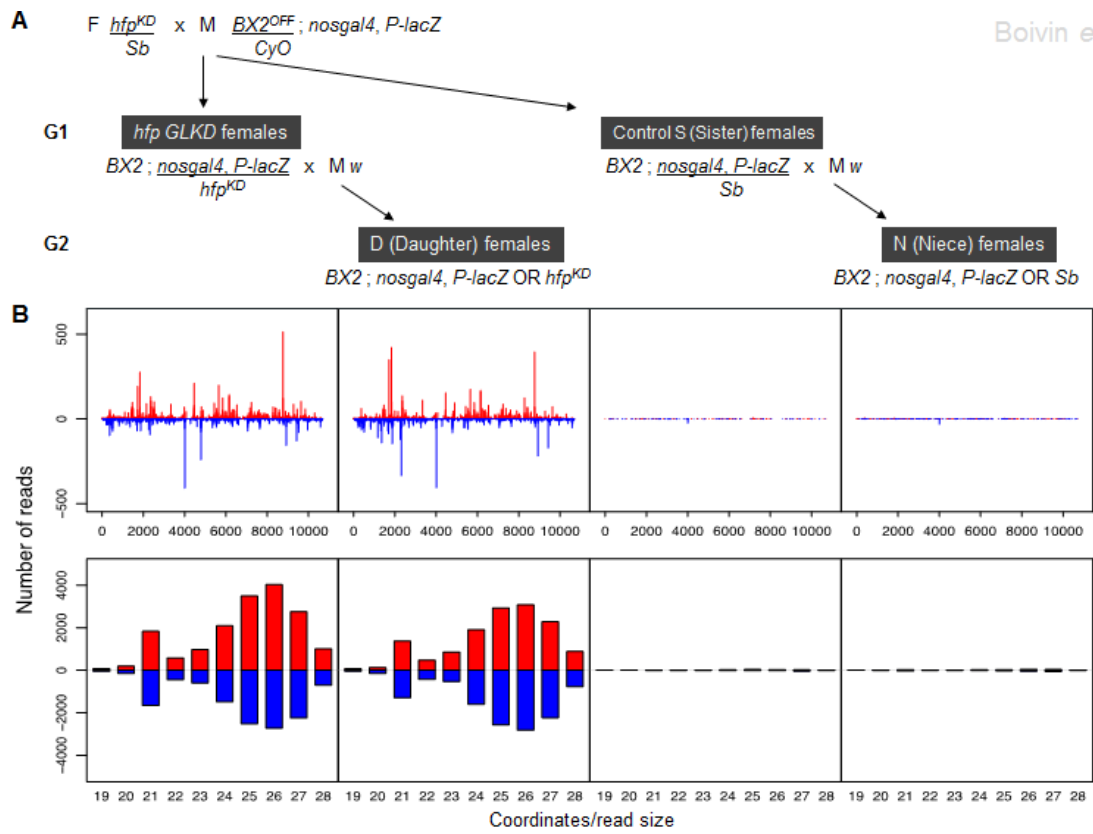
**Figure 2.**



**Figure 3.**

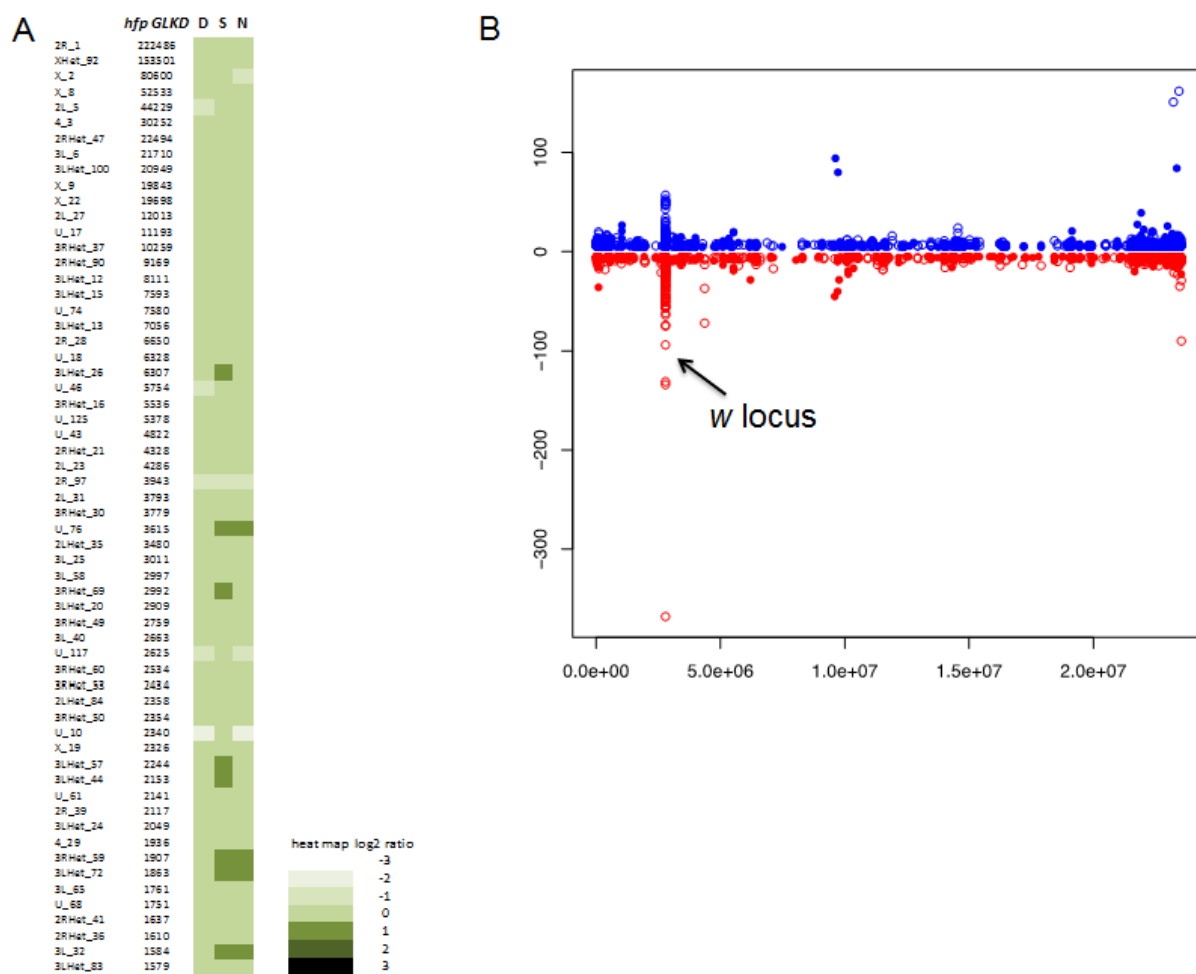


**Figure 4.**





**Figure 5.**



**Table S1.**

Gene	Stock number	Valium type	Loc	G1 TSE	G1 fertility
Ada2a,Rpb4	50905	20	attP40	nd	sterile (atrophied ovaries)
Ada2b	35334	22	attP2	OFF	fertile
ago1	31700	1	attP2	OFF	fertile
ago1	33727	20	attP2	OFF	sterile
ago2	34799	20	attP2	OFF	fertile
ago3	35232	22	attP2	OFF	sterile
ago3	44543	20	attP40	OFF	sterile
aret	35394	22	attP2	OFF	sterile (atrophied ovaries)
aret	38983	20	attP40	OFF	sterile (atrophied ovaries)
aret	44483	20	attP2	OFF	sterile
ash1	33705	20	attP2	OFF	fertile
ash1	36803	22	attP40	OFF	fertile
Asx	51677	22	attP2	OFF	fertile
Atac2	32890	20	attP2	OFF	fertile
Atac2	53918	20	attP40	nd	sterile (atrophied ovaries)
B52	37519	20	attP40	OFF	fertile
belle	35185	22	attP2	nd	sterile (atrophied ovaries)
belle	35302	22	attP2	OFF	fertile
HnRNP-K	42540	20	attP40	OFF	fertile
blanks	33667	20	attP2	OFF	nt
brm	34520	20	attP2	OFF	fertile
brm	35211	22	attP2	OFF	fertile
Caf1	34069	20	attP2	nd	sterile (atrophied ovaries)
CG10418	44458	22	attP2	nd	sterile (atrophied ovaries)
CG10445	43137	22	attP40	OFF	sterile
CG11266	44431	22	attP2	OFF	fertile
CG12493	42791	22	attP2	OFF	fertile
CG13472	36819	22	attP2	OFF	fertile
CG16941	34840	20	attP2	nd	sterile (atrophied ovaries)
CG17454	43199	22	attP40	nd	sterile (atrophied ovaries)
CG2926	34941	20	attP2	OFF	nt
CG5728	36592	22	attP2	OFF	fertile
CG6506	42552	20	attP40	OFF	fertile
CG6841	51909	20	attP40	nd	sterile (atrophied ovaries)
CG7185	34804	20	attP2	nd	sterile (atrophied ovaries)
CG7878	35229	22	attP2	OFF	fertile
CG8273	34805	20	attP2	OFF	fertile
CG8920	34738	20	attP2	OFF	fertile
CG9684	36880	22	attP2	OFF	fertile
CG9925	35811	20	attP2	OFF	fertile
cup	35406	22	attP2	OFF	sterile
dcr2	33656	20	attP2	OFF	fertile
dhd	41857	22	attP2	OFF	fertile
dik	32451	20	attP2	OFF	fertile
DIP1	35226	22	attP2	OFF	fertile
DIP1	35333	22	attP2	OFF	fertile
DIP2	34918	20	attP2	OFF	fertile
DIP2	42598	20	attP40	OFF	fertile
dl	32934	20	attP2	OFF	sterile
dl	36650	22	attP40	OFF	sterile

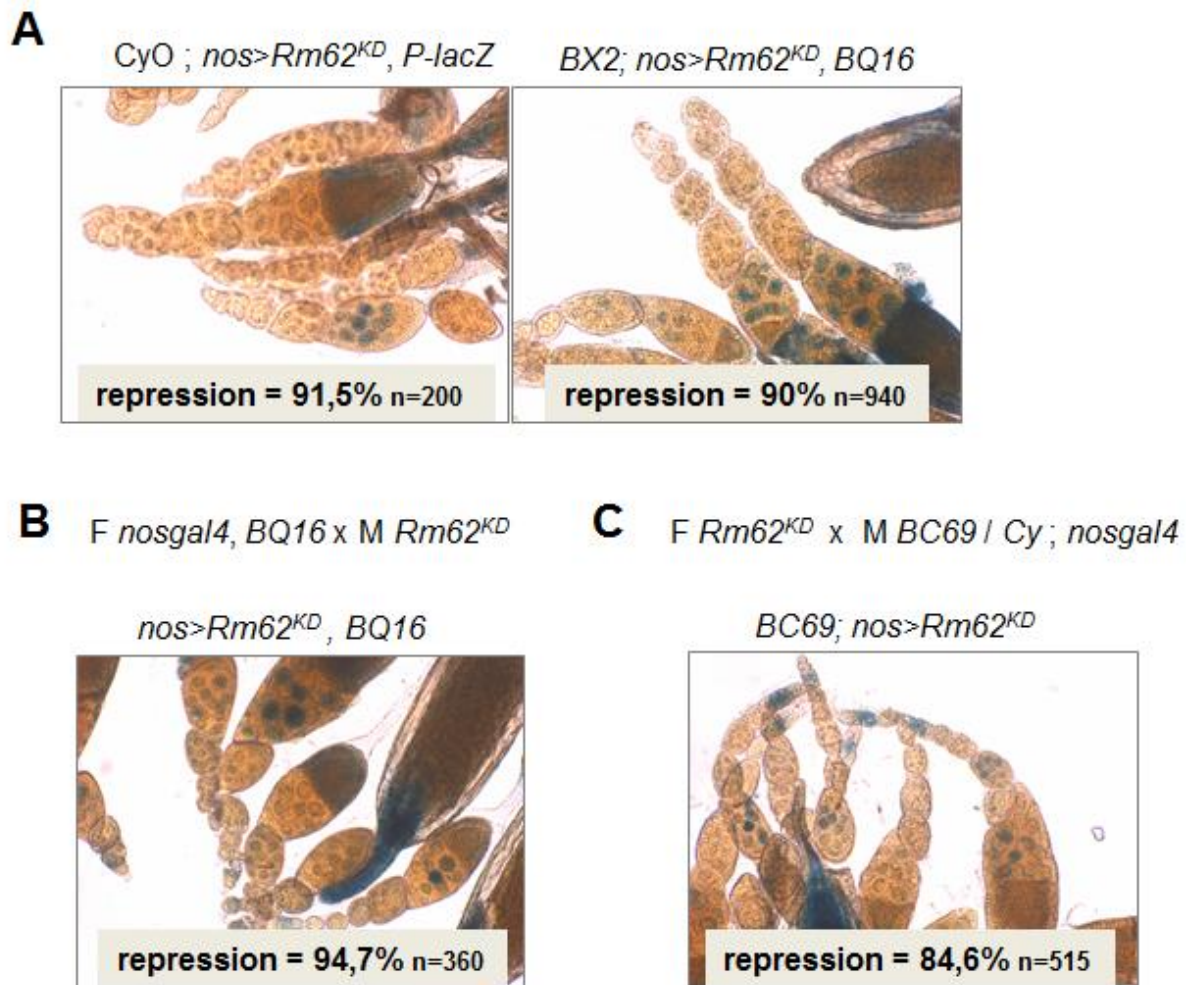
dl	38905	22	attP2	OFF	sterile
Doa	50903	20	attP40	OFF	sterile
dom	34827	20	attP2	nd	sterile (atrophied ovaries)
dom	38385	20	attP2	nd	sterile (atrophied ovaries)
dom	38941	20	attP2	OFF	fertile
dom	40914	20	attP2	OFF	sterile (atrophied ovaries)
dom	41674	20	attP2	nd	sterile (atrophied ovaries)
Dp1	32872	20	attP2	OFF	fertile
drosha	33657	20	attP2	OFF	nt
drosha	35233	22	attP2	OFF	fertile
E(bx)	33658	20	attP2	OFF	fertile
E(bx)	35353	22	attP2	nt	nt
E(Pc)	35271	22	attP2	nd	sterile (atrophied ovaries)
E(z)	33659	20	attP2	OFF	nt
E(z)	36068	22	attP40	OFF	sterile
eIF4AIII	32444	20	attP2	nd	sterile (atrophied ovaries)
eIF4AIII	32907	20	attP2	nd	sterile (atrophied ovaries)
eIF4AIII	38202	22	attP40	nd	sterile (atrophied ovaries)
enc	42797	22	attP2	OFF	fertile
enok	40917	20	attP2	OFF	fertile
enok	41664	20	attP2	OFF	fertile
enok	42941	20	attP40	OFF	fertile
Fmr1	35200	20	attP2	OFF	fertile
fs(1)Yb	35181	22	attP2	OFF	fertile
fs(1)Yb	35301	22	attP2	OFF	fertile
G9a	34817	20	attP2	OFF	fertile
hay	53345	20	attP40	nd	sterile (atrophied ovaries)
Hel25E/ uap56	33666	20	attP2	nd	sterile (atrophied ovaries)
Hrb27C	33716	20	attP2	OFF	sterile
Hrb87F	52937	20	attP40	OFF	fertile
Hrb98DE	32351	20	attP2	OFF	sterile (atrophied ovaries)
Hsc70-1	34527	20	attP2	OFF	fertile
Hsc70-2	42014	20	attP40	OFF	fertile
Hsc70-2	44485	20	attP2	OFF	fertile
Hsc70- 2,Hsp70Ba,Hsp70Bb,Hsp70Bbb,Hsp70Bc	32997	20	attP2	OFF	fertile
Hsc70-3	32402	20	attP2	OFF	sterile
Hsc70-4	34836	20	attP2	nd	sterile (atrophied ovaries)
Hsc70-4	35684	21	attP2	nd	sterile (atrophied ovaries)
Hsf	41581	22	attP2	OFF	nt
Hsp23	44029	20	attP40	OFF	fertile
Hsp26	35408	22	attP2	OFF	fertile
Hsp26	42610	20	attP40	OFF	fertile
Hsp27	33007	20	attP2	OFF	fertile
Hsp27	33922	20	attP2	OFF	fertile
Hsp67Ba	41962	20	attP2	OFF	fertile
Hsp67Bb,Hsp22	41709	20	attP40	OFF	fertile
Hsp67Bb,Hsp22	51397	22	attP2	OFF	fertile
Hsp67Bc	35452	22	attP2	OFF	fertile
Hsp67Bc	42607	20	attP40	OFF	fertile
Hsp68	50637	20	attP2	OFF	fertile
Hsp83	32996	20	attP2	OFF	sterile
Iswi	32845	20	attP2	OFF	nt
Iswi	51931	22	attP40	OFF	fertile
kis	34908	20	attP2	OFF	fertile

kis	36597	22	attP40	OFF	sterile
kis	44542	20	attP40	OFF	sterile
krimper	35230	22	attP2	OFF	sterile
krimper	35230	22	attP2	OFF	sterile
krimper	35231	22	attP2	OFF	sterile
krimper	37511	20	attP40	OFF	sterile
kto	34588	20	attP2	OFF	sterile
l(3)72Ab	34024	20	attP2	nd	sterile (atrophied ovaries)
l(3)72Ab	50716	20	attP40	nd	sterile (atrophied ovaries)
loqs	32955	20	attP2	OFF	fertile
loqs	33427	20	attP2	OFF	fertile
loqs	34779	20	attP2	OFF	fertile
loqs	34780	20	attP2	OFF	fertile
loqs	34781	20	attP2	OFF	fertile
loqs	34782	20	attP2	OFF	fertile
loqs	34851	20	attP2	OFF	fertile
mago	35453	22	attP2	OFF	fertile
Me31b	33675	20	attP2	OFF	sterile (atrophied ovaries)
Me31b	38923	22	attP40	OFF	fertile
metro	35810	22	attP2	OFF	fertile
Mi-2	33419	20	attP2	nd	sterile (atrophied ovaries)
Mi-2	35398	22	attP2	nd	sterile (atrophied ovaries)
Mi-2	51774	20	attP40	nd	sterile (atrophied ovaries)
mle	34864	20	attP2	OFF	nt
mor	34919	20	attP2	nt	nt
mor	35630	22	attP40	OFF	sterile
mor	35662	21	attP2	OFF	sterile
mub	34870	20	attP2	OFF	fertile
nej	36682	20	attP2	OFF	fertile
nej	37489	20	attP2	OFF	sterile
Nurf-38	35444	22	attP2	OFF	sterile
osa	35447	22	attP2	OFF	fertile
osa	38285	20	attP40	OFF	fertile
otu	34065	20	attP2	OFF	sterile (atrophied ovaries)
pacman	34690	20	attP2	OFF	fertile
papi	34932	20	attP2	OFF	fertile
papi	35450	22	attP2	OFF	fertile
papi	37513	20	attP40	OFF	fertile
papi	38216	22	attP40	OFF	fertile
pasha	33972	20	attP2	OFF	nt
Pc	33622	20	attP2	OFF	fertile
Pc	33964	20	attP2	OFF	fertile
Pc	36070	22	attP40	OFF	fertile
Pcaf	33981	20	attP2	OFF	fertile
Pcaf	35601	22	attP40	OFF	fertile
Pcl	33945	20	attP2	OFF	fertile
Pcl	33946	20	attP2	OFF	fertile
pea	32838	20	attP2	nd	sterile (atrophied ovaries)
ph-p	33669	20	attP2	OFF	fertile
pps	38529	20	attP2	OFF	fertile
pps	38912	22	attP40	OFF	fertile
Psi	34825	20	attP2	OFF	fertile
<b>pUf68</b>	<b>34785</b>	<b>20</b>	<b>attP2</b>	<b>ON</b>	<b>fertile</b>
qin	37475	22	attP2	OFF	fertile

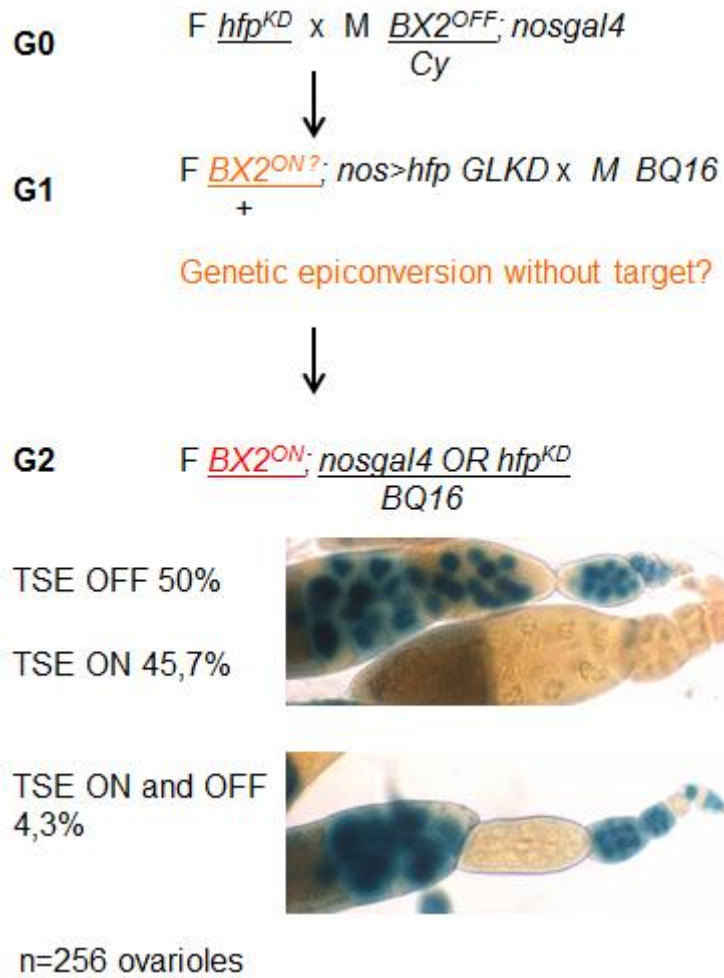
qin	41662	22	attP2	OFF	nt
r2d2	34784	20	attP2	OFF	fertile
Rab5	34832	20	attP2	OFF	sterile
Rab5	51847	20	attP40	OFF	fertile
<b>Rm62</b>	<b>34829</b>	<b>20</b>	<b>attP2</b>	<b>ON</b>	<b>fertile</b>
Rox8	32472	20	attP2	OFF	fertile
rpd3	33725	20	attP2	OFF	sterile
rpd3	34846	20	attP2	OFF	fertile
Rrp1	35420	22	attP2	OFF	fertile
rrp6	34809	20	attP2	lethal	-
rrp6	42064	20	attP40	OFF	sterile
Saf-B	51759	20	attP2	OFF	sterile
Sbf	32419	20	attP2	OFF	fertile
Sbf	44004	20	attP40	OFF	fertile
SF2	32367	20	attP2	OFF	fertile
shutdown	35454	22	attP2	OFF	sterile
skd	34630	20	attP2	OFF	fertile
smaug	35477	22	attP2	OFF	fertile
Smn	36621	22	attP2	OFF	fertile
snf	34593	20	attP2	OFF	fertile
snf	51459	20	attP40	nd	sterile (atrophied ovaries)
Snr1	32372	20	attP2	OFF	sterile
snRNP-U1-70K	33396	20	attP2	nd	sterile (atrophied ovaries)
snRNP-U1-C	34822	20	attP2	nd	sterile (atrophied ovaries)
Spf45	41954	20	attP2	OFF	fertile
spn-E / homeless	35303	22	attP2	OFF	sterile
Spt-l	55685	20	attP40	OFF	fertile
Spt3	35148	22	attP2	OFF	fertile
spt4	32896	20	attP2	OFF	sterile (atrophied ovaries)
Spt5	34837	20	attP2	nd	sterile (atrophied ovaries)
Spt6	32373	20	attP2	nd	sterile (atrophied ovaries)
sqd	35627	22	attP2	OFF	sterile
sqd	53891	20	attP40	OFF	fertile
stauen	35690	21	attP2	OFF	fertile
stauen	43187	22	attP40	OFF	sterile
su(sable)	33982	20	attP2	OFF	fertile
Su(var)205	33400	20	attP2	nd	sterile (atrophied ovaries)
Su(var)205	36792	22	attP40	nd	sterile (atrophied ovaries)
Sxl	34393	20	attP2	OFF	sterile (atrophied ovaries)
Sxl	38195	22	attP40	OFF	sterile (atrophied ovaries)
Taf1	32421	20	attP2	nd	sterile (atrophied ovaries)
Taf1	35314	22	attP2	nd	sterile (atrophied ovaries)
Taf10	35239	22	attP2	OFF	fertile
Taf12	34852	20	attP2	OFF	fertile
Taf4	35427	22	attP2	OFF	sterile
Taf4	50985	20	attP40	nd	sterile (atrophied ovaries)
Taf5	35367	22	attP2	nd	sterile (atrophied ovaries)
Taf7	55216	22	attP2	OFF	fertile
tej	36879	22	attP2	OFF	sterile
tej	41929	20	attP40	OFF	sterile
tra2	28018	10	attP2	OFF	fertile
trx	33703	20	attP2	OFF	fertile
tsunagi	36585	22	attP2	OFF	sterile (atrophied ovaries)
tudor	42800	22	attP2	OFF	fertile

Tudor-SN	34865	20	attP2	OFF	nt
U2af38	50561	22	attP2	nd	sterile (atrophied ovaries)
U2af50	50521	22	attP2	OFF	sterile
Upf1	43144	22	attP2	OFF	sterile
Upf2	31095	1	attP2	OFF	fertile
Upf3	44565	20	attP40	OFF	nt
vasa	32434	20	attP2	OFF	sterile
vasa intronic gene	35184	22	attP2	OFF	fertile
vasa intronic gene	35183	22	attP2	OFF	sterile

**Figure S1.**

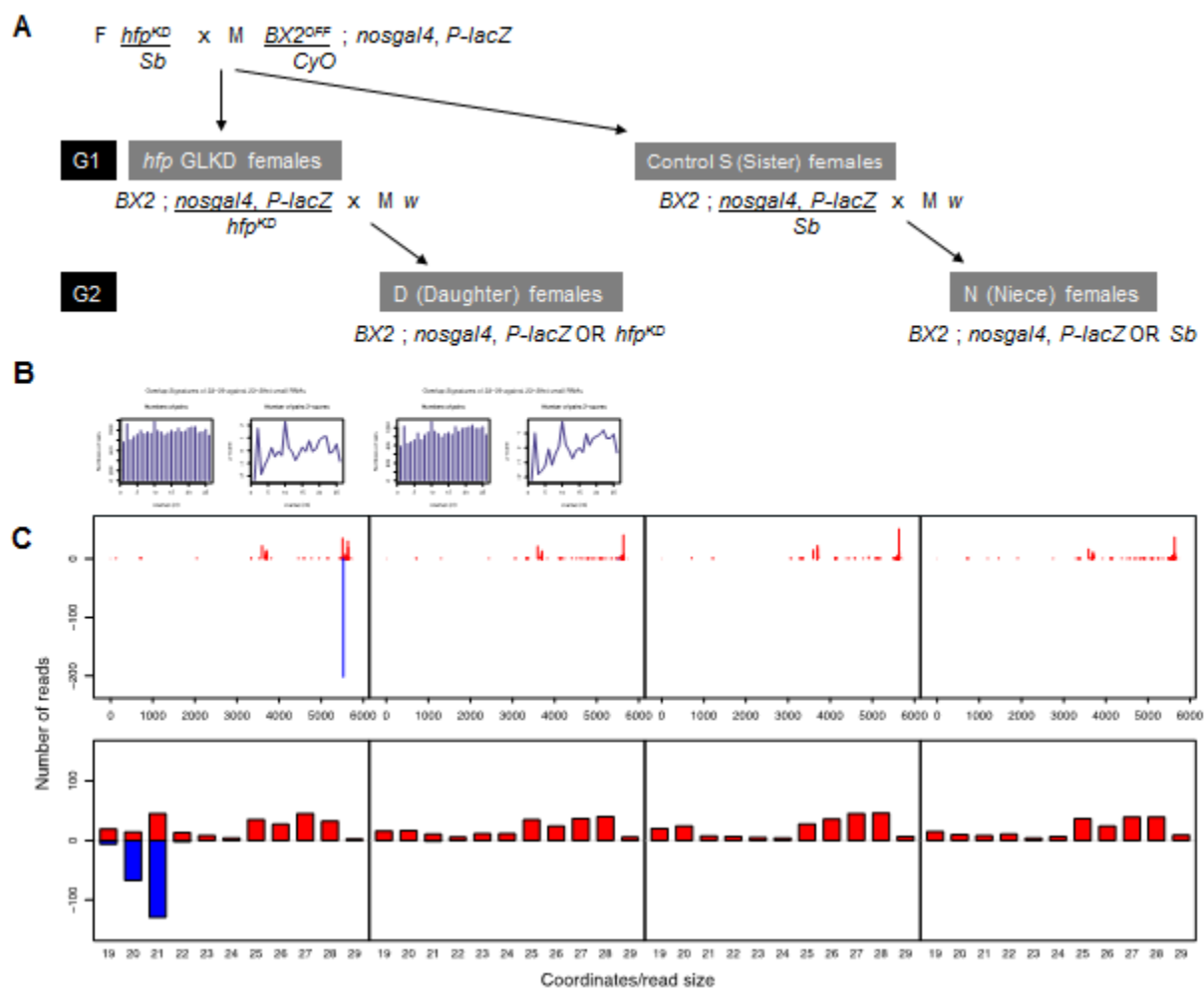


**Figure S2.**





**Figure S3.**

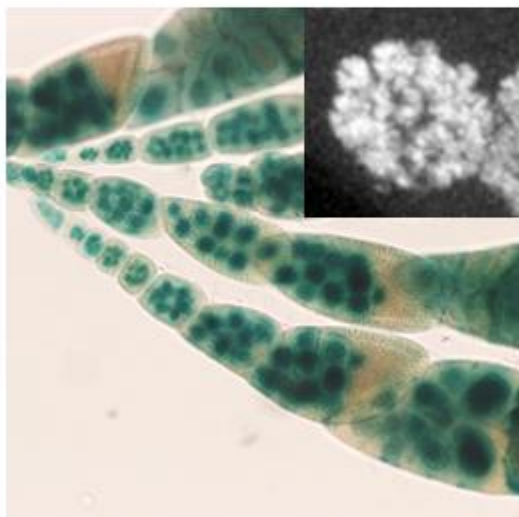


**Figure S4.**

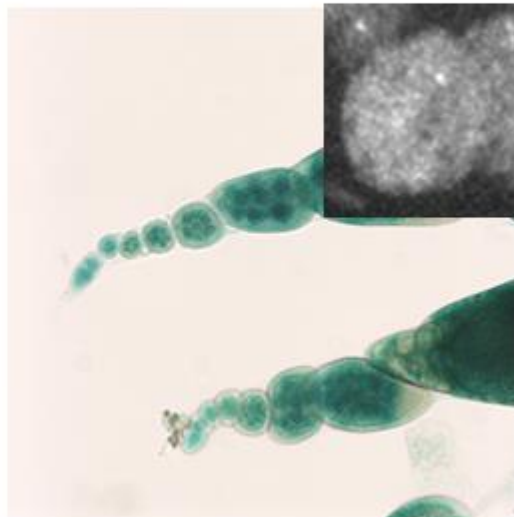
$otu^{11} / M5 ; BQ16$  x  $otu^{11(7)} ; BX2 / Cy$

$otu^{11/11} ; BX2^{pat} ; BQ16$

$otu^{11/+} ; BX2^{pat} ; BQ16$



TSE = 0% (n=565)



TSE = 0% (n= 1720)

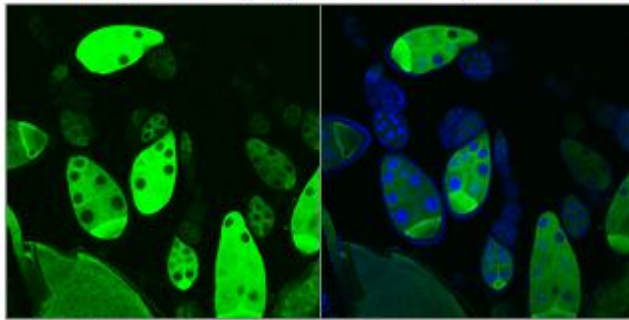


## Résultats additionnels

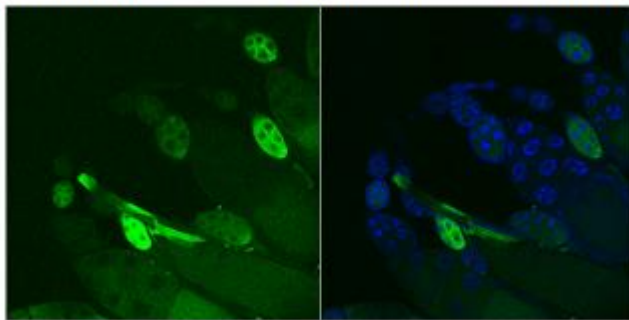
Si il est clair que *hfp*<sup>GLKD</sup> est capable d'induire la production de piARN à partir d'un locus *BX2* hérité paternellement, des questions demeurent : par quels mécanismes *hfp* agit-il sur *BX2* ? Et à partir de quand au cours du développement, le locus *BX2* produit-il des piARN fonctionnels en contexte *hfp*<sup>GLKD</sup> ?

Des expériences de RNA seq sont en cours pour essayer de répondre à la première question. Pour répondre à la seconde, j'ai réalisé l'expérience suivante : j'ai utilisé des femelles portant le shARN de *hfp* que j'ai croisé avec des males portants le cluster *BX2* inactivé et le système de cible. Les ovaires adultes présentent un grand nombre de chambres réprimées pour la GFP (Fig. 48) comme pour la cible *PlacZ*<sup>BQ16</sup> dans l'article. Ce résultat confirme, avec une autre cible, l'effet de la perte de fonction de *hfp* sur la conversion épigénétique du cluster *BX2* inerte en cluster producteur de piARN.

*shARN (w) x BX2<sup>OFF</sup>/cy ; nosGal4<sup>32180</sup>, UASpGFP<sup>7373</sup>*



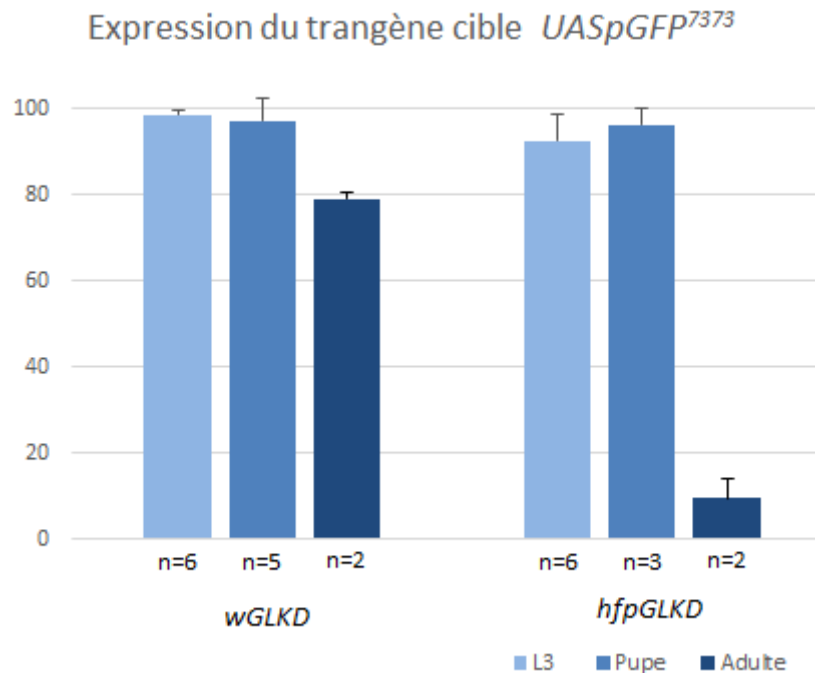
*shARN (hfp) x BX2<sup>OFF</sup>/cy ; nosGal4<sup>32180</sup>, UASpGFP<sup>7373</sup>*



**Figure 48 - Activation du locus producteur de piARN *BX2* dans le contexte perte de fonction pour *hfp*.** Le nombre de chambres réprimées (GFP-) augmente en l'absence de la protéine Hfp par rapport au croisement contrôle avec le shARN du gène *white*.

En revanche, les gonades de larves et de pupes issues de croisements *shARN (hfp) x BX2<sup>OFF</sup>/cy ; nosGal4<sup>32180</sup>, UASpGFP<sup>7373</sup>* qui possèdent donc le cluster inactif, hérité paternellement, en contexte *hfp*<sup>GLKD</sup> ne montrent aucune répression significative du transgène cible *UASpGFP<sup>7373</sup>*. Ainsi, j'ai pu montrer que l'inactivation de *hfp* n'a pas d'effet sur la

répression d'un transgène cible *UASpGFP*<sup>7373</sup> avant la métamorphose. Les résultats sont regroupés dans la figure suivante.



**Figure 49– Effet du *hfp*<sup>GLKD</sup> au cours du développement.** Comparaison des stades L3, Pupe et adulte pour la répression du transgène cible *UASpGFP* par le silencer *BX2* apporté par voie mâle. Pourcentage de PGC/gonade exprimant la GFP pour les stades L3 et Pupe et de chambres ovariennes exprimant la GFP pour le stade adulte. On voit l'activation du cluster de piARN *BX2* dans le mutant *hfp*<sup>GLKD</sup> uniquement au stade adulte avec la baisse du pourcentage de chambres ovariennes exprimant le transgène *UASpGFP*.

Ceci permet de constater que le gène *hfp* est requis pour le maintien du statut inactif de *BX2* à partir du début de l'ovogenèse. Ces changements sont maintenus au fil des générations, malgré le retour à une dose normale de *hfp* dès la génération suivante. La perte de fonction d'un gène endogène permet donc une conversion épigénétique stable. Ceci démontre que l'état inactif d'un cluster producteur de piARN n'est pas un état «par défaut», mais un choix maintenu au niveau de la cellule et stable à travers les générations. De plus, ceci ouvre une piste de réflexion sur la différence entre le mécanisme de répression par les piARN durant le développement de la lignée germinale et durant l'ovogenèse.

## Conclusions

Cette levée de l'état inactif du cluster *BX2* en contexte mutant *hfp* suggère qu'il existe un ou des mécanismes assurant le maintien de l'état non producteur de piARN de *BX2*. Au cours de ma thèse, j'ai pu montrer que certains facteurs directement reliés à la voie de biosynthèse des piARN sont absolument nécessaires durant le développement. L'implication de *hfp* ajoute un niveau supplémentaire de régulation, cette fois au niveau des clusters. Il semblerait qu'un mécanisme encore inconnu bloque le changement d'état des séquences productrices de piARN (en tout cas, celui du *locus BX2*) pendant le développement. La protéine Hfp/Puf68 prendrait ensuite le relais au moment de l'ovogenèse et son absence entraîne l'activation de *BX2* uniquement au stade adulte. Pour l'instant, on ne connaît que la Piwi-Less Pocket (Fig. 33), cette fenêtre de dérégulation modérée des ET, localisée dans le germarium, comme endroit/moment où la variabilité entraînée par la mobilisation des ET est autorisée. Il est probable que la frontière entre les deux types de régulation pour l'inaptitude des séquences répétées à faire des piARN soit également localisée après les cellules souches germinales adultes, au niveau des cystes germinaux à 2-16 cellules.



### III – Discussion et Perspectives

#### A. Les piARN larvaires différents des piARN adultes ?

##### A.1 Relâcher la pression ?

La « piwi-less pocket » observée dans le germarium de l'adulte semble montrer qu'une certaine latitude existe dans la répression des éléments transposables au niveau de la lignée germinale. Nous voulions voir si de telles fenêtres de dérégulation des ET étaient présentes au cours du développement de la lignée germinale. D'après mes résultats, il semble que ce ne soit pas le cas. Le silencer *RS3* ne montre pas de faiblesse de silencing au cours du développement à l'exception des cystes germinaux à 2-16 cellules de la PILP (Dufourt et al, 2014).

Les résultats sur le gène *hfp* obtenus par A. Boivin suggèrent que certains gènes peuvent être impliqués dans la conservation de l'identité des séquences répétées, qu'elles soient productrices ou non-productrices de piARN. *hfp* maintient le cluster *BX2* dans un état non producteur de piARN, mais l'abolition de cette répression en l'absence de *hfp*, entraîne un changement stable des propriétés épigénétiques du *locus*. Néanmoins, d'après mes résultats, la mutation d'*hfp* ne rend pas *BX2* actif avant le germarium. Un verrou extrêmement robuste semble être mis sur le développement, à la fois au niveau des clusters et des éléments transposables.

##### A.2 Utilisation du crible shARN pour étudier l'effet des mutations de gènes de la voie des piARN

L'étude de la production des piARN n'est pas toujours évidente car la mutation de gènes de la voie des piARN donne souvent lieu à des individus stériles aux ovaires atrophiés. Ceci empêche le suivi des mutants sur plusieurs générations et donc l'analyse d'éventuels effets épigénétiques médiés par les piARN. Durant ma thèse, j'ai pu montrer que la répression par les piARN fonctionne durant le développement de la lignée germinale depuis l'embryogenèse. Le crible de gènes candidats que j'ai mené ensuite en L3 a vérifié qu'un certain nombre de protéines partenaires de la voie des piARN en larve sont les mêmes qu'en adulte : la répression d'un transgène cible est abolie dès la larve en l'absence des principaux



acteurs de la voie des piARN. Un point intéressant est que l'ARN interférence médiée par les shARN ne change ni la morphologie des cellules germinales larvaires ni leur quantité. Ceci peut être attribué à l'efficacité moindre du shARN comparé à un mutant nul. J'ai pu observer la disparition de la protéine ciblée par les shARN dans les PGC de plusieurs lignées grâce à des immunomarquages, mais il est possible que des niveaux résiduels de protéine persistent et qu'ils ne soient pas détectables par cette méthode. Une autre explication de ce phénotype d'ovaires est l'implication des cellules folliculaires dans la forme de l'ovaire. En effet, certains partenaires de la voie des piARN sont nécessaires aussi dans les cellules somatiques de l'ovaire pour son bon développement. Par exemple, la disparition de Piwi uniquement dans les cellules folliculaire est suffisante pour induire une stérilité complète (Rozhkov, Hammell, et Hannon 2013) et l'absence de production de piARN de *flamenco* dans les cellules folliculaires empêche leur action de protection de la lignée germinale (Mével-Ninio et al. 2007). Ces cellules participent donc à façonner la structure de l'ovaire. Contrairement aux PGC, elles subissent de nombreuses mitoses et se différencient durant la vie larvaire (Chap. 4 A.2.3). La mutation de gènes de la voie des piARN dans les cellules folliculaires a donc un fort impact sur la forme de l'ovaire adulte, tandis que les PGC commencent à être vraiment actives durant l'ovogenèse quand elles se différencient. Au stade adulte en revanche, les ovaires des mouches contenant les shARN sont plus rarement atrophiés même si plusieurs lignées montrent de la stérilité (14/28). L'interférence ARN localisée uniquement dans les PGC grâce à des pilotes *gal4*, alliée à un système rapporteur pour la répression par les piARN, pourrait permettre de faire des recherches sur les gènes qui donnent des ovaires atrophiés. L'observation de la répression dans les gonades larvaires serait un bon indicateur de leur influence sur le TSE.

D'autre part, nous avons observé l'effet de ces mutations sur la répression médiée par les piARN du *TAS 3R* mais il est possible d'utiliser de cette façon tous les transgènes silenciers localisés sur différents chromosomes que nous possédons. On pourrait alors comparer les clusters de piARN. L'étude du silencer *P-1152* (localisé sur le chromosome X) et du cluster de transgène répresseur *BX2\** (localisé dans l'euchromatine) permettrait de mieux comprendre les différents mécanismes de répression. Il est probable que des différences existent entre les *locus* producteurs de piARN, certaines sont d'ailleurs déjà connues: clusters unidirectionnels / double brins, tandis que d'autres sont en cours d'étude: clusters télomériques / centromériques.

L'utilisation d'autres transgènes pilotes *gal4* est également possible, pour diriger l'expression du shARN et de la cible dans le tissu étudié. Le pilote *bamGal4* permettrait d'induire l'expression d'un shARN précisément dans les cystes germinaux à 2-16 cellules de l'ovariole pour visualiser l'effet d'un KD durant la différenciation des cellules germinales chez l'adulte. Au stade adulte également, on sait que le pilote *nosGal4* ne s'exprime qu'au niveau du germarium et des stades tardifs de l'ovariole, la combinaison de plusieurs transgènes *gal4* permettrait une activation du shARN plus complète lors de l'ovogenèse comme dans la lignée 31777 qui en contient trois (Ni et al. 2011).

Enfin, tester l'effet de l'absence de plusieurs protéines de la voie des piARN en même temps grâce à ce système n'est pas toujours possible. Certains partenaires des piARN ont des fonctions redondantes ou complémentaires (Tej et Tap par exemple) et l'étude de doubles mutants pourrait être intéressante. Malheureusement, la majorité des lignées shARN ont un site d'insertion fixe du transgène shARN sur le chromosome 3.

### **A.3 Quelle différence entre les transgènes silenciers *RS3* et *P-1152* ?**

Une question centrale qui se pose encore est la différence entre les silenciers *RS3* et *P-1152*. Sachant que l'homologie de séquence qu'ils partagent avec le transgène cible est la même (environ 500pb Fig. 1 Article 1), pourquoi n'induisent-ils pas un pourcentage de répression identique ? Nous nous sommes donc intéressés à la variabilité induite par le silencer *P-1152*.

La première hypothèse serait une quantité différente de piARN produite par les deux silenciers. Nous avons analysé les piARN du 5' de l'élément *P* produits par les deux transgènes silenciers. Il semble que *RS3* ne fasse pas plus de piARN de cette séquence que *P-1152* en adulte, voire moins. Des séquençages de piARN de larves L3 sont en cours pour voir si c'est également le cas en larve.

La seconde hypothèse serait une nature différente des piARN produits par les deux silenciers. Par exemple, le patron de découpage des précurseurs provenant de *RS3* pourrait produire des piARN plus efficaces. Ce point est difficile à visualiser à cette échelle mais une analyse plus fine avec des piARN uniques pourrait apporter des réponses à cette question.

Les transgènes *RS3* et *P-1152* et le transcrite cible *UASpGFP<sup>7373</sup>* partagent 500 pb d'homologie, un morceau du 5' de l'élément *P*. Néanmoins, en amont du promoteur de

l'*UASpGFP*<sup>7373</sup> on trouve des séquences du gène *white* (Fig. 1 Article). Il est possible que les piARN de *white* provenant de *RS3* entraînent la répression transcriptionnelle du transgène cible de façon plus efficace grâce à un phénomène d'hétérochromatinisation. Néanmoins, cette hypothèse semble peu probable car elle n'explique pas pourquoi les transgènes cibles *PlacZ* qui partagent plus d'homologie avec le silencer *P-1152* (qui contient des séquences de *lacZ*) présentent une expression variée alors qu'ils sont totalement réprimés par *RS3*. Cela va dans le sens que la variabilité est une propriété intrinsèque à *P-1152*.

Une dernière idée implique que la localisation des différents *TAS* (chromosomes *X* et *3R*) pourrait avoir un rôle dans leur efficacité. On peut imaginer que le regroupement de certains *TAS* dans un compartiment spécifique du noyau ait une incidence sur la production de piARN (Voir hypothèse 1) ou pour la répression (hypothèse 2). Par exemple, les protéines du groupe Polycomb (PcG) sont présentes sous forme de foyers dans le noyau. Ces protéines permettent de maintenir un état répressif des gènes (marques H3K27me3) en compactant la chromatine (Bantignies et Cavalli 2011). Or les *TAS* possèdent certains motifs pouvant être reconnus par la famille PcG (Boivin et al. 2003). Il a été montré que la marque répressive H3K27me3 était présente sur les *TAS* par des expériences de FISH sur chromosomes polytènes (Andreyeva et al. 2005). Il est possible que les *TAS* des chromosomes *X* et *3* soient localisés dans deux foyers distincts. Cela dit, il semblerait que l'effet des PcG soit limité aux cellules somatiques car des lignées GLKD pour *Pc*, *sce*, *ph*, *E(z)*, *Pcl* et *Rpd3* ne montrent pas de diminution de la répression par les silencers *P-1152* et *BX2*\*

Enfin, les résultats que j'ai obtenus sur l'implication de HP1a dans la production de piARN télomériques amènent une autre hypothèse que je développerai dans la partie suivante.

## B. Le rôle des protéines Rhino et HP1a dans la production de piARN

### B.1 Rôle de Rhino et de HP1a dans la transcription des clusters de piARN

L'analyse de Deep-seq des petits ARN provenant de corps gras et de gonades de descendants femelles (L3) *rhino*<sup>GLKD</sup> montre une forte baisse de la présence de piARN par rapport au contrôle *w*<sup>GLKD</sup>. On note la perte d'environ 50% des piARN correspondant au *locus* *42AB* et une baisse de 80% des piARN du *TAS 3R* (et de *RS3*). Rhino affecte donc la production de piARN télomériques et péricentromériques provenant de dual-strand clusters ce qui est cohérent avec le modèle actuel selon lequel cette protéine est nécessaire à leur transcription et constituerait le premier maillon de la chaîne d'évènement permettant la transformation de leurs transcrits en piARN (Klattenhoff et al. 2009; Mohn et al. 2014). Les résultats suivants suggèrent qu'une étape antérieure a lieu au niveau de la chromatine des clusters télomériques.

Les résultats obtenus avec le shARN de HP1a (*Su(var)205*<sup>GLKD</sup>) sont plus surprenants car ils ne sont pas homogènes pour les deux types de dual-strand clusters que nous étudions. Dans le crible shARN, la mutation des gènes *H2Av* et *Su(var)205* montrent que l'hétérochromatinisation des *loci* producteurs de piARN est essentielle à la transcription du cluster télomérique du *3R*. Le variant de l'histone H2A, H2Av, semble nécessaire pour attirer HP1a sur l'ADN. En effet, les mutants *H2Av* présentent une diminution de la marque H4K12ac, une baisse de H3K9me3 et une réduction du recrutement de HP1a centromérique par immunomarquage sur chromosomes polytènes (Swaminathan, Baxter, et Corces 2005), plaçant ce variant d'histone en amont de la formation de l'hétérochromatine. Comme H2Av semble présent aux télomères (Fig. 17), on peut penser que ce mécanisme décrit pour l'hétérochromatine centromérique est également valable pour les régions sub-télomériques et donc pour les clusters de piARN télomériques. Ceci pourrait suggérer que H2Av soit requis pour la production de piARN télomériques expliquant ainsi la perte de la répression du transgène *UASpGFP* par le silencer *RS3* dans les gonades larvaires. Des analyses de Deep-seq de *H2Av*<sup>GLKD</sup> permettraient de répondre à ces hypothèses.

D'autre part, nos résultats montrent que HP1a est impliquée dans la production de piARN à partir du *locus* télomérique *TAS-3R* dans lequel est inséré le transgène silencer *RS3*. L'analyse de Deep-seq des petits ARN des descendants issus du croisement de femelles *RS3 ; nosGal4* par des mâles *sh(Su(var)205)* montre que la production de piARN

téломériques dépend de la présence de HP1a. Plus de 70 % des piARN du *TAS 3R* et de *RS3* sont perdus dans *Su(var)205<sup>GLKD</sup>* alors que la transcription du *42AB* est inchangée (moins de 15% de perte des piARN homologues).

Jusqu'à présent, dans la voie des piARN adulte, HP1a a surtout été impliquée *via* son rôle répressif, également au niveau des *loci* producteurs de piARN, avec les expériences de Yin et Lin en 2007. La présence d'HP1 sur les TAS, mesurée par ChIP-seq, augmente en condition mutante pour *piwi* (Yin et Lin 2007). La perte de Qin entraîne également une plus forte présence de HP1a dans les stades précoces, spécifiquement sur les dual-strand clusters (Anand et Kai 2012), l'hypothèse que Qin agirait en empêchant HP1a de s'y fixer pour faciliter leur transcription a été évoquée par Anand et Kai en 2012. Rôle répressif également au niveau des *loci* cible, où Piwi dirigerait l'hétérochromatinisation des séquences *via* un mécanisme qui implique l'accumulation de HP1a et leur silencing transcriptionnel (S. H. Wang et Elgin 2011; Sienski et al. 2015). HP1a et plusieurs gènes dans la voie piARN ont été impliqués dans l'effet de variévation lié à la diffusion de l'hétérochromatine depuis des régions centromériques et téломériques (S. H. Wang et Elgin 2011; Sienski et al. 2015) sur des gènes euchromatiques, le PEV (Swaminathan et al, 2005). En parallèle, la mutation de Rhino ne produit pas de changements dans l'organisation de l'hétérochromatine en modifiant l'expression des gènes au cours de l'ovogenèse (Klattenhoff et al. 2009), l'effet de HP1a sur la production des piARN est donc probablement spécifique aux TAS et sûrement en amont de Rhino.

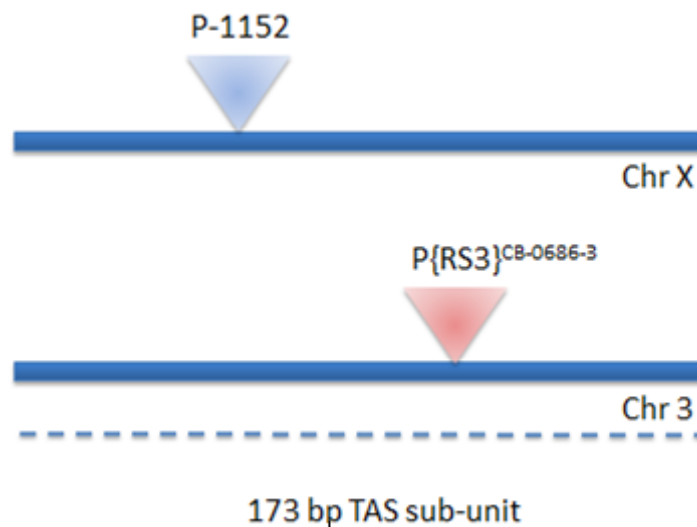
Ces résultats confortent les observations du rôle de HP1a dans la répression par les piARN (Ronsseray et al. 1996) et la disparition des piARN en contexte *Su(var)205/+* (Todeschini et al. 2010) montrant le rôle « positif » de HP1a dans la transcription des clusters de piARN.

## **B.2 Modèle proposé pour la variévation**

Le rôle de modificateur de la chromatine de HP1a pourrait influencer l'expression de transgènes insérés dans les régions sub-téломériques en affectant les niveaux de production piARN à des degrés divers, en fonction de la position ou l'environnement des transgènes. Dans le PEV, la présence de HP1a induit de la répression transcriptionnelle de gènes euchromatiques relocalisés vers des régions hétérochromatiques. On peut imaginer un modèle dans lequel la présence de HP1a pourrait fluctuer au niveau des bornes de l'hétérochromatine,

mimant ce phénomène. La protéine aurait alors un tout autre rôle car nos résultats montrent que HP1a favorise la production des piARN télomériques.

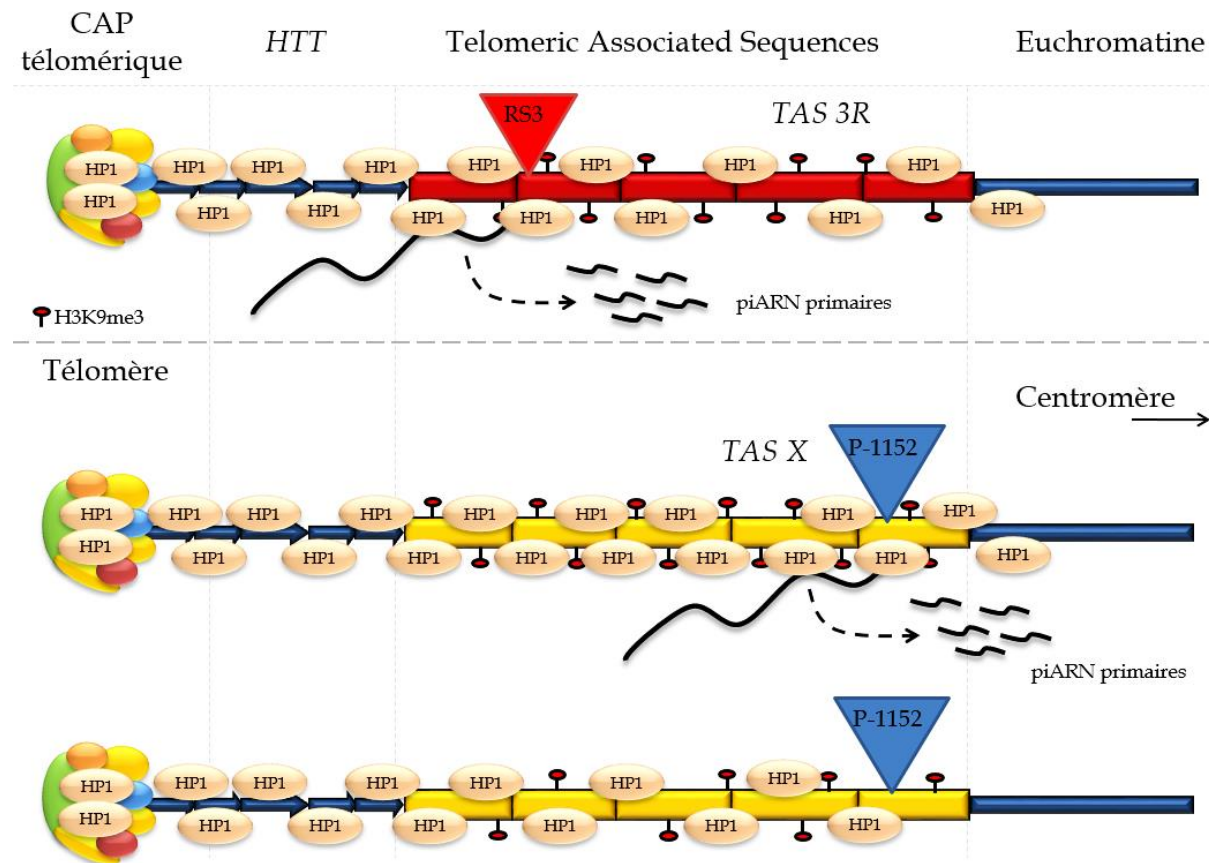
Si HP1a est particulièrement présente à l'extrémité du chromosome (Perrini et al. 2004), on sait qu'elle est également associée aux séquences *HTT* et *TAS* (Fanti et al. 1998; Frydrychova et al. 2008; Klattenhoff et al. 2009). On connaît la position des transgènes silenciers *RS3* (Ryder et al. 2004) et *P-1152* (Josse et al. 2008) au sein d'une unité *TAS* (Fig. 50), mais ceux-ci sont insérés dans deux chromosomes différents. De plus, les *TAS* sont constituées de plusieurs sous-unités (*TAS A* à *D*) et la position relative de la sous-unité dans laquelle est inséré chaque transgène silencer n'est pas connue.



**Figure 50 – Position respective des transgènes silenciers *RS3* et *P-1152* au sein d'une unité *TAS*.**

On peut proposer que *P-1152* soit dans un environnement où la présence de HP1a fluctue alors que *RS3* pourrait être inséré dans un locus où la quantité de HP1a varie peu. Par exemple, *P-1152* pourrait être proche de la frontière de l'hétérochromatine. En émettant l'hypothèse que l'activation de la production de piARN par *P-1152* est différente d'une cellule à l'autre, cela pourrait expliquer le phénomène de variéation, où une cible euchromatique s'exprime de façon ON-OFF dans les cellules germinales souches (PGC, GSC) et les chambres à œuf qui en descendent. Cet effet, couplé à l'implication de HP1a dans la production de piARN télomériques pourrait donc expliquer que le transgène silencer *P-1152* ne réprime pas la cible dans toutes les cellules alors que *RS3* est activé à chaque fois (Fig. 51).

Il est également possible que la quantité de HP1a sur les *TAS* dépende de la quantité de piARN capturée par les PGC lors de la cellularisation. Certaines PGC des embryons de la lignée *P-1152* auraient donc l'extrémité du *TAS X* recouverte de HP1a dès le début du développement tandis que d'autres en auraient une quantité moindre.



**Figure 51 - Modèle basé sur la présence du transgène silencer *P-1152* à l'extrémité des séquences *TAS*.** La concentration de HP1a à un *locus* influence sa transcription et le devenir de ses transcrits. (En haut) *RS3* pourrait être situé proche du télomère dans le *TAS 3R* tandis que (En bas) *P-1152* localisé au *TAS X*, pourrait être proche de l'euchromatine et donc sujet aux fluctuations de la présence de HP1a cellule par cellule.

### B.3 Éviter les piARN

Pourquoi, malgré la présence supposée de piARN dans toutes les cellules des stades tardifs de l'ovogenèse (en tout cas suffisamment pour que 100% des descendants présentent de la répression), la cible reste exprimée dans certaines cellules? En effet, on observe de la variabilité jusqu'au stade adulte mais les individus femelles issus de chambres exprimant la cible présentent autant de répression que ceux issus de chambres réprimant la cible (~50%). Ceci suggère que la présence tardive de piARN ne suffit pas à réprimer une cible pour laquelle la décision d'expression a été prise précocement. L'activation retardée du transgène

silencer *P-1152* aurait donc un effet seulement sur certaines cibles dont le statut transcriptionnel serait décidé très tôt et ensuite maintenu au cours du développement. Le transgène silencer finirait par être activé dans toutes les cellules germinales, HP1a pouvant par exemple s'accumuler sur le TAS durant le développement, et les piARN produits activeraient le *locus* silencer dans la descendance. Ce qui mène à une autre question: quels mécanismes/facteurs peuvent contrer l'action des piARN homologues ?

Cette question se rapproche des observations du laboratoire de M. Simonelig et de Z. Mourelatos sur l'embryogenèse. Des piARN liés à Aubergine entraînent la dégradation de transcrits de *nanos* dans la partie antérieure de l'embryon tout en permettant une accumulation de ceux-ci dans le pôle postérieur (Rouget et al. 2010). Cette étude a été étendue pour visualiser les autres ARNm concernés: des centaines d'ARNm interagissent avec Aub dans les deux premières heures du développement. Un tiers de ces ARNm subirait la MTZ tandis que le reste consiste en transcrits enrichis dans le germlasme: leur dégradation sélective permettrait leur localisation (Barckmann et al. 2015). Chez la Drosophile, les ARNm nécessaires à la mise en place des cellules germinales sont plus longs et plus abondants que d'autres ARNm. La coopération entre Aub et Tudor ainsi que des piARN partiellement homologues avec les ARNm qu'ils séquestrent serait responsable de leur accumulation au pôle postérieur (Vourekas et al. 2016). On peut imaginer un mécanisme où les transcrits de la cible seraient également « protégés » dans certaines cellules mais ceci n'explique pas comment ils sont traduits par la suite.

Il serait pourtant tentant d'imaginer un mécanisme alternatif impliquant que l'absence de piARN produits par *P-1152*, dans les cellules où le silencer n'est pas activé précocement, permettrait la mise en place d'un contexte chromatinien permissif sur les séquences cibles, par exemple juste après la cellularisation. Ce contexte chromatinien serait ensuite verrouillé durant le développement et les gènes transcrits et traduits en dépit de la présence de piARN homologues. Une réinitialisation des marques épigénétiques semble néanmoins probable dans les cystoblastes car c'est à cet endroit que l'on détecte un relâchement dans la répression des éléments transposables (Dufourt et al. 2014). Les résultats de l'article d'A. Boivin indiquent également que le gène qui maintient le *locus BX2* à l'état inactif n'est requis qu'à partir du germarium. Il n'y a pas d'incidence de la mutation d'*x* plus tôt au cours du développement. Enfin, le transgène cible UASpGFP<sup>7373</sup> est totalement réprimé dans l'ovariole adulte. Ceci rend probable un blocage transcriptionnel par des marques épigénétiques durant le



développement et la mise en place de la lignée germinale qui serait suivi d'une "relecture" au moment de la différenciation en cystoblaste. Par la suite, dans les chambres à œuf de l'ovariole, les piARN homologues répriment leurs cibles par PTGS.

## C. Perspectives : observer le mécanisme moléculaire de la Variégation

La variégation de la répression du transgène P-GFP que j'ai utilisé pendant ma thèse ou de transgènes P-LacZ observées précédemment est très intrigant. En tant que phénomène stochastique la question était de savoir si cette décision est maintenue ou re-décidée à chacune des mitoses de la vie des cellules germinale. Les premières tentatives que nous avons faites pour y répondre reposaient sur le suivi des mitoses *in vivo*. Nous avons utilisé des souches contenant une protéine de fusion pour suivre les cellules germinales (*vasaGFP*) et une cible fluorescente RFP. Nous avons disséqué des gonades de L2 pour tenter de filmer les mitoses et établir si les cellules filles d'une PGC étaient toujours dans le même état répressif que la cellule mère (hypothèse de la mémoire). Malheureusement, si les gonades disséquées restent en vie un moment (jusqu'à 8h), nous n'avons jamais été en mesure de distinguer une mitose. Soit les PGC ne se divisent plus après dissection des corps gras, le milieu de montage n'étant peut-être pas adapté, soit nous avons été incapables de remarquer ces événements cellulaires. Le marquage *vasaGFP* étant situé au niveau du nuage, la structure qui entoure le noyau, la déstabilisation de celui-ci durant la mitose rend l'observation difficile. Des expériences de lignage cellulaire ont permis d'éclaircir ce point durant ma thèse: j'ai pu montrer qu'une mémoire est maintenue à travers le développement de la lignée germinale (Dernière partie de l'article 1).

Cette mémoire est probablement associée à un état chromatinien différent du transgène cible selon les PGC/chambres à œufs. Dans les ovaires adultes, des expériences de ChIP-seq ont montré que la présence de piARN de *P-1152* conduit à une augmentation de la marque H3K9me3 et une diminution de la marque H3K4me sur le *locus* cible *BC69* qui n'a lieu que lorsqu'il y a héritage de piARN homologues par la mère (Le Thomas et al. 2013). Cette étude n'est pas assez résolutive pour répondre à la question, étant donné qu'on obtient un niveau global de l'occupation d'H3K9me3 sur le transgène *BC69* dans toutes les cellules de l'ovaire. Pour prouver que c'est bien une marque chromatinienne qui génère la variégation, il faudrait donc pouvoir observer la chromatine du *locus* cible cellule par cellule tout en visualisant sa transcription. Des ChIP-seq sur cellules après tri des populations Cible+ Cible – seraient l'idéal, mais la séparation des PGC de gonades larvaires est problématique. J'ai mis au point un protocole mais les rendements que j'ai obtenus sont très faibles pour le moment.

Pour l'observation de la variégation au stade adulte, je suis actuellement en train de fabriquer une lignée mutante pour *Bam* produisant des cancers de cellules germinales et qui

contiendra un traceur fluorescent pour les GSC couplé au système *P-1152/Cible* fluorescente. Si je réussis à récupérer suffisamment de matériel, ceci permettrait d'étudier la chromatine des *locus* Cibles et Silencer et de comparer la chromatine des GSC ON et OFF.

Dans les ovaires larvaires, une expérience alternative couplant FISH (localisation du *locus BC69*) et immunomarquage ( $\beta$ gal + différentes modifications d'histones) serait envisageable. On pourrait ainsi observer la colocalisation entre le transgène cible *BC69* et une augmentation de H3K4me dans les cellules exprimant la cible ou bien un éventuel remplacement de H2A par H2Av entraînant une forte concentration de la marque H3K9me3 et le recrutement de la protéine HP1a dans les cellules réprimant la cible.

Pour finir, l'étude de la variégation pourrait permettre de mieux comprendre la régulation des gènes. En effet, en dehors de leur rôle de répression d'ET, les piARN peuvent aussi réguler l'expression de gènes codants par des mécanismes variés. Cette régulation peut se faire aussi bien à un niveau transcriptionnel que post-transcriptionnel. La compréhension des mécanismes moléculaires de la régulation stricte des séquences productrices de piARN qui a lieu au cours du développement est un point essentiel pour appréhender leur fonctionnement.

## IV – Matériel et Méthodes

### A. Souches utilisées

Les mouches proviennent du Bloomington Drosophila Stock Center (#7373, #32180, #32249, #6420 et TRiP lines) et du Kyoto Drosophila Genomics and Genetic Resources (#123282). Elles ont été élevées à 25°C et les croisements ont également été faits à 25°C. Les chromosomes balanceurs portant les marqueurs phénotypiques dominants suivants ont été utilisés : [Cy] pour balancer le chromosome 2, [Sb] et [Ser] pour le chromosome 3 et [Xa] pour les chromosomes attachés 2-3.

Types de souches utilisés:

- Les souches dites « Silencer », possèdent un transgène inséré dans l'hétérochromatine sub-télomérique (TAS). Ce transgène produit des piARN capables d'éteindre l'expression d'un autre transgène possédant une séquence homologue.
  - La souche *P-1152* contient deux transgènes *PlacZ-rosy* au télomère du chromosome X, et produit des piARN de P, lacZ et rosy (de Vanssay et al. 2012).
  - La souche *RS3* contient un transgène P-FRT-white (FBms0003945) inséré au télomère du bras chromosomique *3R* au site 100E3. Cette souche est homozygote viable et fertile (Bloomington #123282), elle produit des piARN de P, w+ et de FRT (Dufourt et al. 2014).
  - La souche *BX2* contient 7 transgènes *P-lacW*(séquences *lacZ* et *white*) non télomériques insérés en tandem sur le bras droit du chromosome 2. Il ne produit pas de piARN homologues à ces transgènes.
  - La souche *T-1* contient les mêmes transgènes *P-lacW* que la souche *BX2*. Elle a été traitée aux rayons X et ce choc a modifié ses propriétés épigénétiques. C'est un répresseur fort producteur de piARN capables d'éteindre l'expression de transgènes cibles euchromatiques contenant des séquences P, LacZ ou white (Ronsseray et al. 2001).
  - La souche *BX2\**a été obtenue par le croisement de femelle *T-1* et de mâles *BX2* : les piARN de *T-1* ont été transférés par croisement aux descendants possédant le cluster *BX2* ce qui a activé la production de piARN à partir de ce cluster. C'est également un répresseur fort des transgènes cibles contenant des séquences P, LacZ ou white.

- Les souches portant un transgène dit « Cible ». Plusieurs types de cibles existent suivant le gène rapporteur utilisé et leur mode d'expression :
  - Les transgènes « Enhancer-trap », dont l'expression du gène rapporteur dépend des régions flanquant l'insertion. *BC69* contient un transgène enhancer trap *P-lacZ* *ry+* (*P{A92}*) constitué d'une fusion en phase le deuxième exon de la transposase de l'élément *P* et entre le gène *lacZ* d'*E. coli*. Il contient le marqueur *rosy* (FBtp0000154) et est inséré dans le gène endogène *vasa* : les femelle homozygotes sont stériles mais pas les mâles (Lemaitre et al, 1993).
  - Les transgènes dont l'expression du gène rapporteur dépend d'un promoteur spécifique qui dirige son expression comme la souche *vasaGFP*<sup>109171</sup>.
  - Les transgènes « UAS », pour lesquelles l'expression du rapporteur dépend de la présence de la protéine GAL4, elle-même pouvant être produite grâce à un transgène Enhancer-trap ou promoteur spécifique. Exemple *UASpGFP*<sup>7373</sup>.
- Le dernier type de souche contient un transgène portant le gène Gal4. La protéine GAL4 permet l'activation transcriptionnelle d'un transgène UAS. Sa séquence codante est fusionnée avec différents promoteurs tels que *nosGal4*<sup>32179</sup>, *nosGal4*<sup>32180</sup> ou *vasaGal4*.

**Pondoirs** : Récupération des embryons et des larves de premier stade (L1).

5 mâles x 20 femelles - Mise en tube préalable 12 à 24h pour fécondation.

Les pondoirs sont constitués d'une coupole en plastique reposant sur une boîte de Pétri démontable remplie d'un milieu de ponte spécifique (pour un litre : 25g Agar-Agar, 162ml de vinaigre de vin, 88ml d'eau, 5g de saccharose et 250g de levure fraîche). Il est possible de changer les boîtes à intervalles réguliers (2h) pour récupérer des embryons « stadés » c'est-à-dire des embryons ayant le même temps de développement après la ponte (+/- deux heures). Attention néanmoins à ne pas secouer les mouches, elles peuvent faire de la rétention d'œuf sous l'effet du stress. Les embryons qui seront finalement pondus auront mûri in-utero, faussant ainsi l'appréciation de l'âge de l'embryon.

## **B. Détection d'une cible lacZ : coloration Xgal**

### **Stade embryonnaire et L1**

Décoller les œufs/larves à l'eau et au pinceau, les prélever à l'aide d'un petit tamis.

Déchorionner les embryons 4 min dans la Javel 50%. Rincer à l'eau.

Coloration Xgal : Fixation 10 minutes Glutaraldéhyde 0.5%,

3 Rinçages au PBT 0.3%,

Coloration ON (Ferricyanide/Ferrocyanide 0.1M + Xgal 0.3%) à 37°C.

### **Stade larvaires**

Sexer et disséquer les larves femelles dans du PBS 1X

Coloration Xgal : Fixation, rinçages et coloration idem embryon.

Montage : aligner les corps gras en les sortants de la goutte de PBS avec une minutie de manière à ce que le tissu ne fasse pas de plis et soit étiré sur la lame. Ajouter éventuellement un peu de PBS une fois les corps alignés pour qu'ils ne sèchent pas. Observer à la lumière blanche au grossissement x20.

### **Stade adulte**

Disséquer les ovaires dans du PBS 1X

Coloration Xgal : Fixation, rinçages et coloration idem embryon.

Montage : aligner les ovaires en les sortants de la goutte de PBS avec les minuties de manière à se débarrasser des tissus parasites (tubes de Malpighi ou fragments de cuticules) pouvant être restés lors de la dissection. Observer à la lumière blanche au grossissement x10 ou x20.

## **C. Immunohistochimie**

### **C.1 Immunomarquage d'embryons**

Les embryons ont été collectés dans des boîtes de Pétri 18 à 22h AEL, après avoir été rincés à l'eau, ils sont déchorionés 4' dans la javel 50% puis rincés à l'eau de nouveau.

Les œufs sont fixés 5' sous agitation (PBS + Triton X-100 = PBT + 0.8 Formaldéhyde 3.7%), ils sont soniqués 2x7 secondes à intensité maximum par un BIORUPTOR (Diagenode) puis ils restent sous agitation 15' de plus dans la solution de fixation.

Ils sont ensuite rincés trois fois 15' dans du PBT (0.4% Triton) et pré-hybridés dans du PBT contenant 1% de Bovine Serum Albumin (PBTB) pendant 20 minutes et incubés avec les anticorps primaires dilués dans du PBTB over night (ON) à 4°C.

Après trois rinçages au PBT de 15', ils sont incubés avec les anticorps secondaires dilués dans du PBTB et du DAPI (Sigma) (1 :1000) ON à 4°C ou 3h à température ambiante.

Trois derniers rinçages au PBT de 15' puis ils sont montés dans du Citifluor (Biovalley) entre lame et lamelle avec du scotch double face comme spacer pour ne pas les écraser.

Les embryons ont été observés grâce à un microscope confocal Leica TCS SP5 Droit, utiliser le plus petit grossissement.

## **C.2 Immunomarquage de gonades larvaires**

**Premier stade Larvaire** : Elles sont récoltées sur boîtes de Pétri 24 à 28h AEL et traitées comme les embryons avec une période de sonication de 2x12 secondes.

**Stades L2/L3 /Pupes** : Ces stades peuvent être sexés sous loupe binoculaire. Seules les femelles sont disséquées dans du PBS 1X et les corps gras sont récupérés dans un Ependorf puis traités comme les embryons sans sonication et du PBT 0.4% Triton est utilisé tout du long.

Pour le montage, poser les corps gras sur un grand couvre-objet dans une goutte de PBT, placer une goutte de Citifluor au milieu de la lame, faire passer rapidement chaque corps gras contenant une gonade femelle dans le Citifluor et les aligner de l'autre côté de la lame. Attention, les tissus se fragilisent rapidement dans ce produit.

Ne pas mettre de lamelle ou utiliser des spacers. Observation au grossissement x40 sur le Leica TCS SP5 inversé puis zoom numérique sur l'objet.

Les préparations se conservent plusieurs jours en chambre humide à 4°C.

Pour les gonades larvaires et pupales, les objets sont trop gros pour être mis entre lame et lamelle, on les place donc directement sur un grand couvre-objet que l'on observe grâce au microscope confocal Leica TCS SP5 inversé. Pour l'acquisition des images, des pas de 1µm à 1.5µm ont été choisis suivant la taille des gonades et les 4 Photomultiplicateurs suivants : Dapi (415-470nm), Vert (500-550nm), Rouge (570-620nm) et Rouge lointain (670-770nm) ont été utilisés. Les acquisitions sont faites en niveaux de gris (obtention de stacks d'images), elles sont traitées et mises en couleur avec Fiji.

## **C.3 Immunomarquage d'ovaires adultes**

Les ovaires sont disséqués dans du PBS 1X sous loupe binoculaire puis directement fixés au (PBT + 0.8 Formaldéhyde 3.7%) pendant 20 minutes. Rinçages et anticorps comme pour les embryons avec du PBT 0.4% Triton tout du long.

Faire le montage des ovaires entre lames et lamelles dans une goutte de Citifluor et les aligner sur la lame. Attention, les tissus se fragilisent rapidement dans ce produit et deviennent transparents, il faut essayer d'écarter au maximum les ovarioles sans détruire leur structure en séparant doucement les germariums les uns des autres avec les minuties. Pour une observation à fort grossissement, par exemple zoom sur les germariums, séparer complètement les

ovarioles en plantant la minutie dans le stade le plus tardif et en tirant l'ovariole à travers le Citifluor, ceci le débarrassera de la gaine musculaire et permettra une meilleure résolution.

Observation au microscope confocal droit ou inversé au plus petit grossissement.

#### **C.4 Liste et concentration des anticorps utilisés**

Anticorps primaires de DSHB: rat anti-Vasa (1:1000), souris anti-HP1 (1:1000), souris anti-1B1 (1:1000); don de T. Kai: lapin anti-Piwi (1:1000) et lapin anti-Ago3 (1:1000); don de B. Theurkauf: lapin anti-Aub (1:1000) et cochon d'inde anti-Rhino (1:500) et de Rockland immunochemicals Inc.: lapin anti- $\beta$ galactosidase (1:1000).

Anticorps secondaires d'Invitrogen: Alexa Fluor 633 chèvre anti-rat IgG (H +L) (1:1000), Alexa Fluor 647 chèvre anti-rat IgM ( $\mu$ chaîne) (1:1000), Alexa Fluor 594 chèvre anti-rat IgM ( $\mu$ chaîne) (1:1000), Alexa Fluor 568 chèvre anti-lapin IgG (H+L) (1:1000); Life Technologies: Alexa Fluor 594 chèvre anti-lapin IgG (H+L) (1:1000), Alexa Fluor 594 chèvre anti-souris IgG (H+L) (1:1000), Alexa Fluor 594 chèvre anti-cochon d'inde IgG (H+L) (1:1000) et de Chromotek: GFP-Booster\_Atto488 (1:1000).

#### **C.5 Détection de la fluorescence endogène (GFP ou RFP)**

Idem que pour l'immunomarquage des ovaires adultes/ovaires larvaires exception faite que la fixation durera 15 minutes et que seul du DAPI sera utilisé pour le marquage. Le booster GFP peut être éventuellement utilisé (incubation 2h RT).

### **D. Tri cellulaire**

Récupération des larves, sexage, dissection des corps gras dans milieu S2 glacé. Ne pas trop charger les tubes à cause du gras (4 Ependorfs de 100 larves).

Broyage des tissus au Dounce Homogenizer et ressuspension dans 1 ml de PBS sans calcium + 1% de trypsine + 2.5mg collagénase. Agitation (termomixer) 30 minutes à 37°C.

Centrifugation à 1000 rpm pendant 10 minutes à 4 ° C. Élimination du gras et du surnageant.

Ressuspension dans 1ml de S2 20% de sérum de veau fœtal (pour mettre fin à la protéolyse) + 3 unités de DNase I. Incubation à RT pendant 15 minutes.

Filtration dans un tamis 125 $\mu$ m pour enlever les débris (Montage du filtre en nylon entre deux cônes bleus coupés).

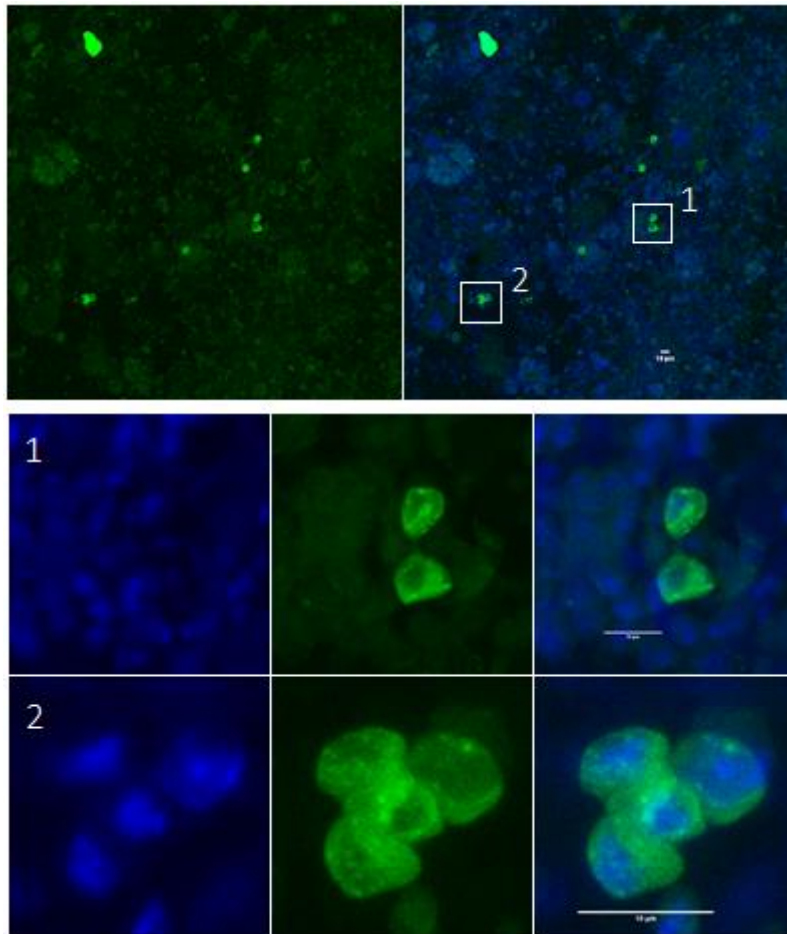
Centrifugation à 1000 rpm pendant 10 minutes à 4 ° C. Resuspendre dans 1 ml de PBS glacé sans calcium.

Centrifugation à 1000 rpm pendant 10 minutes à 4 ° C. Resuspendre dans 1 ml de S2.



Stocker sur la glace jusqu'au FACs.

Avant le tri, filtrer à travers un disque de nylon 40  $\mu\text{m}$  et ajouter du DAPI au 1/1000 (les cellules vivantes ne seront pas marquées).



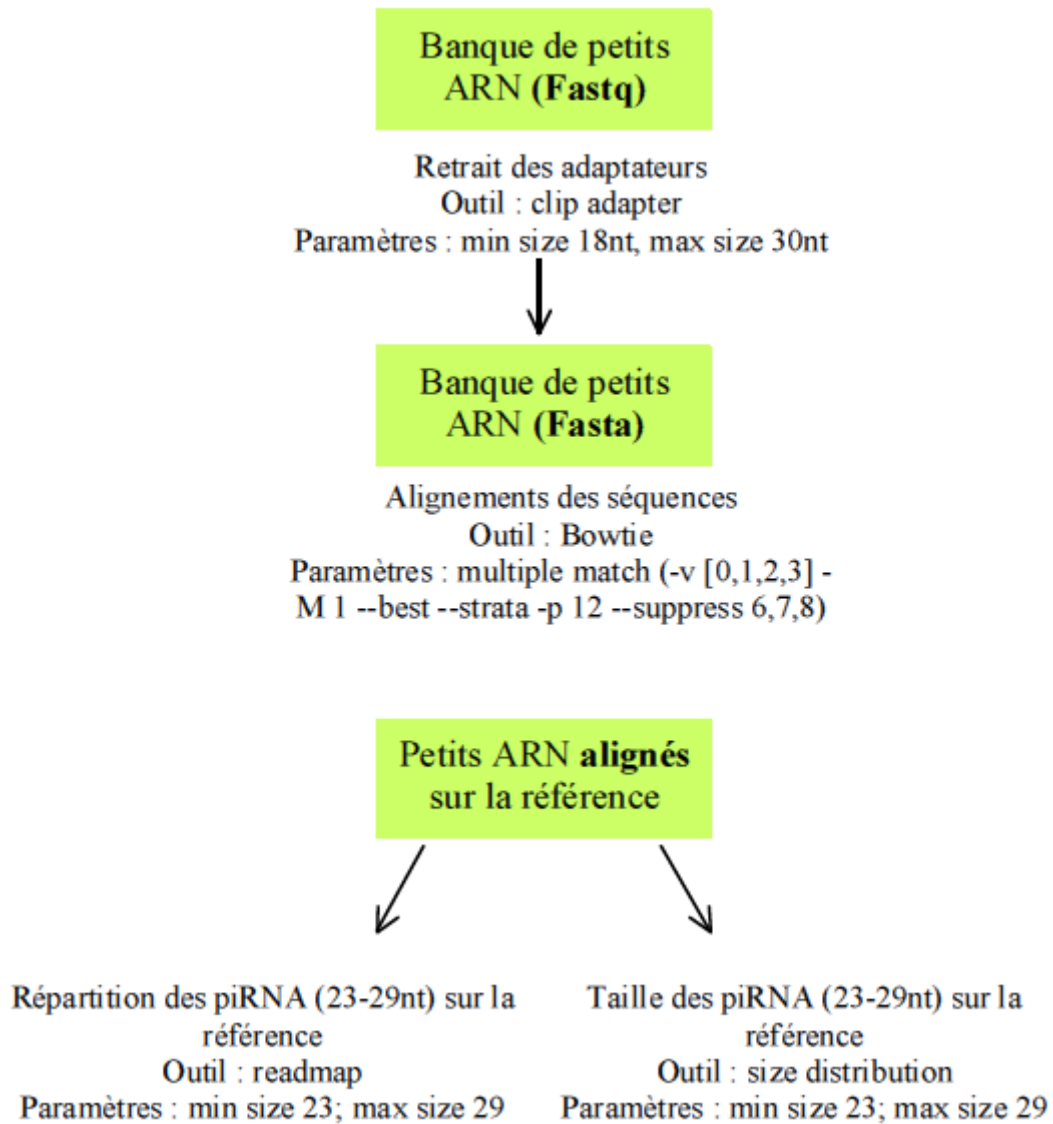
**Figure 52 – Observation au microscope à confocal des cellules récupérées.** Cellules obtenues à partir de corps gras de L3 *vasaGFP<sup>109171</sup>vasaGFP<sup>109172</sup>* homozygotes. Elles ont été rincées dans une solution PBT + Triton et marquées au DAPI. ADN (DAPI) en bleu, GFP en vert. Les deux dernières lignes sont un élargissement des zones encadrées sur la première ligne. Échelle 10  $\mu\text{m}$ .

Visuellement, les cellules vertes ressemblent à des cellules germinales de larves (Fig. 52 1-2), les rendements des expériences de tri cellulaire que j'ai effectué étaient très bas (8% des cellules théoriquement présentes). Probablement en raison d'une mauvaise séparation des cellules (Fig. 52 2), en revanche ce traitement semble affecter très peu leur viabilité ce qui est encourageant.

## E. Extraction d'ARN et Deep-Sequencing des petits ARN

Les extractions d'ARN ont été faites comme dans (Hermant et al. 2015) sauf que les tissus de départ sont des corps gras de larves contenant les gonades femelles. Cent larves de troisième stade ont été utilisées disséquées à la main pour chaque expérience. Au laboratoire,

nous disposons d'un accès au serveur Galaxy grâce à la plateforme bio-informatique (ARTbio) de l'Institut de Biologie Paris-Seine. Cet outil permet l'analyse de données de séquençage à haut débit. Les différentes étapes pour l'analyse sont présentées dans la figure suivante.



**Figure 53 – Pipeline des différentes étapes pour l'analyse des données de séquençage des petits ARN.**

## F. RTqPCR

### F.1 Ovaires Adultes (détection des transcrits de nanos et GFP)

Dissection de 30 paires d'ovaires dans du PBS 1X. Extraction des ARN totaux grâce au kit RNeasy (Qiagen). Après dosage au nanodrop, 10 µg d'ARN total sont traités à la DNase (Fermentas). Puis, 1 µg d'ARN DNasé sert de matrice à la reverse transcriptase en utilisant des hexamères comme amorces (Fermentas). Les ADNc obtenus sont ensuite quantifiés par PCR quantitative en utilisant des amorces spécifiques (gène de référence otu, séquences testées gfp, gal4) et du SYBR green (kit EVAGREEN BioRAD).

Amorces utilisées :

Nom	Séquence
otu For	5'-TGGGTGTAGAGAACGACGTG-3'
otu Rev	5'-ATTACATGCCATTCGGTGGTT-3'
gfpint1For	5'-ACGTAAACGGCCACAAGTTC-3'
gfpint1Rev	5'-AAGTCGTCTGCTTCATGTG-3'
gal4int1For	5'-CATGCGATATTTGCCGACTT-3'
gal4int1Rev	5'-GCCTTGATTCCACTTCTGTCA-3'

### F.2 Gonades larvaires (détection des transcrits de différents éléments transposables)

Dissection de 100 femelles L3 dans du PBS 1X par réplica biologique. Extraction des ARN totaux comme pour les expériences de Deep-seq. Ensuite, comme pour les ovaires adultes avec les gènes de référence otu, nos et Rp49. Les séquences testées sont dans la Figure 41.

Amorces utilisées :

Nom	Séquence
rp49U	5'-TCTGCATGAGCAGGACCTC-3'
rp49L	ATCGGTTACGGATCGAACAA-3'
otu_for	5'-TGGGTGTAGAGAACGACGTG-3'
otu_rev	5'-ATTACATGCCATTCGGTGGTT-3'
nanos_forward	5'-CGCGATCCTTGAAAATCTTTGC-3'
nanos_reverse	5'-AACTCCTGCATCACATCCTG-3'
max_for	5'-TCTAGCCAGTCGAGGCGTAT-3'
max_rev	5'-TGGAAGAGTGTGCTTTGTG-3'
inv1_for	5'-GTACCGTTTTTGAGCCCGTA-3'
inv1_rev	5'-GCGAAGTAGCCTCCTTGATG-3'
x_for	5'-TAAATTCGGAGGTGGACTCG-3'
x_rev	5'-ATCGATGTACCCTGGTGAGC-3'
blood_for	5'-CAGCGAATGTCTACCAAGCA-3'
blood_rev	5'-TATGGTCAAGGGCTTTGGTC-3'
tabor_for	5'-GGCGCCGATATATCAGTCAT-3'
tabor_rev	5'-TGAACCCTTCCAATTCCTG-3'

I-element_forward	5'-GACCAAATAAAAATAATACGACTT-3'
I-element_reverse	5'-AACTAATTGCTGGCTTGTTATG-3'
TAHRE_forward	5'-CTGTTGCACAAAGCCAAGAA-3'
TAHRE_reverse	5'-GTTGGTAATGTTTCGCGTCCT-3'
HeTA_forward	5'-CGCGCGGAACCCATCTTCAGA-3'
HeTA_reverse	5'-CGCCGCAGTCGTTTGGTGAGT-3'
burdock_forward	5'-AGGGAAATATTTGGCCATCC-3'
burdock_reverse	5'-TTTTGGCCCTGTAAACCTTG-3'
blood_forward	5'-CCAACAAAGAGGCAAGACcG-3'
blood_reverse	5'-TCGAGCTGCTTACGCATACTGTC-3'
Max_forward	5'-ATCTAGCCAGTCGAGGCGTA-3'
Max_reverse	5'-TGGAAGAGTGTCGCTTTGTG-3'
Zam_forward	5'-CTACGAAATGGCAAGATTAATTCCTTCC-3'
ZAM_reverse	5'-CCCGTTTCCTTTATGTCGCAGTAGCT-3'



## V – Références Bibliographiques

- Aasland, R, et A F Stewart. 1995. « The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. » *Nucleic Acids Research* 23 (16): 3168–3173.
- Adams, Mark D., Susan E. Celniker, Robert A. Holt, Cheryl A. Evans, Jeannine D. Gocayne, Peter G. Amanatides, Steven E. Scherer, et al. 2000. « The Genome Sequence of *Drosophila Melanogaster* ». *Science* 287 (5461): 2185–2195. doi:10.1126/science.287.5461.2185.
- Ahmad, Kami, et Steven Henikoff. 2002. « The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly ». *Molecular Cell* 9 (6): 1191–1200.
- AL-Khayat, H. A., D. Bhella, J. M. Kenney, J. F. Roth, A. J. Kingsman, E. Martin-Rendon, et H. R. Saibil. 1999. « Yeast Ty Retrotransposons Assemble into Virus-like Particles Whose T-Numbers Depend on the C-Terminal Length of the Capsid Protein ». *Journal of Molecular Biology* 292 (1): 65–73. doi:10.1006/jmbi.1999.3055.
- 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): 870–882. doi:10.1038/emboj.2011.449.
- Andress, Arlise, Yanxia Bei, Bryan R. Fonslow, Ritika Giri, Yilong Wu, John R. Yates, et Richard W. Carthew. 2016. « Spindle-E Cycling between Nuage and Cytoplasm Is Controlled by Qin and PIWI Proteins ». *The Journal of Cell Biology* 213 (2): 201–211. doi:10.1083/jcb.201411076.
- Andreyeva, Evgenia N., Elena S. Belyaeva, Valerii F. Semeshin, Galina V. Pokholkova, et Igor F. Zhimulev. 2005. « Three Distinct Chromatin Domains in Telomere Ends of Polytene Chromosomes in *Drosophila Melanogaster* Tel Mutants ». *Journal of Cell Science* 118 (Pt 23): 5465–5477. doi:10.1242/jcs.02654.
- Anne, Joël. 2010. « Targeting and Anchoring Tudor in the Pole Plasm of the *Drosophila* Oocyte ». *PLoS One* 5 (12): e14362. doi:10.1371/journal.pone.0014362.
- Aravin, Alexei A., Mikhail S. Klenov, Vasilii V. Vagin, Frédéric Bantignies, Giacomo Cavalli, et Vladimir A. Gvozdev. 2004. « Dissection of a Natural RNA Silencing Process in the *Drosophila Melanogaster* Germ Line ». *Molecular and Cellular Biology* 24 (15): 6742–6750. doi:10.1128/MCB.24.15.6742-6750.2004.
- Aravin, Alexei, Dimos Gaidatzis, Sébastien Pfeffer, Mariana Lagos-Quintana, Pablo Landgraf, Nicola Iovino, Patricia Morris, et al. 2006. « A Novel Class of Small RNAs Bind to MILI Protein in Mouse Testes ». *Nature* 442 (7099): 203–7.
- Baldi, Sandro, et Peter B. Becker. 2013. « The Variant Histone H2A.V of *Drosophila*—three Roles, Two Guises ». *Chromosoma* 122 (4): 245–258. doi:10.1007/s00412-013-0409-x.
- Ballut, Lionel, Brice Marchadier, Aurélie Baguet, Catherine Tomasetto, Bertrand Séraphin, et Hervé Le Hir. 2005. « The Exon Junction Core Complex Is Locked onto RNA by Inhibition of eIF4AIII ATPase Activity ». *Nature Structural & Molecular Biology* 12 (10): 861–869. doi:10.1038/nsmb990.
- Bantignies, Frédéric, et Giacomo Cavalli. 2011. « Polycomb Group Proteins: Repression in 3D ». *Trends in Genetics: TIG* 27 (11): 454–464. doi:10.1016/j.tig.2011.06.008.
- Barckmann, Bridlin, Stéphanie Pierson, Jérémy Dufourt, Catherine Papin, Claudia Armenise, Phillip Port, Thomas Grentzinger, et al. 2015. « Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo ». *Cell Reports* 12 (7): 1205–1216. doi:10.1016/j.celrep.2015.07.030.

- Barckmann, Bridlin, et Martine Simonelig. 2013. « Control of Maternal mRNA Stability in Germ Cells and Early Embryos ». *Biochimica Et Biophysica Acta* 1829 (6-7): 714–724. doi:10.1016/j.bbagr.2012.12.011.
- Bartel, David P. 2004. « MicroRNAs: Genomics, Biogenesis, Mechanism, and Function ». *Cell* 116 (2): 281–97.
- Batista, Pedro J., Graham Ruby, Julie M. Claycomb, Rosaria Chiang, Noah Fahlgren, Kristin D. Kasschau, Daniel A. Chaves, et al. 2008. « PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans* ». *Molecular cell* 31 (1): 67–78.
- Baudry, Céline, Sophie Malinsky, Matthieu Restituito, Aurélie Kapusta, Sarah Rosa, Eric Meyer, et Mireille Bétermier. 2009. « PiggyMac, a Domesticated piggyBac Transposase Involved in Programmed Genome Rearrangements in the Ciliate *Paramecium Tetraurelia* ». *Genes & Development* 23 (21): 2478–2483. doi:10.1101/gad.547309.
- Baulcombe, David. 2004. « RNA Silencing in Plants ». *Nature* 431 (7006): 356–63.
- Benachenhou, Farid, Göran O. Sperber, Erik Bongcam-Rudloff, Göran Andersson, Jef D. Boeke, et Jonas Blomberg. 2013. « Conserved Structure and Inferred Evolutionary History of Long Terminal Repeats (LTRs) ». *Mobile DNA* 4 (1): 5.
- Bhattacharyya, M., C. Martin, et A. Smith. 1993. « The Importance of Starch Biosynthesis in the Wrinkled Seed Shape Character of Peas Studied by Mendel ». *Plant Molecular Biology* 22 (3): 525–31.
- Biémont, Christian. 2010. « A Brief History of the Status of Transposable Elements: From Junk DNA to Major Players in Evolution ». *Genetics* 186 (4): 1085–93.
- Billmyre, R. Blake, Silvia Calo, Marianna Feretzaki, Xuying Wang, et Joseph Heitman. 2013. « RNAi Function, Diversity, and Loss in the Fungal Kingdom ». *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology* 21 (0). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3874831/>.
- Boivin, Antoine, Christelle Gally, Sophie Netter, Dominique Anxolabéhère, et Stéphane Ronsseray. 2003. « Telomeric Associated Sequences of *Drosophila* Recruit Polycomb-Group Proteins in Vivo and Can Induce Pairing-Sensitive Repression ». *Genetics* 164 (1): 195–208.
- 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): 1387–1392. doi:10.1126/science.1165171.
- 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.
- Callaini, G., M. G. Riparbelli, et R. Dallai. 1995. « Pole Cell Migration through the Gut Wall of the *Drosophila* Embryo: Analysis of Cell Interactions ». *Developmental Biology* 170 (2): 365–375. doi:10.1006/dbio.1995.1222.
- Carr, Martin, Douda Bensasson, et Casey M. Bergman. 2012. « Evolutionary Genomics of Transposable Elements in *Saccharomyces cerevisiae* ». *PLOS ONE* 7 (11): e50978.

- Cerutti, Heriberto, et J. Armando Casas-Mollano. 2006. « On the origin and functions of RNA-mediated silencing: from protists to man ». *Current genetics* 50 (2): 81–99. doi:10.1007/s00294-006-0078-x.
- Chambeyron, S., et A. Bucheton. 2005. « I Elements in Drosophila: In Vivo Retrotransposition and Regulation ». *Cytogenetic and Genome Research* 110 (1-4): 215–222. doi:10.1159/000084955.
- Chen, Yung-Chia Ariel, Evelyn Stuwe, Yicheng Luo, Maria Ninova, Adrien Le Thomas, Ekaterina Rozhavskaya, Sisi Li, et al. 2016. « Cutoff Suppresses RNA Polymerase II Termination to Ensure Expression of piRNA Precursors ». *Molecular Cell*, juin. doi:10.1016/j.molcel.2016.05.010.
- Cheng, Chao-Yin, Alexander Vogt, Kazufumi Mochizuki, et Meng-Chao Yao. 2010. « A Domesticated piggyBac Transposase Plays Key Roles in Heterochromatin Dynamics and DNA Cleavage during Programmed DNA Deletion in Tetrahymena Thermophila ». *Molecular Biology of the Cell* 21 (10): 1753–1762. doi:10.1091/mbc.E09-12-1079.
- 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.
- Clark, Josef P., et Nelson C. Lau. 2014. « Piwi proteins and piRNAs step onto the systems biology stage ». *Advances in experimental medicine and biology* 825: 159–97.
- Clelland, Brett W., et Michael C. Schultz. 2010. « Genome Stability Control by Checkpoint Regulation of tRNA Gene Transcription ». *Transcription* 1 (3): 115–125. doi:10.4161/trns.1.3.13735.
- Coffman, Clark R, Rachel C Strohm, Fredrick D Oakley, Yukiko Yamada, Danielle Przychodzin, et Robert E Boswell. 2002. « Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. » *Genetics* 162 (1): 273–284.
- Cox, D. N., A. Chao, J. Baker, L. Chang, D. Qiao, et H. Lin. 1998. « A Novel Class of Evolutionarily Conserved Genes Defined by Piwi Are Essential for Stem Cell Self-Renewal ». *Genes & Development* 12 (23): 3715–3727.
- Cryderman, Diane E., Stephanie K. Grade, Yuhong Li, Laura Fanti, Sergio Pimpinelli, et Lori L. Wallrath. 2005. « Role of Drosophila HP1 in Euchromatic Gene Expression ». *Developmental Dynamics* 232 (3): 767–774. doi:10.1002/dvdy.20310.
- « Current genetics - Repetitive DNA ». 2016. Consulté le juin 3. [http://biol.lf1.cuni.cz/ucebnice/en/repetitive\\_dna.htm](http://biol.lf1.cuni.cz/ucebnice/en/repetitive_dna.htm).
- 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): 749–761. doi:10.1016/j.molcel.2013.04.007.
- 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): 112–115. doi:10.1038/nature11416.
- . 2013. « piRNAs and Epigenetic Conversion in Drosophila ». *Fly* 7 (4): 237–241. doi:10.4161/fly.26522.
- Dehghani, Mehrnoush, et Paul Lasko. 2015. « In Vivo Mapping of the Functional Regions of the DEAD-Box Helicase Vasa ». *Biology Open* 4 (4): 450–462. doi:10.1242/bio.201410579.
- Dennis, Cynthia, Vanessa Zanni, Emilie Brassset, 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.
- Ding, Shou-Wei, et Olivier Voinnet. 2007. « Antiviral Immunity Directed by Small RNAs ». *Cell* 130 (3): 413-26.
- Ding, Xiaofang, Huangtao Guan, et Honggang Li. 2013. « Characterization of a piRNA Binding Protein Miwi in Mouse Oocytes ». *Theriogenology* 79 (4): 610–615.e1. doi:10.1016/j.theriogenology.2012.11.013.
- Dönertas, Derya, Grzegorz Sienski, et Julius Brennecke. 2013. « *Drosophila* Gtsf1 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.
- . 1997. « Transgene Repeat Arrays Interact with Distant Heterochromatin and Cause Silencing in Cis and Trans ». *Genetics* 147 (3): 1181–1190.
- Driever, W., et C. Nüsslein-Volhard. 1989. « The Bicoid Protein Is a Positive Regulator of Hunchback Transcription in the Early *Drosophila* Embryo ». *Nature* 337 (6203): 138–143. doi:10.1038/337138a0.
- Drinnenberg, Ines A., Gerald R. Fink, et David P. Bartel. 2011. « Compatibility with Killer explains the Rise of RNAi-deficient fungi ». *Science (New York, N.Y.)* 333 (6049): 1592.
- Du, W. W., W. Yang, J. Xuan, S. Gupta, S. N. Krylov, X. Ma, Q. Yang, et B. B. Yang. 2016. « Reciprocal Regulation of miRNAs and piRNAs in Embryonic Development ». *Cell Death and Differentiation*, mars.
- Dufourt, Jérémy, Cynthia Dennis, Antoine Boivin, Nathalie Gueguen, Emmanuelle Théron, Coline Goriaux, Pierre Pouchin, Stéphane Ronsseray, Emilie Brassat, et Chantal Vaury. 2014. « Spatio-temporal requirements for transposable element piRNA-mediated silencing during *Drosophila* oogenesis ». *Nucleic Acids Research* 42 (4): 2512–2524. doi:10.1093/nar/gkt1184.
- Eissenberg, J C, T C James, D M Foster-Hartnett, T Hartnett, V Ngan, et S C Elgin. 1990. « Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. » *Proceedings of the National Academy of Sciences of the United States of America* 87 (24): 9923–9927.
- Engels, W. R. 1984. « A Trans-Acting Product Needed for P Factor Transposition in *Drosophila* ». *Science (New York, N.Y.)* 226 (4679): 1194-96.
- Engels, William R., et Christine R. Preston. 1979. « Hybrid Dysgenesis in *DROSOPHILA MELANOGASTER*: The Biology of Female and Male Sterility ». *Genetics* 92 (1): 161–174.
- Ephrussi, A., et R. Lehmann. 1992. « Induction of Germ Cell Formation by Oskar ». *Nature* 358 (6385): 387–392. doi:10.1038/358387a0.
- Fanti, L., G. Giovino, M. Berloco, et S. Pimpinelli. 1998a. « The Heterochromatin Protein 1 Prevents Telomere Fusions in *Drosophila* ». *Molecular Cell* 2 (5): 527–538.
- Feltzin, Virzhiniya L., Mugdha Khaladkar, Masashi Abe, Michael Parisi, Gert-Jan Hendriks, Junhyong Kim, et Nancy M. Bonini. 2015. « The Exonuclease Nibbler Regulates Age-Associated Traits and Modulates piRNA Length in *Drosophila* ». *Aging Cell* 14 (3): 443–452. doi:10.1111/acel.12323.

- Finnegan, D. J. 1992. « Transposable Elements ». *Current Opinion in Genetics & Development* 2 (6): 861-67.
- 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.
- Frydrychova, Radmila Capkova, James M. Mason, et Trevor K. Archer. 2008a. « HP1 Is Distributed Within Distinct Chromatin Domains at *Drosophila* Telomeres ». *Genetics* 180 (1): 121–131. doi:10.1534/genetics.108.090647.
- 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): 1716–1721.
- Gavis, E. R., et R. Lehmann. 1992. « Localization of Nanos RNA Controls Embryonic Polarity ». *Cell* 71 (2): 301–313.
- Gebert, Daniel, et David Rosenkranz. 2015. « RNA-Based Regulation of Transposon Expression ». *Wiley Interdisciplinary Reviews. RNA* 6 (6): 687-708.
- 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.
- Gilboa, Lilach. 2015. « Organizing Stem Cell Units in the *Drosophila* Ovary ». *Current Opinion in Genetics & Development* 32 (juin): 31–36. doi:10.1016/j.gde.2015.01.005.
- Gilboa, Lilach, et Ruth Lehmann. 2004. « Repression of Primordial Germ Cell Differentiation Parallels Germ Line Stem Cell Maintenance ». *Current Biology: CB* 14 (11): 981–986. doi:10.1016/j.cub.2004.05.049.
- . 2006. « Soma-Germline Interactions Coordinate Homeostasis and Growth in the *Drosophila* Gonad ». *Nature* 443 (7107): 97–100. doi:10.1038/nature05068.
- Girard, Angélique, Ravi Sachidanandam, Gregory J. Hannon, et Michelle A. Carmell. 2006. « A Germline-Specific Class of Small RNAs Binds Mammalian Piwi Proteins ». *Nature* 442 (7099): 199-202.
- Goldberg, R. B., G. Hoschek, et L. O. Vodkin. 1983. « An Insertion Sequence Blocks the Expression of a Soybean Lectin Gene ». *Cell* 33 (2): 465–475.
- Goriaux, Coline, Sophie Desset, Yoan Renaud, Chantal Vauray, et Emilie Brassat. 2014. « Transcriptional Properties and Splicing of the Flamenco piRNA Cluster ». *EMBO Reports* 15 (4): 411–418. doi:10.1002/embr.201337898.
- Greil, Frauke, Ineke van der Kraan, Jeffrey Delrow, James F. Smothers, Elzo de Wit, Harmen J. Bussemaker, Roel van Driel, Steven Henikoff, et Bas van Steensel. 2003. « Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location ». *Genes & Development* 17 (22): 2825–2838. doi:10.1101/gad.281503.
- Grimson, Andrew, Mansi Srivastava, Bryony Fahey, Ben J. Woodcroft, H. Rosaria Chiang, Nicole King, Bernard M. Degnan, Daniel S. Rokhsar, et David P. Bartel. 2008. « The early origins of microRNAs and Piwi-interacting RNAs in animals ». *Nature* 455 (7217). doi:10.1038/nature07415.
- 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): 1587–1590. doi:10.1126/science.1140494.
- 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): 762–777. 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): 3977–3993. doi:10.1038/emboj.2011.308.
- Hanyu-Nakamura, Kazuko, Hiroko Sonobe-Nojima, Akie Tanigawa, Paul Lasko, et Akira Nakamura. 2008. « Drosophila Pgc Protein Inhibits P-TEFb Recruitment to Chromatin in Primordial Germ Cells ». *Nature* 451 (7179): 730–733. doi:10.1038/nature06498.
- 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): 2823–2832.
- He, Li, Xiaobo Wang, et Denise J. Montell. 2011. « Shining Light on Drosophila Oogenesis: Live Imaging of Egg Development ». *Current Opinion in Genetics & Development* 21 (5): 612–619. doi:10.1016/j.gde.2011.08.011.
- Hermant, Catherine, Antoine Boivin, Laure Teyssset, Valérie Delmarre, Amna Asif-Laidin, Marius van den Beek, Christophe Antoniewski, et Stéphane Ronsseray. 2015. « Paramutation in Drosophila Requires Both Nuclear and Cytoplasmic Actors of the piRNA Pathway and Induces Cis-Spreading of piRNA Production ». *Genetics* 201 (4): 1381–1396. doi:10.1534/genetics.115.180307.
- Hirakata, Shigeki, et Mikiko C. Siomi. 2016. « piRNA Biogenesis in the Germline: From Transcription of piRNA Genomic Sources to piRNA Maturation ». *Biochimica Et Biophysica Acta* 1859 (1): 82–92. doi:10.1016/j.bbagr.2015.09.002.
- Huang, Xiao A., Hang Yin, Sarah Sweeney, Debasish Raha, Michael Snyder, et Haifan Lin. 2013. « A Major Epigenetic Programming Mechanism Guided by piRNAs ». *Developmental Cell* 24 (5): 502–516. doi:10.1016/j.devcel.2013.01.023.
- Ibarra, Christian A., Xiaoqi Feng, Vera K. Schoft, Tzung-Fu Hsieh, Rie Uzawa, Jessica A. Rodrigues, Assaf Zemach, et al. 2012. « Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes ». *Science (New York, N.Y.)* 337 (6100): 1360–64.
- 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): 279–283. doi:10.1038/nature11502.
- Ito, Hidetaka. 2012. « Small RNAs and Transposon Silencing in Plants ». *Development, Growth & Differentiation* 54 (1): 100–107.
- Jaglarz, M. K., et K. R. Howard. 1995. « The Active Migration of Drosophila Primordial Germ Cells ». *Development (Cambridge, England)* 121 (11): 3495–3503.
- James, T. C., J. C. Eissenberg, C. Craig, V. Dietrich, A. Hobson, et S. C. Elgin. 1989. « Distribution Patterns of HP1, a Heterochromatin-Associated Nonhistone Chromosomal Protein of Drosophila ». *European Journal of Cell Biology* 50 (1): 170–180.
- Jenkins, Allison B., J. Michael McCaffery, et Mark Van Doren. 2003. « Drosophila E-Cadherin Is Essential for Proper Germ Cell-Soma Interaction during Gonad Morphogenesis ». *Development (Cambridge, England)* 130 (18): 4417–4426.
- Jensen, S., M. P. Gassama, et T. Heidmann. 1999. « Taming of Transposable Elements by Homology-Dependent Gene Silencing ». *Nature Genetics* 21 (2): 209–212. doi:10.1038/5997.
- Jenuwein, Thomas. 2001. « Re-SET-Ting Heterochromatin by Histone Methyltransferases ». *Trends in Cell Biology* 11 (6): 266–273. doi:10.1016/S0962-8924(01)02001-3.

- Jeske, Mandy, Matteo Bordi, Sebastian Glatt, Sandra Müller, Vladimir Rybin, Christoph W. Müller, et Anne Ephrussi. 2015. « The Crystal Structure of the Drosophila Germline Inducer Oskar Identifies Two Domains with Distinct Vasa Helicase- and RNA-Binding Activities ». *Cell Reports* 12 (4): 587–598. doi:10.1016/j.celrep.2015.06.055.
- Joly, Willy, Aymeric Chartier, Patricia Rojas-Rios, Isabelle Busseau, et Martine Simonelig. 2013. « The CCR4 Deadenylase Acts with Nanos and Pumilio in the Fine-Tuning of Mei-P26 Expression to Promote Germline Stem Cell Self-Renewal ». *Stem Cell Reports* 1 (5): 411–424. doi:10.1016/j.stemcr.2013.09.007.
- Jongens, T. A., L. D. Ackerman, J. R. Swedlow, L. Y. Jan, et Y. N. Jan. 1994. « Germ Cell-Less Encodes a Cell Type-Specific Nuclear Pore-Associated Protein and Functions Early in the Germ-Cell Specification Pathway of Drosophila ». *Genes & Development* 8 (18): 2123–2136.
- Joshua-Tor, Leemor, et Gregory J. Hannon. 2011. « Ancestral Roles of Small RNAs: An Ago-Centric Perspective ». *Cold Spring Harbor Perspectives in Biology* 3 (10). doi:10.1101/cshperspect.a003772.
- 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). doi:10.1371/journal.pgen.0030158.
- Juliano, Celina E., Adrian Reich, Na Liu, Jessica Götzfried, Mei Zhong, Selen Uman, Robert A. Reenan, Gary M. Wessel, Robert E. Steele, et Haifan Lin. 2014. « PIWI Proteins and PIWI-Interacting RNAs Function in Hydra Somatic Stem Cells ». *Proceedings of the National Academy of Sciences* 111 (1): 337–342. doi:10.1073/pnas.1320965111.
- Kalmykova, Alla I., Mikhail S. Klenov, et Vladimir A. Gvozdev. 2005. « Argonaute protein PIWI controls mobilization of retrotransposons in the Drosophila male germline ». *Nucleic Acids Research* 33 (6): 2052–2059. doi:10.1093/nar/gki323.
- Karpen, G. H., et A. C. Spradling. 1992. « Analysis of Subtelomeric Heterochromatin in the Drosophila Minichromosome Dp1187 by Single P Element Insertional Mutagenesis ». *Genetics* 132 (3): 737–753.
- Kidwell, Margaret G., James F. Kidwell, et John A. Sved. 1977. « Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and Male Recombination ». *Genetics* 86 (4): 813–833.
- Kidwell, Margaret G., et Damon R. Lisch. 2000. « Transposable Elements and Host Genome Evolution ». *Trends in Ecology & Evolution* 15 (3): 95–99.
- Kim, A., C. Terzian, P. Santamaria, A. Pélisson, N. Purd'homme, et A. Bucheton. 1994. « Retroviruses in Invertebrates: The Gypsy Retrotransposon Is Apparently an Infectious Retrovirus of Drosophila Melanogaster ». *Proceedings of the National Academy of Sciences of the United States of America* 91 (4): 1285–89.
- Kim, V. Narry. 2006. « Small RNAs Just Got Bigger: Piwi-Interacting RNAs (piRNAs) in Mammalian Testes ». *Genes & Development* 20 (15): 1993–1997. doi:10.1101/gad.1456106.
- Kirino, Yohei, et Zissimos Mourelatos. 2007. « The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs ». *RNA* 13 (9): 1397–1401.
- Klattenhoff, Carla, Diana P. Bratu, Nadine McGinnis-Schultz, Birgit S. Koppetsch, Heather A. Cook, et William E. Theurkauf. 2007. « Drosophila rasiRNA Pathway Mutations

- Disrupt Embryonic Axis Specification through Activation of an ATR/Chk2 DNA Damage Response ». *Developmental Cell* 12 (1): 45-55.
- Klattenhoff, Carla, Hualin Xi, Chengjian Li, Soohyun Lee, Jia Xu, Jaspreet S. Khurana, Fan Zhang, et al. 2009. « The Drosophila HP1 homologue Rhino is required for transposon silencing and piRNA production by dual strand clusters ». *Cell* 138 (6): 1137–1149. doi:10.1016/j.cell.2009.07.014.
- 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): 18760–18765. doi:10.1073/pnas.1106676108.
- Kunwar, Prabhat S., Daria E. Siekhaus, et Ruth Lehmann. 2006. « In Vivo Migration: A Germ Cell Perspective ». *Annual Review of Cell and Developmental Biology* 22: 237–265. doi:10.1146/annurev.cellbio.22.010305.103337.
- Laidin, A. V. Delmarre, J. Laurentie, W. J. Miller, S. Ronsseray, et L. Teyssset. en préparation. « Structure, dynamic and properties of subtelomeric piRNA clusters in Drosophila ».
- Laski, F. A., D. C. Rio, et G. M. Rubin. 1986. « Tissue Specificity of Drosophila P Element Transposition Is Regulated at the Level of mRNA Splicing ». *Cell* 44 (1): 7-19.
- Lau, Nelson C., Nicolas Robine, Raquel Martin, Wei-Jen Chung, Yuzo Niki, Eugene Berezikov, et Eric C. Lai. 2009. « Abundant Primary piRNAs, Endo-siRNAs, and microRNAs in a Drosophila Ovary Cell Line ». *Genome Research* 19 (10): 1776–1785. doi:10.1101/gr.094896.109.
- Lau, Nelson C., Anita G. Seto, Jinkuk Kim, Satomi Kuramochi-Miyagawa, Toru Nakano, David P. Bartel, et Robert E. Kingston. 2006. « Characterization of the piRNA Complex from Rat Testes ». *Science (New York, N.Y.)* 313 (5785): 363–367. doi:10.1126/science.1130164.
- Le Bras, Stéphanie, et Mark Van Doren. 2006. « Development of the Male Germline Stem Cell Niche in Drosophila ». *Developmental Biology* 294 (1): 92–103. doi:10.1016/j.ydbio.2006.02.030.
- Le Hir, Hervé, Jérôme Saulière, et Zhen Wang. 2016. « The Exon Junction Complex as a Node of Post-Transcriptional Networks ». *Nature Reviews. Molecular Cell Biology* 17 (1): 41–54. doi:10.1038/nrm.2015.7.
- Le Thomas, Adrien, Georgi K. Marinov, et Alexei A. Aravin. 2014. « A Transgenerational Process Defines piRNA Biogenesis in Drosophila Virilis ». *Cell Reports* 8 (6): 1617–1623. doi:10.1016/j.celrep.2014.08.013.
- 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): 390–399. 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): 1667–1680. doi:10.1101/gad.245514.114.
- Le Thomas, Adrien, Katalin Fejes Tóth, et Alexei A. Aravin. 2014. « To Be or Not to Be a piRNA: Genomic Origin and Processing of piRNAs ». *Genome Biology* 15 (1): 204. doi:10.1186/gb4154.

- Leatherman, Judith L., Lissa Levin, Julie Boero, et Thomas A. Jongens. 2002. « Germ Cell-Less Acts to Repress Transcription during the Establishment of the Drosophila Germ Cell Lineage ». *Current Biology: CB* 12 (19): 1681–1685.
- Lee, Miler T., Ashley R. Bonneau, et Antonio J. Giraldez. 2014. « Zygotic Genome Activation during the Maternal-to-Zygotic Transition ». *Annual Review of Cell and Developmental Biology* 30: 581–613. doi:10.1146/annurev-cellbio-100913-013027.
- 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.
- Lemaitre, B., S. Ronsseray, et D. Coen. 1993. « Maternal Repression of the P Element Promoter in the Germline of Drosophila Melanogaster: A Model for the P Cytotype ». *Genetics* 135 (1): 149–160.
- Lesch, Bluma J., et David C. Page. 2012. « Genetics of Germ Cell Development ». *Nature Reviews. Genetics* 13 (11): 781–794. doi:10.1038/nrg3294.
- Lewandowski, Jordan P., Kathy B. Sheehan, Paul E. Bennett, et Robert E. Boswell. 2010. « Mago Nashi, Tsunagi/Y14 and Ranshi form a complex that influences oocyte differentiation in Drosophila melanogaster ». *Developmental biology* 339 (2): 307–319. doi:10.1016/j.ydbio.2009.12.035.
- Li, Yang, Chaoqun Li, Jie Xia, et Youxin Jin. 2011. « Domestication of Transposable Elements into MicroRNA Genes in Plants ». *PloS One* 6 (5): e19212.
- Lin, H., et A. C. Spradling. 1993. « Germline Stem Cell Division and Egg Chamber Development in Transplanted Drosophila Germaria ». *Developmental Biology* 159 (1): 140–152. doi:10.1006/dbio.1993.1228.
- . 1997. « A Novel Group of Pumilio Mutations Affects the Asymmetric Division of Germline Stem Cells in the Drosophila Ovary ». *Development (Cambridge, England)* 124 (12): 2463–2476.
- Liu, Li, Hongying Qi, Jianquan Wang, et Haifan Lin. 2011. « PAPI, a Novel TUDOR-Domain Protein, Complexes with AGO3, ME31B and TRAL in the Nuage to Silence Transposition ». *Development (Cambridge, England)* 138 (9): 1863–1873. doi:10.1242/dev.059287.
- Lu, Xuemin, Jennifer M. Li, Olivier Elemento, Saeed Tavazoie, et Eric F. Wieschaus. 2009. « Coupling of Zygotic Transcription to Mitotic Control at the Drosophila Mid-Blastula Transition ». *Development (Cambridge, England)* 136 (12): 2101–2110. doi:10.1242/dev.034421.
- Mahowald, A. P. 1971. « Polar Granules of Drosophila. 3. The Continuity of Polar Granules during the Life Cycle of Drosophila ». *The Journal of Experimental Zoology* 176 (3): 329–43.
- 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.
- Malone, Colin D., Claire Mestdagh, Junaid Akhtar, Nastasja Kreim, Pia Deinhard, Ravi Sachidanandam, Jessica Treisman, et Jean-Yves Roignant. 2014. « The Exon Junction Complex Controls Transposable Element Activity by Ensuring Faithful Splicing of the Piwi Transcript ». *Genes & Development* 28 (16): 1786–1799. doi:10.1101/gad.245829.114.
- 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): 1841–1854.

- Marracci, S., R. Batistoni, G. Pesole, L. Citti, et I. Nardi. 1996. « Gypsy/Ty3-like Elements in the Genome of the Terrestrial Salamander *Hydromantes* (Amphibia, Urodela) ». *Journal of Molecular Evolution* 43 (6): 584-93.
- 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.
- Matsumoto, Naoki, Kaoru Sato, Hiroshi Nishimasu, Yurika Namba, Kana Miyakubi, Naoshi Dohmae, Ryuichiro Ishitani, Haruhiko Siomi, Mikiko C. Siomi, et Osamu Nureki. 2015. « Crystal Structure and Activity of the Endoribonuclease Domain of the piRNA Pathway Factor Maelstrom ». *Cell Reports* 11 (3): 366-375. doi:10.1016/j.celrep.2015.03.030.
- McDonald, J. F., L. V. Matyunina, S. Wilson, I. K. Jordan, N. J. Bowen, et W. J. Miller. 1997. « LTR Retrotransposons and the Evolution of Eukaryotic Enhancers ». *Genetica* 100 (1-3): 3-13.
- 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): 1615-1624. doi:10.1534/genetics.106.068106.
- Mochizuki, Kazufumi, et Martin A. Gorovsky. 2004. « Small RNAs in Genome Rearrangement in *Tetrahymena* ». *Current Opinion in Genetics & Development* 14 (2): 181-87.
- Mohn, Fabio, Dominik Handler, et Julius Brennecke. 2015. « Noncoding RNA. piRNA-Guided Slicing Specifies Transcripts for Zucchini-Dependent, Phased piRNA Biogenesis ». *Science (New York, N.Y.)* 348 (6236): 812-17. doi:10.1126/science.aaa1039.
- 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): 1364-1379. doi:10.1016/j.cell.2014.04.031.
- Mohr, Stephanie E., Simon T. Dillon, et Robert E. Boswell. 2001. « The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during *Drosophila* oogenesis ». *Genes & Development* 15 (21): 2886-2899. doi:10.1101/gad.927001.
- Molla-Herman, Anahi, Ana Maria Vallés, Carine Ganem-Elbaz, Christophe Antoniewski, et Jean-René Huynh. 2015. « tRNA Processing Defects Induce Replication Stress and Chk2-Dependent Disruption of piRNA Transcription ». *The EMBO Journal* 34 (24): 3009-3027. doi:10.15252/embj.201591006.
- Montpetit, Ben, Nathan D. Thomsen, Kara J. Helmke, Markus A. Seeliger, James M. Berger, et Karsten Weis. 2011. « A Conserved Mechanism of DEAD-Box ATPase Activation by Nucleoporins and InsP6 in mRNA Export ». *Nature* 472 (7342): 238-242. doi:10.1038/nature09862.
- 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): 736-748. doi:10.1016/j.molcel.2013.04.006.
- Mugat, Bruno, Abdou Akkouche, Vincent Serrano, Claudia Armenise, Blaise Li, Christine Brun, Tudor A. Fulga, David Van Vactor, Alain Péllisson, et Séverine Chambeyron. 2015. « MicroRNA-Dependent Transcriptional Silencing of Transposable Elements in *Drosophila* Follicle Cells ». *PLoS Genetics* 11 (5). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4451950/>.

- Murota, Yukiko, Hirotsugu Ishizu, Shinichi Nakagawa, Yuka W. Iwasaki, Shinsuke Shibata, Miharu 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): 103–113. doi:10.1016/j.celrep.2014.05.043.
- Murray, Sheryl M., Shu Yuan Yang, et Mark Van Doren. 2010. « Germ Cell Sex Determination: A Collaboration between Soma and Germline ». *Current Opinion in Cell Biology* 22 (6): 722–729. doi:10.1016/j.ceb.2010.09.006.
- Nakamura, A., R. Amikura, M. Mukai, S. Kobayashi, et P. F. Lasko. 1996. « Requirement for a Noncoding RNA in Drosophila Polar Granules for Germ Cell Establishment ». *Science (New York, N.Y.)* 274 (5295): 2075–2079.
- Ni, Jian-Quan, Rui Zhou, Benjamin Czech, Lu-Ping Liu, Laura Holderbaum, Donghui Yang-Zhou, Hye-Seok Shim, et al. 2011. « A Genome-Scale shRNA Resource for Transgenic RNAi in Drosophila ». *Nature Methods* 8 (5): 405–407. doi:10.1038/nmeth.1592.
- Ohlstein, B., et D. McKearin. 1997. « Ectopic Expression of the Drosophila Bam Protein Eliminates Oogenic Germline Stem Cells ». *Development (Cambridge, England)* 124 (18): 3651–3662.
- Okamura, Katsutomo, Akira Ishizuka, Haruhiko Siomi, et Mikiko C. Siomi. 2004. « Distinct Roles for Argonaute Proteins in Small RNA-Directed RNA Cleavage Pathways ». *Genes & Development* 18 (14): 1655–66.
- 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): 3301–3317. doi:10.1038/emboj.2010.212.
- Palakodeti, Dasaradhi, Magda Smielewska, Yi-Chien Lu, Gene W. Yeo, et Brenton R. Graveley. 2008. « The PIWI Proteins SMEDWI-2 and SMEDWI-3 Are Required for Stem Cell Function and piRNA Expression in Planarians ». *RNA (New York, N.Y.)* 14 (6): 1174–1186. doi:10.1261/rna.1085008.
- 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.
- Paro, R., et D. S. Hogness. 1991. « The Polycomb Protein Shares a Homologous Domain with a Heterochromatin-Associated Protein of Drosophila ». *Proceedings of the National Academy of Sciences of the United States of America* 88 (1): 263–267.
- Patil, Veena S, Amit Anand, Alisha Chakrabarti, et Toshie Kai. 2014. « The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of Drosophila melanogaster ». *BMC Biology* 12 (octobre). doi:10.1186/s12915-014-0061-9.
- Patil, Veena S., et Toshie Kai. 2010. « Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas ». *Current Biology: CB* 20 (8): 724–730. doi:10.1016/j.cub.2010.02.046.
- Péllisson, Alain, Emeline Sarot, Geneviève Payen-Groschêne, et Alain Bucheton. 2007. « A Novel Repeat-Associated Small Interfering RNA-Mediated Silencing Pathway Downregulates Complementary Sense Gypsy Transcripts in Somatic Cells of the Drosophila Ovary ». *Journal of Virology* 81 (4): 1951–60.
- 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.



- Perrini, Barbara, Lucia Piacentini, Laura Fanti, Fabio Altieri, Silvia Chichiarelli, Maria Berloco, Carlo Turano, Anna Ferraro, et Sergio Pimpinelli. 2004. « HP1 Controls Telomere Capping, Telomere Elongation, and Telomere Silencing by Two Different Mechanisms in *Drosophila* ». *Molecular Cell* 15 (3): 467–476. doi:10.1016/j.molcel.2004.06.036.
- Pöyhönen, Maria, Augustin de Vanssay, Valérie Delmarre, Catherine Hermant, Anne Laure Todeschini, Laure Teyssset, et Stéphane Ronsseray. 2012. « Homology-Dependent Silencing by an Exogenous Sequence in the *Drosophila* Germline ». *G3: Genes\textbarGenomes\textbarGenetics* 2 (3): 331–338. doi:10.1534/g3.111.001925.
- 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* 18 (8): 1446–1457. doi:10.1261/rna.034405.112.
- Qi, Yijun, Ahmet M. Denli, et Gregory J. Hannon. 2005. « Biochemical Specialization within *Arabidopsis* RNA Silencing Pathways ». *Molecular Cell* 19 (3): 421–28.
- Rajasethupathy, Priyamvada, Igor Antonov, Robert Sheridan, Sebastian Frey, Chris Sander, Thomas Tuschl, et Eric R. Kandel. 2012. « A Role for Neuronal piRNAs in the Epigenetic Control of Memory-Related Synaptic Plasticity ». *Cell* 149 (3): 693–707. doi:10.1016/j.cell.2012.02.057.
- Rangan, Prashanth, Matthew DeGennaro, Kean Jaime-Bustamante, Rémi-Xavier Coux, Rui G. Martinho, et Ruth Lehmann. 2009. « Temporal and Spatial Control of Germ-Plasm RNAs ». *Current Biology: CB* 19 (1): 72–77. doi:10.1016/j.cub.2008.11.066.
- 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): 1373–1379. doi:10.1016/j.cub.2011.06.057.
- 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): 1839–1855.
- Rongo, C., H. T. Broihier, L. Moore, M. Van Doren, A. Forbes, et R. Lehmann. 1997. « Germ Plasm Assembly and Germ Cell Migration in *Drosophila* ». *Cold Spring Harbor Symposia on Quantitative Biology* 62: 1–11.
- 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): 1631–1642.
- 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): 1663–1674.
- 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): 115–123.
- Ronsseray, S., L. Marin, M. Lehmann, et D. Anxolabéhère. 1998. « Repression of Hybrid Dysgenesis in *Drosophila Melanogaster* by Combinations of Telomeric P-Element Reporters and Naturally Occurring P Elements ». *Genetics* 149 (4): 1857–1866.
- Roovers, Elke F., David Rosenkranz, Mahdi Mahdipour, Chung-Ting Han, Nannan He, Susana M. Chuva de Sousa Lopes, Lucette A. J. van der Westerlaken, et al. 2015. « Piwi Proteins and piRNAs in Mammalian Oocytes and Early Embryos ». *Cell Reports* 10 (12): 2069–2082. doi:10.1016/j.celrep.2015.02.062.
- Rosenkranz, David, Chung-Ting Han, Elke F. Roovers, Hans Zischler, et René F. Ketting. 2015. « Piwi Proteins and piRNAs in Mammalian Oocytes and Early Embryos: From

- Sample to Sequence ». *Genomics Data* 5 (septembre): 309–313.  
doi:10.1016/j.gdata.2015.06.026.
- Rothe, M., M. Pehl, H. Taubert, et H. Jäckle. 1992. « Loss of Gene Function through Rapid Mitotic Cycles in the *Drosophila* Embryo ». *Nature* 359 (6391): 156–159.  
doi:10.1038/359156a0.
- Rouget, Christel, Catherine Papin, Anthony Boureux, Anne-Cécile Meunier, Bénédicte Franco, Nicolas Robine, Eric C. Lai, Alain Pelisson, et Martine Simonelig. 2010. « Maternal mRNA Deadenylation and Decay by the piRNA Pathway in the Early *Drosophila* Embryo ». *Nature* 467 (7319): 1128–1132. doi:10.1038/nature09465.
- Rozhkov, Nikolay V., Molly Hammell, et Gregory J. Hannon. 2013. « Multiple roles for Piwi in silencing *Drosophila* transposons ». *Genes & Development* 27 (4): 400–412.  
doi:10.1101/gad.209767.112.
- Ryder, Edward, Fiona Blows, Michael Ashburner, Rosa Bautista-Llacer, Darin Coulson, Jenny Drummond, Jane Webster, et al. 2004. « The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. » *Genetics* 167 (2): 797–813. doi:10.1534/genetics.104.026658.
- 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): 1296–1299. doi:10.1038/nature08501.
- Saito, Kuniaki, Hirotsugu Ishizu, Miharuru 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): 2493–2498. 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.
- 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): 1603–1608. doi:10.1101/gad.1563607.
- Santos, Ana C., et Ruth Lehmann. 2004. « Germ Cell Specification and Migration in *Drosophila* and beyond ». *Current Biology: CB* 14 (14): R578–589.  
doi:10.1016/j.cub.2004.07.018.
- Sarot, Emeline, Geneviève Payen-Groschêne, Alain Bucheton, et Alain Pelisson. 2004. « Evidence for a Piwi-Dependent RNA Silencing of the Gypsy Endogenous Retrovirus by the *Drosophila* *Melanogaster* Flamenco Gene ». *Genetics* 166 (3): 1313–1321.
- Sato, Kaoru, Yuka W. Iwasaki, Aoi Shibuya, Piero Carninci, Yuuta Tsuchizawa, Hirotsugu Ishizu, Mikiko C. Siomi, et Haruhiko Siomi. 2015. « Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the *Drosophila* Germline ». *Molecular Cell* 59 (4): 553–563. doi:10.1016/j.molcel.2015.06.024.
- Sato, Kaoru, Kazumichi M. Nishida, Aoi Shibuya, Mikiko C. Siomi, et Haruhiko Siomi. 2011. « Maelstrom Coordinates Microtubule Organization during *Drosophila* Oogenesis through Interaction with Components of the MTOC ». *Genes & Development* 25 (22): 2361–2373. doi:10.1101/gad.174110.111.
- Savitsky, Mikhail, Oksana Kravchuk, Larisa Melnikova, et Pavel Georgiev. 2002. « Heterochromatin Protein 1 Is Involved in Control of Telomere Elongation in

- Drosophila melanogaster* ». *Molecular and Cellular Biology* 22 (9): 3204–3218. doi:10.1128/MCB.22.9.3204-3218.2002.
- Schoeberl, Ursula E., Henriette M. Kurth, Tomoko Noto, et Kazufumi Mochizuki. 2012. « Biased Transcription and Selective Degradation of Small RNAs Shape the Pattern of DNA Elimination in Tetrahymena ». *Genes & Development* 26 (15): 1729–42.
- Senti, Kirsten-André, Daniel Jurczak, Ravi Sachidanandam, et Julius Brennecke. 2015. « piRNA-Guided Slicing of Transposon Transcripts Enforces Their Transcriptional Silencing via Specifying the Nuclear piRNA Repertoire ». *Genes & Development* 29 (16): 1747–1762. doi:10.1101/gad.267252.115.
- Seydoux, G., et M. A. Dunn. 1997. « Transcriptionally Repressed Germ Cells Lack a Subpopulation of Phosphorylated RNA Polymerase II in Early Embryos of *Caenorhabditis Elegans* and *Drosophila Melanogaster* ». *Development (Cambridge, England)* 124 (11): 2191–2201.
- Shermoen, A. W., et P. H. O’Farrell. 1991. « Progression of the Cell Cycle through Mitosis Leads to Abortion of Nascent Transcripts ». *Cell* 67 (2): 303–310.
- Siddiqui, Najeeb U., Xiao Li, Hua Luo, Angelo Karaiskakis, Huayun Hou, Thomas Kislinger, J. Timothy Westwood, Quaid Morris, et Howard D. Lipshitz. 2012. « Genome-Wide Analysis of the Maternal-to-Zygotic Transition in *Drosophila* Primordial Germ Cells ». *Genome Biology* 13 (2): R11. doi:10.1186/gb-2012-13-2-r11.
- Sienski, Grzegorz, Julia Batki, Kirsten-André Senti, Derya Dönertas, Laszlo Tirian, Katharina Meixner, et Julius Brennecke. 2015. « Silencio/CG9754 connects the Piwi–piRNA complex to the cellular heterochromatin machinery ». *Genes & Development* 29 (21): 2258–2271. doi:10.1101/gad.271908.115.
- 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): 964–980. doi:10.1016/j.cell.2012.10.040.
- Simmons, M. J., J. D. Raymond, C. D. Grimes, C. Belinco, B. C. Haake, M. Jordan, C. Lund, T. A. Ojala, et D. Papermaster. 1996. « Repression of Hybrid Dysgenesis in *Drosophila Melanogaster* by Heat-Shock-Inducible Sense and Antisense P-Element Constructs ». *Genetics* 144 (4): 1529–1544.
- Simonelig, Martine. 2012. « Maternal-to-Zygotic Transition: Soma versus Germline ». *Genome Biology* 13 (2): 145. doi:10.1186/gb-2012-13-2-145.
- Siomi, Haruhiko, et Mikiko C. Siomi. 2015. « RNA. Phased piRNAs Tackle Transposons ». *Science (New York, N.Y.)* 348 (6236): 756–757. doi:10.1126/science.aab3004.
- Slaidina, Maija, et Ruth Lehmann. 2014. « Translational Control in Germline Stem Cell Development ». *The Journal of Cell Biology* 207 (1): 13–21. doi:10.1083/jcb.201407102.
- Soares, Zamira Guerra, André Nicolau Aquime Gonçalves, Karla Pollyanna Vieira de Oliveira, et João Trindade Marques. 2014. « Viral RNA recognition by the *Drosophila* small interfering RNA pathway ». *Microbes and Infection, Special Issue: Pathogen nucleic acid sensing by innate immune receptors*, 16 (12): 1013–21.
- Song, Ji-Joon, Stephanie K. Smith, Gregory J. Hannon, et Leemor Joshua-Tor. 2004. « Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity ». *Science (New York, N.Y.)* 305 (5689): 1434–1437. doi:10.1126/science.1102514.
- Song, S. U., T. Gerasimova, M. Kurkulos, J. D. Boeke, et V. G. Corces. 1994. « An Env-like Protein Encoded by a *Drosophila* Retroelement: Evidence That Gypsy Is an Infectious Retrovirus ». *Genes & Development* 8 (17): 2046–57.
- Song, Yan, Lanette Fee, Tammy H. Lee, et Robin P. Wharton. 2007. « The Molecular Chaperone Hsp90 Is Required for mRNA Localization in *Drosophila Melanogaster* Embryos ». *Genetics* 176 (4): 2213–2222. doi:10.1534/genetics.107.071472.

- Spradling, A. C., et G. M. Rubin. 1982. « Transposition of Cloned P Elements into *Drosophila* Germ Line Chromosomes ». *Science (New York, N.Y.)* 218 (4570): 341-47.
- Spradling, Allan C., Hugo J. Bellen, et Roger A. Hoskins. 2011. « *Drosophila* P Elements Preferentially Transpose to Replication Origins ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (38): 15948–15953. doi:10.1073/pnas.1112960108.
- Swaminathan, Jyothishmathi, Ellen M. Baxter, et Victor G. Corces. 2005. « The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of *Drosophila* heterochromatin ». *Genes & Development* 19 (1): 65–76. doi:10.1101/gad.1259105.
- Swarts, Daan C., Matthijs M. Jore, Edze R. Westra, Yifan Zhu, Jorijn H. Janssen, Ambrosius P. Snijders, Yanli Wang, et al. 2014. « DNA-Guided DNA Interference by a Prokaryotic Argonaute ». *Nature* 507 (7491): 258-61.
- Talbert, Paul B., et Steven Henikoff. 2010. « Histone Variants—ancient Wrap Artists of the Epigenome ». *Nature Reviews. Molecular Cell Biology* 11 (4): 264–275. doi:10.1038/nrm2861.
- Tange, Thomas Ø, Toshiharu Shibuya, Melissa S. Jurica, et Melissa J. Moore. 2005. « Biochemical Analysis of the EJC Reveals Two New Factors and a Stable Tetrameric Protein Core ». *RNA (New York, N.Y.)* 11 (12): 1869–1883. doi:10.1261/rna.2155905.
- Tautz, D. 1988. « Regulation of the *Drosophila* Segmentation Gene Hunchback by Two Maternal Morphogenetic Centres ». *Nature* 332 (6161): 281–284. doi:10.1038/332281a0.
- Tautz, D., et C. Pfeifle. 1989. « A Non-Radioactive in Situ Hybridization Method for the Localization of Specific RNAs in *Drosophila* Embryos Reveals Translational Control of the Segmentation Gene Hunchback ». *Chromosoma* 98 (2): 81–85.
- Teyssset, L., J. C. Burns, H. Shike, B. L. Sullivan, A. Bucheton, et C. Terzian. 1998. « A Moloney Murine Leukemia Virus-Based Retroviral Vector Pseudotyped by the Insect Retroviral Gypsy Envelope Can Infect *Drosophila* Cells ». *Journal of Virology* 72 (1): 853-56.
- 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). doi:10.1371/journal.pone.0011032.
- Tolia, Niraj H., et Leemor Joshua-Tor. 2007. « Slicer and the Argonautes ». *Nature Chemical Biology* 3 (1): 36–43. doi:10.1038/nchembio848.
- Upton, Kyle R., Daniel J. Gerhardt, J. Samuel Jesuadian, Sandra R. Richardson, Francisco J. Sánchez-Luque, Gabriela O. Bodea, Adam D. Ewing, et al. 2015. « Ubiquitous L1 Mosaicism in Hippocampal Neurons ». *Cell* 161 (2): 228-39.
- Vagin, Vasily V., Alla Sigova, Chengjian Li, Hervé Seitz, Vladimir Gvozdev, et Phillip D. Zamore. 2006. « A Distinct Small RNA Pathway Silences Selfish Genetic Elements in the Germline ». *Science* 313 (5785): 320-24.
- Vagin, Vasily V., Yang Yu, Anna Jankowska, Yicheng Luo, Kaja A. Wasik, Colin D. Malone, Emily Harrison, et al. 2013. « Minotaur is critical for primary piRNA biogenesis ». *RNA* 19 (8): 1064–1077. doi:10.1261/rna.039669.113.
- van Daal, A., et S. C. Elgin. 1992. « A Histone Variant, H2AvD, Is Essential in *Drosophila Melanogaster* ». *Molecular Biology of the Cell* 3 (6): 593–602.
- Van Doren, M., A. L. Williamson, et R. Lehmann. 1998a. « Regulation of Zygotic Gene Expression in *Drosophila* Primordial Germ Cells ». *Current Biology: CB* 8 (4): 243–246.

- Van Doren, Mark, Wendy R. Mathews, Monique Samuels, Lisa A. Moore, Heather Tarczy Broihier, et Ruth Lehmann. 2003. « Fear of Intimacy Encodes a Novel Transmembrane Protein Required for Gonad Morphogenesis in *Drosophila* ». *Development (Cambridge, England)* 130 (11): 2355–2364.
- 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.
- Voronina, Ekaterina, Geraldine Seydoux, Paolo Sassone-Corsi, et Ippei Nagamori. 2011. « RNA Granules in Germ Cells ». *Cold Spring Harbor Perspectives in Biology* 3 (12). doi:10.1101/cshperspect.a002774.
- Vourekas, Anastassios, Panagiotis Alexiou, Nicholas Vrettos, Manolis Maragkakis, et Zissimos Mourelatos. 2016. « Sequence-Dependent but Not Sequence-Specific piRNA Adhesion Traps mRNAs to the Germ Plasm ». *Nature* 531 (7594): 390–394. doi:10.1038/nature17150.
- Walter, M. F., C. Jang, B. Kasravi, J. Donath, B. M. Mechler, J. M. Mason, et H. Biessmann. 1995. « DNA Organization and Polymorphism of a Wild-Type *Drosophila* Telomere Region ». *Chromosoma* 104 (4): 229–241.
- Wang, Guilin, et Valerie Reinke. 2008. « A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis ». *Current biology : CB* 18 (12): 861–867. doi:10.1016/j.cub.2008.05.009.
- Wang, Hui, Zaijun Ma, Kongyan Niu, Yi Xiao, Xiaofen Wu, Chenyu Pan, Yun Zhao, Kai Wang, Yaoyang Zhang, et Nan Liu. 2016. « Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila* ». *Development (Cambridge, England)* 143 (3): 530–539. doi:10.1242/dev.128116.
- 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): 21164–21169. doi:10.1073/pnas.1107892109.
- Wang, Wei, Bo W. Han, Cindy Tipping, Daniel Tianfang Ge, Zhao Zhang, Zhiping Weng, et Phillip D. Zamore. 2015. « Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms ». *Molecular Cell* 59 (5): 819–830. doi:10.1016/j.molcel.2015.08.007.
- Webster, Alexandre, Sisi Li, Junho K. Hur, Malte Wachsmuth, Justin S. Bois, Edward M. Perkins, Dinshaw J. Patel, et Alexei A. Aravin. 2015. « Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper ». *Molecular Cell* 59 (4): 564–575. doi:10.1016/j.molcel.2015.07.017.
- Williamson, A., et R. Lehmann. 1996. « Germ Cell Development in *Drosophila* ». *Annual Review of Cell and Developmental Biology* 12: 365–391. doi:10.1146/annurev.cellbio.12.1.365.
- Wolfe, Kenneth H., David Armisén, Estelle Proux-Wera, Seán S. ÓhÉigeartaigh, Haleema Azam, Jonathan L. Gordon, et Kevin P. Byrne. 2015. « Clade- and species-specific features of genome evolution in the Saccharomycetaceae ». *FEMS Yeast Research* 15 (5). doi:10.1093/femsyr/fov035.
- 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.
- Zalokar, M. 1976. « Autoradiographic Study of Protein and RNA Formation during Early Development of *Drosophila* Eggs ». *Developmental Biology* 49 (2): 425–437.
- 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): 4039–4050. 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): 871–884. doi:10.1016/j.cell.2012.09.040.
- Zhang, Liang, Michelle Beaucher, Yan Cheng, et Yikang S. Rong. 2014. « Coordination of Transposon Expression with DNA Replication in the Targeting of Telomeric Retrotransposons in *Drosophila* ». *The EMBO Journal* 33 (10): 1148–1158. doi:10.1002/embj.201386940.
- 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): 1353–1363. doi:10.1016/j.cell.2014.04.030.
- Zhang, Zhao, Jia Xu, Birgit S. Koppetsch, Jie Wang, Cindy Tipping, Shengmei Ma, Zhiping Weng, William E. Theurkauf, et Phillip D. Zamore. 2011. « Heterotypic piRNA Ping-Pong Requires Qin, a Protein with Both E3 Ligase and Tudor Domains ». *Molecular Cell* 44 (4): 572–584. doi:10.1016/j.molcel.2011.10.011.



## VI – Annexe

**Tableau 2 - Liste des souches testées dans le cadre du crible.** En bleu les lignées Enhancer-trap lacZ sélectionnées par coloration Xgal (« ubiquitaire »). En orange les lignées Enhancer-trap Gal4 sélectionnées par coloration Xgal (« ubiquitaire ») Voir Résultats.

Code Flybase	Numéro de stock	Génotype	Type d'insertion	Gène affecté par l'insertion
FBst0302438	103569	<i>y<sup>*</sup> w<sup>*</sup>; P{GawB}ex<sup>NP0254</sup> / CyO, P{UAS-lacZ.UW14}UW14</i>	Gal4	ex
FBst0315087	112347	<i>P{GawB}NP0818 w<sup>*</sup> / FM7c</i>	Gal4	Inconnu
FBst0303338	104472	<i>w<sup>*</sup>; P{GawB}Hsp68NP3363</i>	Gal4	Hsp68
FBst0012542	12542	<i>w<sup>1118</sup>; P{GT1}Hsp23<sup>BG01483</sup></i>	Gal4	Hsp23
FBst0304347	105489	<i>y<sup>*</sup> w<sup>*</sup>; P{GawB}mfrn<sup>NP7499</sup> / TM6, P{UAS-lacZ.UW23-1}UW23-1</i>	Gal4	mfrn
FBst0315040	112300	<i>P{GawB}sgg<sup>NP0715</sup> w<sup>*</sup> / FM7c</i>	Gal4	sgg
FBst0316836	114115	<i>P{w[+mW.hs]=GawB}sgg[NP7167] w[*] / FM7c</i>	Gal4	sgg
FBst0316821	114100	<i>P{GawB}sgg<sup>NP7142</sup> w<sup>*</sup> / FM7c</i>	Gal4	sgg
FBst0314794	112052	<i>P{GawB}sgg<sup>NP0122</sup> w<sup>*</sup> / FM7c</i>	Gal4	sgg
FBst0302705	103837	<i>w<sup>*</sup>; P{GawB}NP0967 / CyO</i>	Gal4	Inconnu
FBst0302970	104103	<i>w<sup>*</sup>; P{GawB}Nmdmc<sup>NP2138</sup> / TM3, Sb<sup>l</sup> Ser<sup>l</sup></i>	Gal4	Nmdmc
FBst0302620	103752	<i>w<sup>*</sup>; P{GawB}CG2310<sup>NP0767</sup> / TM3, Ser<sup>l</sup></i>	Gal4	CG2310
FBst0302929	104062	<i>w<sup>*</sup>; P{GawB}NP2010 / CyO</i>	Gal4	Inconnu
FBst0316941	114221	<i>w<sup>*</sup>; P{GawB}chic<sup>NP7317</sup> / CyO, P{UAS-lacZ.UW14}UW14</i>	Gal4	Inconnu
FBst0303674	104811	<i>y<sup>*</sup> P{GawB}NP5028 w<sup>*</sup> / FM7c</i>	Gal4	Inconnu
FBst0303223	104356	<i>y<sup>*</sup> w<sup>*</sup>; P{GawB}Prosa5<sup>NP3047</sup> / CyO, P{UAS-lacZ.UW14}UW14</i>	Gal4	Prosa5
FBst0304258	105400	<i>y<sup>*</sup> w<sup>*</sup> P{GawB}dlgINP7225</i>	Gal4	dlg1
FBst0303481	104616	<i>y<sup>*</sup> w<sup>*</sup> P{GawB}dlgI<sup>NP4134</sup> / FM7c</i>	Gal4	dlg1
FBst0303788	104929	<i>y<sup>*</sup> w<sup>*</sup>; P{GawB}NP5267 / CyO, P{UAS-lacZ.UW14}UW14</i>	Gal4	Inconnu
FBst0013122	13122	<i>w<sup>1118</sup>; P{GT1}CG17618<sup>BG02691</sup> CG6982<sup>BG02691</sup></i>	Gal4	CG6982, CG17618
FBst0012771	12771	<i>w<sup>1118</sup>; P{GT1}BG02176/CyO</i>	Gal4	Prosa5
FBst0012450	12450	<i>w<sup>1118</sup>; P{GT1}CG3964<sup>BG00933</sup></i>	Gal4	CG3964
FBst0011877	11877	<i>w<sup>67c23</sup> P{lacW}Kmn1<sup>G0237</sup>/FM7c</i>	placZ	Kmn1
FBst0012179	12179	<i>y<sup>l</sup> w<sup>67c23</sup>; P{lacW}mRpL4<sup>k14608</sup>/CyO</i>	placZ	Ecol lacZ, mRpL4
FBst0012259	12259	<i>P{lacW}l(1)G0451<sup>G0451</sup> w<sup>67c23</sup>/FM7c</i>	placZ	l(1)G0451, RpL22
FBst0012253	12253	<i>P{lacW}l(1)G0422<sup>G0422</sup> w<sup>67c23</sup>/FM7c</i>	placZ	l(1)G0422, RpL22



FBst0325542	122105	$w^*$ ; $P\{lacW\}l(2)SH1594SH1594$	placZ	Ehbp1, l(2)SH1594
FBst0010931	10931	$y^I w^{67c23}$ ; $P\{lacW\}Ehbp1^{k09837}/CyO$	placZ	Ecol.lacZ,Ehbp1,l(2)k0983 7
FBst0011973	11973	$P\{lacW\}sgg^{G0335} w^{67c23}/FM7c$	placZ	sgg
FBst0010092	10092	$P\{lacW\}sgg^{G0183} w^{67c23}/FM7c$	placZ	sgg
FBst0011845	11845	$P\{lacW\}sgg^{G0263} w^{67c23}/FM7c$	placZ	sgg
FBst0011596	11596	$P\{lacW\}sgg^{G0055} w^{67c23}/FM7c$	placZ	sgg
FBst0010630	10630	$y^I w^{67c23}$ ; $P\{lacW\}l(2)^{k06506}/CyO$	placZ	Ecol\lacZ, l(2)k06506
FBst0010667	10667	$y^I w^{67c23}$ ; $P\{lacW\}l(2)^{k07308}/CyO$	placZ	Ecol\lacZ, l(2)k07308
FBst0011067	11067	$y^I w^{67c23}$ ; $P\{lacW\}ex^{k12913}/CyO$	placZ	Ecol\lacZ, ex
FBst0010651	10651	$y^I w^{67c23}$ ; $P\{lacW\}l(2)^{k07005}/CyO$	placZ	Ecol\lacZ, l(2)k07005
FBst0012049	12049	$w^{67c23} P\{lacW\}l(1)^{G0444}/FM7c$	placZ	Kmn1, l(1)G0444
FBst0011823	11823	$P\{lacW\}l(1)^{G0115} w^{67c23}/FM7c$	placZ	l(1)G0115, RpL22

**De l'œuf à l'adulte : étude moléculaire et fonctionnelle de la répression des éléments transposables par les piARN au cours du développement chez *Drosophila melanogaster***

Chez les métazoaires, la mobilisation des éléments transposables est régulée par de petits ARN non codants appelés piARN pour "PIWI interacting RNA". Cette répression est très étudiée dans la lignée germinale adulte où elle est particulièrement efficace. Néanmoins, la mobilisation de ces éléments doit être régulée tout au long du développement de la lignée germinale, qui transmet l'information génétique à travers les générations.

Durant ma thèse, j'ai utilisé le modèle *D. melanogaster* pour étudier cette répression au cours du développement de la lignée germinale femelle. J'ai ainsi pu montrer qu'une répression fonctionnelle par les piARN existe dès la fin de l'embryogenèse et que les gènes liés à la régulation chez l'adulte sont également nécessaires pour la répression au cours du développement. L'analyse de données de séquençage haut débit m'a permis de mettre en évidence la production *de novo* de piARN fonctionnels dans les gonades en formation. De plus, comme dans les ovaires adultes, une répression incomplète ressemblant à la variégation, existe à tous les stades du développement. Des expériences de lignage cellulaire suggèrent fortement qu'une mémoire épigénétique précoce est initiée dans les cellules germinales embryonnaires et maintenue jusqu'au stade adulte. L'implication de l'Heterochromatin Protein 1a (HP1a) dans la production des piARN télomériques montrée par séquençage des piARN pourrait expliquer ce phénomène. L'étude de la variégation permettrait de mieux comprendre la régulation des gènes. En effet, en dehors de leur rôle de répression des éléments transposables, les piARN régulent aussi l'expression de gènes codants.

Mots clés : [Élément transposable ; Développement ; Petits ARN non codants ; piARN ; Drosophile]

**From egg to adult: molecular and functional study of piRNA-mediated repression during germline development in *Drosophila melanogaster***

In metazoan germ cells, transposable element activity is repressed by small noncoding PIWI-associated RNAs (piRNAs). Numerous investigations in *Drosophila* have enlightened the mechanism of this repression in the adult germline. However, very little is known about piRNA-mediated repression during germline development. Nevertheless, to maintain the integrity of the genome, repression should occur throughout the lifespan of germ cells.

During my PhD, I show that piRNA-mediated repression is active in the female germline, from late embryonic to pupal primordial germ cells, and that genes related to the adult piRNA pathway are required for repression during development. *rhino*-dependent piRNAs, exhibiting the molecular signature of the piRNA pathway "ping-pong" amplification step, are detected in larval gonads, arguing for *de novo* biogenesis of functional piRNAs during development. I also show that production of telomeric piRNAs depends on Heterochromatin Protein 1a (HP1a). Furthermore, as in adult ovaries, I observe an incomplete, bimodal and stochastic repression resembling variegation at all developmental stages. Clonal analyzes of this incomplete silencing strongly suggest that a cellular memory of an early repression decision is initiated in embryonic germ cells and further maintained until the adult stage. Taken together, the data presented here show that piRNAs and their associated proteins are epigenetic components of a continuous repression system throughout germ cell development.

Keywords: [Transposable elements; Development; small non-coding RNAs; piRNAs; Drosophila]