



HAL
open science

Epigenetic control of the ovarian transcriptome in *Drosophila melanogaster*

Daniela Torres Campana

► **To cite this version:**

Daniela Torres Campana. Epigenetic control of the ovarian transcriptome in *Drosophila melanogaster*. Genetics. Université de Lyon, 2021. English. NNT : 2021LYSE1176 . tel-03793803

HAL Id: tel-03793803

<https://theses.hal.science/tel-03793803>

Submitted on 2 Oct 2022

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

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



N°d'ordre NNT :2021LYSE1176

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

Ecole Doctorale N° 340
Biologie Moléculaire, Intégrative et Cellulaire

Spécialité de doctorat : Biologie moléculaire et cellulaire
Discipline : Épigenétique

Soutenue publiquement le 28/09/2021, par :
Daniela TORRES CAMPANA

Contrôle épigénétique du transcriptome ovarien chez *Drosophila melanogaster*

Devant le jury composé de :

VIEIRA-HEDDI , Cristina Professeure des Universités, Université Lyon 1, Lyon	Présidente
DI STEFANO , Luisa, Chargée de recherche CNRS, Centre de Biologie Intégrative, Toulouse	Rapporteuse
MARTINEZ , Anne-Marie, Professeure des Universités, Université de Montpellier, Montpellier	Rapporteuse
DOSTATNI , Nathalie, Professeure des Universités, Sorbonne Université, Paris	Examinatrice
LOPPIN Benjamin , Directeur de Recherche CNRS, Laboratoire de Biologie et Modélisation de la Cellule, Lyon	Directeur de thèse
ORSI Guillermo , Chargé de recherche, INSERM Laboratoire de Biologie et Modélisation de la Cellule, Lyon	Co-directeur de thèse

Université Claude Bernard – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

Remerciements

Après une explosion, un déménagement et une pandémie, elle est là cette thèse et je tiens à remercier les personnes qui ont fait ça possible.

Tout d'abord je veux remercier mon directeur, Benjamin Loppin, de m'avoir ouvert les portes de l'équipe Épigénétique et Formation du Zygote pour la réalisation de ma thèse. Je le remercie de m'avoir proposé un projet si intéressant qui m'a passionné du début à la fin.

Ensuite, je tiens à remercier mon co-directeur, Guille. (El gracias se queda corto para todo lo que hiciste por mí durante estos años). Merci pour ton encadrement de qualité, ton soutien, ta patience, ta pédagogie et ta bonne humeur. Ce fût une expérience très enrichissante d'avoir travaillé à tes côtés. Je te souhaite le meilleur pour le #OrsiLab.

Je veux également remercier Béa qui, pour l'anecdote, fût la première personne à m'avoir parlé d'épigénétique quand j'étais une jeune étudiante en première année de licence. Merci pour ton accompagnement, tes conseils et ton aide au cours de ces années. Merci également à Raphaëlle d'avoir pris le temps de répondre à mes questions, d'avoir passé mes commandes etc. Enfin, merci à vous deux de m'avoir intégré dans cette belle équipe.

Je remercie également les anciens membres de l'équipe. Plus particulièrement Daniel qui est arrivé au bon moment pour m'aider à démarrer cette aventure qu'est la thèse et qui s'est converti en un grand ami. Merci aussi à Fabiola, qui a rajouté une dose de bonne humeur aux journées au labo. J'ai mangé des bons burgers avec cette Latino Team mais j'ai surtout vécu des moments inoubliables avec vous.

Mes chers collègues du M4 RdC, merci d'avoir créé une super ambiance au labo. Je parle, en partie, des gens qui sont venus envahir mon bureau, Marion, Jérôme, Agnès, mais avec qui, malgré la pandémie, j'ai partagé de très bons moments. Merci Jérôme pour avoir égayé les longues journées au labo et pour les nombreuses conversations pouvant aller du fruit préféré jusqu'aux questions les plus philosophiques. Merci Marion, déjà, pour avoir été notre guide LBMC et surtout pour lâcher des phrases mythiques que je ne retranscrirais pas. Merci Victor pour ton soutien et ta patience surtout pendant la rédaction de ce manuscrit.

Quiero por supuesto agradecer a mi familia y amigos quienes, de cerca o de lejos, siempre me han brindado su apoyo.

Agradezco a mis padres por su incondicional apoyo, por enseñarme a perseguir mis sueños y por no dejar que un océano les impida estar siempre a mi lado.

Finalmente, quiero agradecer a la persona que ha estado siempre a mi lado apoyándome, aguantándome, consolándome, ayudándome y un sinnúmero mas de cosas: mi hermana. Gracias por ser mi hogar durante todos estos años y por demostrarme lo que es la valentía.

La thèse, comme la transcription, est un processus complexe, nécessitant de plusieurs étapes et acteurs et rencontrant souvent des obstacles à contourner. Néanmoins ça reste une période riche en créations et en rencontres. Pour moi, la thèse, fût non seulement une période d'intense apprentissage scientifique mais également de croissance personnelle qui m'a laissé beaucoup de leçons. C'est une période qui nous met à l'épreuve et qui n'est pas simple mais qu'avec les bonnes personnes (mentionnées ci-dessus) peut devenir une expérience épanouissante comme elle le fût pour moi.

¡ Gracias totales!

*To my sister,
It got better.*

« La curiosité est nécessaire à la nature humaine, sans elle, on croupirait dans l'ignorance » -Le Taureau Blanc, Voltaire

Résumé :

La régulation dynamique de l'épigénome est à la base de la plasticité cellulaire et permet aux cellules de répondre aux programmes développementaux et de différenciation. Les effecteurs épigénomiques entraînent des changements tels que la modification des queues des histones, le positionnement des histones ou l'incorporation de variants d'histones, ce qui impacte la structure de la chromatine et donc l'expression génique. Cependant, les mécanismes moléculaires sous-jacents restent, en grande partie, inconnus. Un défi majeur du domaine est l'établissement de liens fonctionnels directs entre modificateurs épigénétiques, qui ont tendance à avoir un effet global sur le génome, et expression de gènes effecteurs.

Mon projet de thèse s'est justement basé sur la découverte d'une connexion fonctionnelle entre une série d'effecteurs épigénomiques et un effecteur de la formation du zygote chez *Drosophila melanogaster* : la thiorédoxine maternelle Deadhead (Dhd). Dhd est cruciale pour le remodelage de la chromatine paternelle à la fécondation et donc pour la formation du zygote.

Tout d'abord, un crible génétique basé sur l'expression de shRNAs dans la lignée germinale femelle m'a permis d'identifier : (i) l'histone demethylase Lid/dKDM5 (ii) les membres du complexe déacétylase Sin3A et Rpd3 (iii) la sous-unité Snr1 du remodeleur de la chromatine Swi/Snf et (iv) le facteur chromatinien Mod(mdg4), comme étant essentiels pour l'expression de *dhd*. Pour la suite des analyses je me suis focalisée sur Lid, Sin3A, Snr1 et Mod(mdg4). Des analyses transcriptomiques ont montré que *dhd* est parmi les gènes les plus exprimés dans les ovaires et que cette expression est abolie lors du *knockdown* dans la lignée germinale femelle de *lid*, *sin3a*, *snr1* ou *mod(mdg4)*. De façon remarquable la quantité de gènes dérégulés dans les ovaires *knockdown* est limitée. Ceci suggère que ces complexes ubiquitaires et conservés sont dédiés à la régulation de *dhd* dans ce tissu. Ce cas paradigmatique m'a donc offert une opportunité pour disséquer les mécanismes moléculaires agissant sur le contrôle épigénomique de l'établissement de programmes transcriptionnels spécifiques, comme c'est le cas lors de l'ovogénèse.

Ensuite, en utilisant la méthode de profilage de la chromatine Cut&Run et une stratégie d'analyse de données dédiée, j'ai trouvé que *dhd* est intégré dans un mini-domaine enrichi en marques hétérochromatiques H3K27me3/H3K9me3 et délimité par des éléments régulateurs.

De plus, l'élément régulateur à proximité du promoteur de *dhd* s'est avéré essentiel pour son expression. De façon surprenante, Lid, Sin3A, Snr1 et Mod(mdg4) ont des effets différents sur H3K27me3 et sur ces éléments régulateurs. Néanmoins, j'ai mis en évidence que ces effecteurs activent *dhd* de façon indépendante de H3K27me3/H3K9me3. De plus, j'ai trouvé que ces marques ne sont pas nécessaires pour réprimer *dhd* dans les tissus adultes.

Mes travaux de thèse ont donc révélé plusieurs caractéristiques inhabituelles, au niveau génomique et épigénomique, au locus *dhd*. Cependant je n'ai pas établi un trait unique qui conduise à l'hyperactivation de *dhd* dans la lignée germinale femelle. La régulation de *dhd* dépendrait non pas d'une seule caractéristique mais plutôt d'une combinaison de traits particuliers.

À travers l'exemple de *dhd*, mes travaux démontrent la complexité des processus qu'implique l'activation d'un gène au bon endroit, au bon moment et en bonne quantité. Ils illustrent également la difficulté d'établir des règles générales sur les mécanismes de régulation transcriptionnelle médiés par la chromatine, qui sont très souvent contexte-spécifique.

Mots clés: épigénétique, épigénomique, régulation transcriptionnelle, ovogénèse, *Drosophila melanogaster*, Deadhead, transition ovocyte-zygote

Summary:

Dynamic regulation of the epigenome underlies cellular plasticity and allows cells to respond to developmental and differentiation programs. Epigenomic effectors can mediate changes such as modification of histone tails, nucleosome positioning or the incorporation of specific histone variants, altering chromatin structure and thus gene expression. Yet, the molecular mechanisms underpinning these processes remain largely unknown. A major challenge in the field is to establish direct functional connections between upstream chromatin factors, which generally have broad impact on chromatin, and the controlled expression of specific cellular effectors.

My PhD project was precisely based on the discovery of a specific functional link between a series of epigenomic effectors and the highly regulated terminal effector of zygote formation in *Drosophila melanogaster*: the maternal thioredoxin Deadhead (Dhd). Dhd is critical to ensure paternal chromatin remodeling at fertilization, and thus zygote formation.

First, an shRNA-based genetic screen in the female germline led me to identify (i) the H3K4me3 demethylase Lid, (ii) the members of the deacetylase complex Sin3A and Rpd3, (iii) the Snr1 subunit of the chromatin remodeler Swi/Snf and (iv) the chromatin factor Mod(mdg4), as essential for the expression of *dhd*. For further analyses I focused on Lid, Sin3A, Snr1 and Mod(mdg4). Transcriptomic analyses showed that *dhd* is among the most highly expressed genes in ovaries and this expression is completely abolished when I deplete Lid, Sin3A, Snr1 or Mod(mdg4) specifically in the female germline. Remarkably, there is a paucity of misregulated genes in *knockdown* ovaries. This suggested that these broadly conserved, ubiquitous complexes are mostly dedicated to regulation of *dhd* in this tissue. This paradigmatic case presented the opportunity to dissect the mechanisms at play in the epigenomic control of the establishment of specific transcriptional programs, as is the one set during oogenesis.

Next, using Cut&Run chromatin profiling with a dedicated data analysis strategy, I found that *dhd* is embedded in a heterochromatic H3K27me3/H3K9me3-enriched mini-domain flanked by DNA regulatory elements, including a *dhd* promoter-proximal element essential for its expression. Surprisingly, Lid, Sin3A, Snr1 and Mod(mdg4) impact H3K27me3 and this regulatory element in distinct manners. However, I showed that these effectors activate *dhd*

independently of H3K27me3/H3K9me3 and that these marks are not required to repress *dhd* in adult tissues.

Altogether, my work uncovered multiple unusual genomic and epigenomic characteristics at the *dhd* locus, but did not identify a single feature that was truly defining ovarian hyperactivation. The dramatic regulation of *dhd* may rely not on an individual trait but rather on a unique combination of such rare features.

Through the example of *dhd*, my work demonstrates the puzzling process that is gene activation in the right place, at the right time and in the right amount. It also illustrates the difficulty to establish general rules on chromatin-based regulatory systems, which are very often context-dependent.

Keywords: epigenetics, epigenomics, transcriptional regulation, oogenesis, *Drosophila melanogaster*, Deadhead, oocyte-to-zygote transition

Table of Contents

INTRODUCTION

I. PART I: MECHANISMS OF EUKARYOTIC TRANSCRIPTION	16
1. THE PROCESS OF EUKARYOTIC TRANSCRIPTION	16
2. PROMOTER-ENHANCER INTERACTIONS REGULATE TRANSCRIPTION	19
3. REGULATION OF GENE EXPRESSION BY UNTRANSLATED REGIONS.....	24
II. PART II: SHAPING OF THE CHROMATIN REGULATES GENE TRANSCRIPTION	26
1. THE NUCLEOSOME: THE BASIC UNIT OF CHROMATIN	26
2. DIVERSITY OF EPIGENOMIC EFFECTORS	28
3. NUCLEOSOME DYNAMICS IMPACT TRANSCRIPTION	32
4. THE HETEROCHROMATIC H3K9ME3/HP1 PATHWAY.....	37
5. POLYCOMB EFFECTORS MEDIATE GENE SILENCING...AND MORE?	42
6. COUNTERACTING REPRESSION TO PROMOTE ACTIVATION.....	47
7. HISTONE PTMS IN ACTIVE TRANSCRIPTION	49
A. <i>Histone PTMs during transcription elongation</i>	50
B. <i>Is H3K4me3 instructive for transcription activation?</i>	51
8. BIVALENT PROMOTERS	53
9. GENOME FOLDING AND TRANSCRIPTION.....	55
III. PART III EPIGENOMIC REGULATION DURING D. MELANOGASTER OOGENESIS	60
1. MAIN STEPS OF DROSOPHILA OOGENESIS	61
2. TRANSCRIPTIONAL PROGRAMS OF NURSE AND FOLLICLE CELLS ARE ESSENTIAL FOR OOCYTE NUTRITION AND MATURATION .	64
3. MECHANISMS OF EPIGENOMIC REGULATION IN THE FEMALE GERMLINE	65
A. <i>Silencing of bulk chromatin by the piRNA system</i>	65
B. <i>The importance of fine transcriptional regulation during oogenesis</i>	67
IV. PART IV THE PECULIAR CASE OF DHD REGULATION	70
1. THE ESSENTIAL ROLE OF DEADHEAD AT FERTILIZATION.....	71
2. DHD, A THIOREDOXIN NOT LIKE THE OTHERS.....	73
3. THE INTRICATE DHD LOCUS	74
4. STATE-OF-THE ART OF DHD REGULATORS IN TRANSCRIPTION REGULATION.....	77
A. <i>The histone demethylase Lid/dKDM5</i>	77
B. <i>The Sin3-HDAC complex</i>	79
C. <i>The nucleosome remodeler subunit Snr1</i>	80
D. <i>The multifaceted protein Mod(mdg4)</i>	81

RESULTS &DISCUSSION

I. AN SHRNA SCREEN IDENTIFIES MATERNAL CHROMATIN FACTORS REQUIRED FOR THE OOCYTE-TO-ZYGOTE TRANSITION	ERREUR ! SIGNET NON DEFINI.
1. ARTICLE PRESENTATION	ERREUR ! SIGNET NON DEFINI.

II. THE CONCERTED ACTIVITY OF EPIGENOMIC EFFECTORS IS ESSENTIAL TO ESTABLISH TRANSCRIPTIONAL PROGRAMS DURING OOGENESIS.....	ERREUR ! SIGNET NON DEFINI.
1. ARTICLE PRESENTATION	ERREUR ! SIGNET NON DEFINI.
III. ADDITIONAL RESULTS AND DISCUSSION	ERREUR ! SIGNET NON DEFINI.
1. IS THERE AN "OVARIAN HYPERACTIVATION CODE"?	ERREUR ! SIGNET NON DEFINI.
2. AN UNUSUAL HETEROCHROMATIC DOMAIN.....	ERREUR ! SIGNET NON DEFINI.
3. DIFFERENT ROLES FOR EPIGENOMIC REGULATORS.....	ERREUR ! SIGNET NON DEFINI.
4. AN APPROACH TO PROFILE REGULATORY ARCHITECTURE OF CHROMATIN DOMAINS.	ERREUR ! SIGNET NON DEFINI.
GENERAL CONCLUSION	211
REFERENCES.....	212

Abbreviations

CDS: CoDing sequence

ESCs: Embryonic Stem Cells

GSC: Germ Stem Cell

GTF: General Transcription Factors

HAT: Histone Acetyltransferase

HDAC: Histone DeAcetylase

HDMT: Histone DeMethylase

HMT: Histone Methyltransferase

HP1: Heterochromatin Protein 1

IBP: Insulator Binding Protein

NDR: Nucleosome Depleted Region

PcG: Polycomb Group

piRNA: piwi-RNA

PRC: Polycomb Repressive Complex

PRE: Polycomb Response Element

PTM: PostTranslational Modification

RNA Pol: RNA Polymerase

shRNA: short hairpin RNA

siRNA: small interference RNA

SNBPs: Sperm Nuclear Basic Proteins

TAD: Topologically Associated Domain

TE: Transposable Element

TF: Transcription Factor

TrxG: Trithorax Group

TSS: Transcription Start Site

INTRODUCTION

I. PART I: MECHANISMS OF EUKARYOTIC TRANSCRIPTION	16
1. THE PROCESS OF EUKARYOTIC TRANSCRIPTION	16
2. PROMOTER-ENHANCER INTERACTIONS REGULATE TRANSCRIPTION	19
3. REGULATION OF GENE EXPRESSION BY UNTRANSLATED REGIONS.....	24
II. PART II: SHAPING OF THE CHROMATIN REGULATES GENE TRANSCRIPTION	26
1. THE NUCLEOSOME: THE BASIC UNIT OF CHROMATIN	26
2. DIVERSITY OF EPIGENOMIC EFFECTORS	28
3. NUCLEOSOME DYNAMICS IMPACT TRANSCRIPTION	32
4. THE HETEROCHROMATIC H3K9ME3/HP1 PATHWAY.....	37
5. POLYCOMB EFFECTORS MEDIATE GENE SILENCING...AND MORE?	42
6. COUNTERACTING REPRESSION TO PROMOTE ACTIVATION	47
7. HISTONE PTMS IN ACTIVE TRANSCRIPTION	49
A. <i>Histone PTMs during transcription elongation</i>	50
B. <i>Is H3K4me3 instructive for transcription activation?</i>	51
8. BIVALENT PROMOTERS	53
9. GENOME FOLDING AND TRANSCRIPTION.....	55
III. PART III EPIGENOMIC REGULATION DURING <i>D. MELANOGASTER</i> OOGENESIS	60
1. MAIN STEPS OF DROSOPHILA OOGENESIS	61
2. TRANSCRIPTIONAL PROGRAMS OF NURSE AND FOLLICLE CELLS ARE ESSENTIAL FOR OOCYTE NUTRITION AND MATURATION .	64
3. MECHANISMS OF EPIGENOMIC REGULATION IN THE FEMALE GERMLINE	65
A. <i>Silencing of bulk chromatin by the piRNA system</i>	65
B. <i>The importance of fine transcriptional regulation during oogenesis</i>	67
IV. PART IV THE PECULIAR CASE OF <i>DHD</i> REGULATION	70
1. THE ESSENTIAL ROLE OF DEADHEAD AT FERTILIZATION.....	71
2. DHD, A THIOREDOXIN NOT LIKE THE OTHERS.....	73
3. THE INTRICATE <i>DHD</i> LOCUS	74
4. STATE-OF-THE ART OF DHD REGULATORS IN TRANSCRIPTION REGULATION.....	77
A. <i>The histone demethylase Lid/dKDM5</i>	77
B. <i>The Sin3-HDAC complex</i>	79
C. <i>The nucleosome remodeler subunit Snr1</i>	80
D. <i>The multifaceted protein Mod(mdg4)</i>	81

Germ cells retain the potential to reproduce an entire organism upon fertilization. Gametogenesis represents thus an extreme cellular differentiation process where complex transcriptional programs must be tightly regulated. Chromatin-based regulation of gene expression and thus epigenomic effectors are crucial in the establishment of specific transcriptional programs in the female germline. Yet, the molecular mechanisms underpinning this regulation remain largely unknown. A major challenge in the field is to establish direct functional connections between upstream chromatin factors, which generally have a broad impact on chromatin, and the controlled expression of highly specialized terminal effectors. My PhD project was precisely based on the discovery of a specific functional connection between a series of epigenomic effectors and the highly regulated terminal effector of zygote formation in *Drosophila melanogaster*: the maternal thioredoxin Deadhead (Dhd). In this introduction I will first describe the mechanisms of eukaryotic transcription followed by the presentation of chromatin-based regulatory mechanisms of gene expression. Next, I will focus on the importance of transcriptional regulation during oogenesis. I will finish by the presentation of the maternal thioredoxin Deadhead and its epigenomic regulators.

I. PART I: Mechanisms of eukaryotic transcription

The DNA molecule carries the genetic information necessary for an organism's development and functioning. Yet, in multicellular organisms, each cell has specific features that define its identity and function. Therefore, information stored in the DNA molecule needs to be used at the right place at the right time. Indeed, gene expression is temporarily, spatially and quantitatively regulated. Various factors affect gene expression and regulation can be exerted at transcriptional, post transcriptional and post translational levels. My case study, the *dhd* gene, is a hyperactivated gene, it is thus important to put it in perspective the mechanisms of transcription, which is the focus of these paragraphs.

1. The process of eukaryotic transcription

Globally, the process of transcription can be divided into three main steps: initiation, elongation and termination. During initiation, the RNA polymerase (RNA Pol) complex is formed on the gene's promoter with the help of general transcription factors (GTFs), resulting in the assembly of the Pre-initiation complex (PIC), which begins transcription and can go into pause. Next, throughout the elongation step, the RNA Pol moves along the template strand, synthesizing the RNA molecule. The polymerase continues transcribing until it reaches the terminator sequence. Finally, the nascent mRNA and the RNA Pol II are released constituting the termination step (Fig1). Transcription is thus a multistep process and all these steps can undergo regulatory control that depends on the gene, the cell type, as well as internal and external signaling.

In eukaryotes not all genes are transcribed by the same polymerase. Three RNA polymerases (RNA Pol I-III) have been identified, differing in their subunit composition and targeting different classes of genes (Sentenac, 1985) RNA Pol I transcribes ribosomal RNAs, RNA Pol II transcribes protein coding genes and some non-coding RNAs and finally, RNA Pol III transcribes small ribosomal 5S RNA, tRNAs and other small non-coding RNAs. On the contrary, prokaryotes use the same RNA Pol to transcribe all their genes. This reflects the increased complexity that is regulation of gene expression in eukaryote systems. Why develop different enzymes? Differences in the structure and composition of the three RNA Pol are linked to the genes they transcribe. For example, to abundantly synthesize tRNAs, RNA Pol III must efficiently terminate and re-initiate transcription on the short genes it targets. Structural analysis of this enzyme revealed that specific subunits permit a loose binding of the RNA:DNA hybrid thereby allowing a fast termination of transcription (Arimbasseri and Maraia, 2016; Hoffmann et al., 2015). Also, two RNA Pol I-specific subunits are needed for the high polymerase loading rate and enhancement of the rRNA gene transcription cycle (Albert et al., 2011). At last, RNA Pol II targets a large set of differently regulated genes, including protein-encoding genes, which is why it has been of particular interest for the study of the regulation of gene expression and will be the main focus of the next paragraph.

RNA Pol II is composed of 12 subunits Rpb1-12 (RNA polymerase B) and is conserved from yeast to human. The largest subunit Rpb1 has a C-terminal domain (CTD) containing a repeat sequence of 7 amino acids (Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7) necessary for the proper functioning of RNA Pol II (Hsin and Manley, 2012). Specific residues in the CTD repeats can be phosphorylated and these post-translational modifications (PTM) play an important

role in the regulation of RNA Pol II activity during the transcription cycle (Fig1) (Buratowski, 2009; Hsin and Manley, 2012). Deletion of the RNA Pol II's CTD, negatively affects mRNA processing, including capping, cleavage/polyadenylation and splicing (S. McCracken et al., 1997; Susan McCracken et al., 1997), indicating thus a crucial role in mRNA processing. First, phosphorylation of serine 5 (Ser5P) is found concomitant with transcription initiation and it has been shown that this modification recruits and stimulates capping machinery at the nascent pre-mRNA. This mark is also important for initial RNA Pol II progression through DNA by the unbinding of the specific proteins that recruited it. During elongation Ser5P undergoes progressive dephosphorylation coincident with the release of 5' processing factors. Next, phosphorylation of Ser2 marks the elongation phase of transcription and is increasingly enriched towards the ends of genes. It has been observed that Ser2P is required for the recruitment of polyadenylation factors to the 3' end of genes in vivo (Ahn et al., 2004). Additionally, genome-wide ChIP experiments showed that peaks of 3' processing factors followed Ser2P peaks, consistent with the view that CTD Ser2P contributes to the recruitment of the polyadenylation complex (Mayer et al., 2012, 2010). Also, Ser7 phosphorylation is found early in transcription initiation and retained until transcription termination in all RNA Pol II-dependent genes. However, this mark has been found to be particularly required for expression of a sub-class of genes encoding small nuclear (sn)RNAs, where it facilitates recruitment of the gene-specific Integrator complex (Egloff et al., 2012, 2007). Finally, hypophosphorylation of the CTD is a prerequisite for RNA pol II to enter the preinitiation complex therefore at the end of the transcription cycle, the regeneration of the hypophosphorylated state of the CTD is critical for RNA Pol II recycling. In conclusion, the CTD "code" allows the coupling of transcription with co-transcriptional RNA processing through the timely recruitment of the appropriate factors at the right point of the transcription cycle. Transcription dynamics, efficiency and rate can thus be regulated, in part, by modifying the enzyme that synthesizes the RNA molecule.

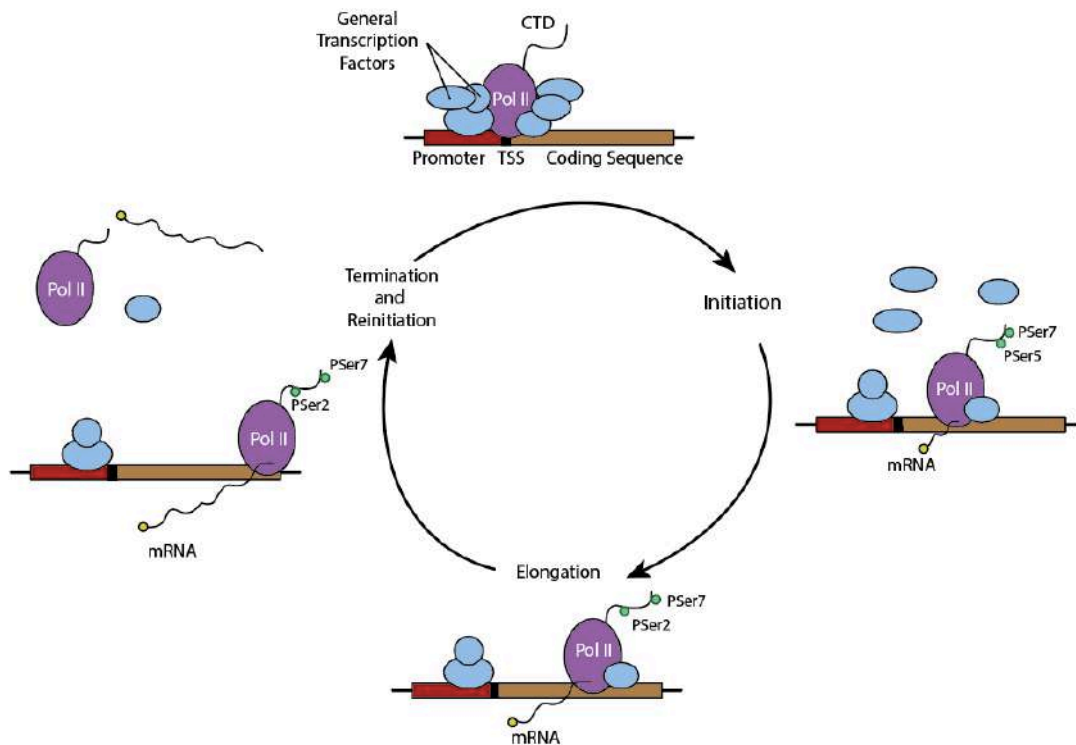


Fig1-Schematic representation of the main steps of the RNA Pol II transcription cycle.

The main steps of the transcription cycle: initiation elongation and termination. Phosphorylations of Serines 2, 5 and 7 on the RNA Pol II C-terminal domain (CTD) are shown. RNA Pol II binds the gene promoter assisted by general transcription factors (GTFs). The polymerase is released and begins active RNA synthesis throughout the elongation phase. Finally, polymerase is dephosphorylated and released along with the mRNA and can be recruited for a new cycle (inspired from Chang Hongh, 2016).

2. Promoter-Enhancer interactions regulate transcription

Promoters are DNA sequences that define where transcription of a gene by the RNA Pol begins. Indeed, transcription initiates at a defined positioned, the transcription start site (TSS) located at the 5' end of a gene. The TSS is embedded within a core promoter which is the minimal stretch of contiguous DNA sequence sufficient to direct transcription initiation. The core promoter serves as a binding platform for the assembly of the transcription machinery. Indeed, general transcription factors (GTFs) bind DNA sequences at the core promoter elements and subsequently recruit the RNA Pol II to form the preinitiation complex (Matsui et al., 1980; Orphanides et al., 1996). However, core promoters generally have low basal activity that can be activated by enhancers (Banerji et al., 1981; Shlyueva et al., 2014). Enhancers are segments of DNA that enable transcriptional regulation in a spatio-temporal

manner. Central to the communication between enhancer and promoters are transcription factors (TFs) and cofactors (Haberle and Stark, 2018). TFs are DNA binding proteins, recognizing specific sequence motifs on promoters and enhancers which in turn recruit cofactors having a variety of biochemical functions. These are thus central effectors of transcriptional regulation. Therefore, to better understand transcriptional regulation it is important to gain insight on all these regulatory elements and their interactions.

Computational analyses of core promoters in fly (Ohler et al., 2002) and human (FitzGerald et al., 2006) have revealed a series of over-represented sequence motifs in these regions. For example, the TATA-box is a core promoter element recognized by one of the GTFs that mediates RNA Pol II recruitment. It has a well-defined position and might thereby determine the choice of TSS at a fixed downstream position. Although it is a well conserved element from yeast to human it is found only at a minority of core promoters, in flies they correspond to ~5% of core promoters (FitzGerald et al., 2006; Ohler et al., 2002). A more widely used element is the initiator (Inr) motif. Although its consensus sequence differs between flies and humans, this element, as the TATA-box, has a well-defined position, that in this case, overlaps with the TSS (FitzGerald et al., 2006). In promoters that lack a TATA-box, the Inr motif is often accompanied by another motif, the downstream promoter element (DPE), which is positioned downstream of the TSS (Burke and Kadonaga, 1996) and functions cooperatively with the Inr for GTFs binding. Increasing or decreasing of one nucleotide the spacing between the Inr and DPE results in reduced binding of GTFs and decrease in transcriptional activity (Kutach and Kadonaga, 2000), indicating thus the importance of a strict Inr–DPE spacing. In addition to these three most abundant core promoter motifs, other motifs either with defined positions relative to the TSS or not, have been identified showing the variety of core promoters.

Interestingly, TATA-box and DPE rarely co-occur in flies, they were thus suggested to be associated with functionally distinct groups of genes (FitzGerald et al., 2006; Ohler et al., 2002). This raised thus the question of the association of core promoters and gene function. Integrating data of sequence composition and motifs with other properties including transcription initiation pattern (focused or dispersed), chromatin configuration and gene function revealed three main types of core promoter in metazoa (Fig2) (Haberle and Stark, 2018; Vo ngoc et al., 2017). First, promoters found at adult tissue-specific genes and differentiated cell-specific genes, characterized by TATA-box and Inr motifs, sharp initiation

patterns and imprecisely positioned nucleosomes (see section II-3). Second, promoters of broadly expressed housekeeping genes, in flies were enriched for the motifs DRE (DNA replication-related element) and ohler 1/6 and in mammals they overlapped with CpG islands. They were also associated with dispersed transcription initiation and a well-defined nucleosome-depleted region. Third, core promoters of developmental TFs, in mammals they resemble to housekeeping gene core promoters. In flies, they tend to contain a DPE motif and focused initiation of transcription. There is thus a correlation between features of core promoters and gene function. Notably, a bias between gene function and specificity of promoter-enhancer interaction has also been observed. Zabidi and colleagues studied this specificity on a genome-wide scale, by high-throughput genome-wide screening of enhancer activity in housekeeping and developmental core promoters (Arnold et al., 2013; Zabidi et al., 2015). Their study revealed different motif enrichment in both core promoters and enhancers between housekeeping and developmental genes (summarized in Table1). Developmental and housekeeping core promoter enhancer elements also show proximity bias. Enhancer sequences activating developmental core promoters appeared to be gene promoter-distal while enhancers activating housekeeping core promoters were generally found proximal to the gene promoter (Arnold et al., 2013). This suggests that some transcriptional enhancers exhibit preferences for certain core promoter elements. Overall, these data show that specific features of cis regulatory elements are linked to gene function. However, the mechanisms of how these features can influence a specific transcription pattern are yet to be elucidated.

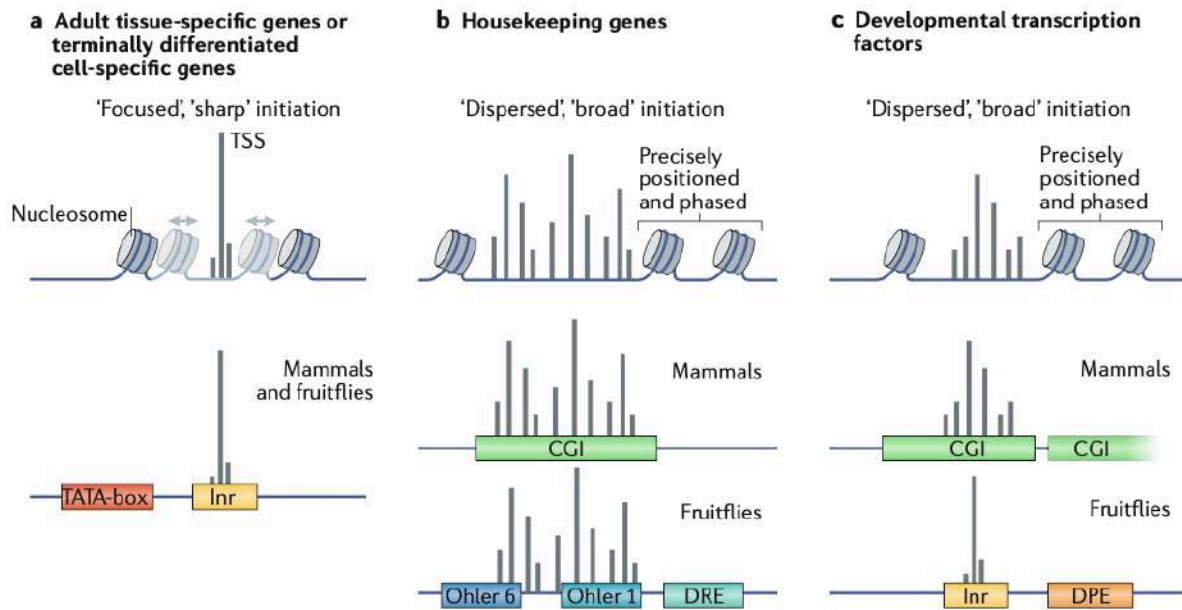


Fig2-Three types of core promoters in metazoa

Main categories of core promoters in *Drosophila* and mammals based on different properties, including initiation pattern (focused or dispersed), sequence composition motifs and gene function (Haberle and Stark, 2018).

	dCP enhancers	hkCP enhancers
Target gene function	Signaling, patterning, morphogenesis	House-keeping
Enhancer location	Distal or intronic	Promoter-proximal
Target gene promoter motifs	TATA, Inr, MTE, DPE	DRE, TCT, others
Enhancer motifs	GAGA (GAGAG)	DRE (TATCGATA)
Transcription factors	Trithorax-like (Trl) / GAF	DRE binding factor (Dref)

Table 1: Summary of differences between developmental and housekeeping core promoters and their enhancers.

(From Lorberbaum and Barolo 2015 based on the result by Zabidi et al., 2015).

Enhancers can be located near the promoter of their cognate gene or be distally positioned. Distances can vary from a few kb to several megabases. Active promoters however, are often in spatial proximity to their enhancer and the establishment of these contacts are tightly linked to 3D organization of DNA in the nucleus (mechanisms to achieve genome folding will be detailed in section II-9).

The regulatory input that core promoters receive from enhancers is mediated by TF binding both of these elements and by transcriptional cofactors, which are recruited by TFs through protein–protein interactions (Fig3). Importantly, cofactors often have enzymatic activities and can post-translationally modify components of the transcription machinery and the surrounding nucleosomes. It has been observed that the Mediator complex, a cofactor, recruited to enhancers and interacting with core promoters can increase RNA Pol II recruitment and PIC assembly (Eychenne et al., 2016), thereby increasing transcription rate. Mediator in yeast can stimulate phosphorylation of Ser5 at the CTD of RNA Pol II (Esnault et al., 2008), which is important for mRNA processing and RNA Pol II progression through the transcription cycle. The interaction between bound TFs and recruited cofactors can thus lead to transcription activation.

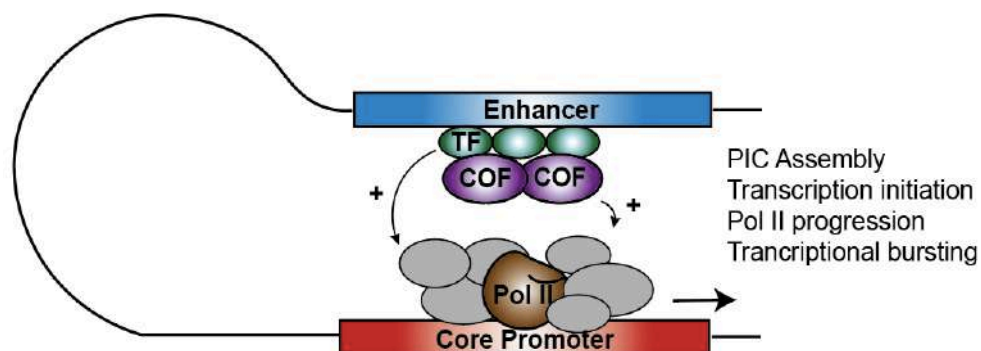


Fig 3-Enhancers communicate with promoters through transcription factors and cofactors. Enhancers and promoters can bind transcriptional factors (TF) which in turn bind cofactors (COF) thereby mediating enhancer-promoter communication. This communication results in stimulation of transcription via different mechanisms such as PIC assembly, transcription initiation, stimulation of RNA Pol II progression or modulation of transcriptional bursts.

Another mechanism through which enhancers can regulate promoter activation is by modulating transcription bursts. Transcription is episodic consisting on short but intense

“bursts”, which comprise initiation events separated by periods of inactivity (Chubb et al., 2006; Raj et al., 2006). Patterns of these bursts can have important roles in development. For example, when visualizing transcription in living fly embryos, it was observed that lengthening of transcriptional activity periods is important for the establishment of boundaries during embryo patterning (Lucas et al., 2013). It is therefore important to understand how these “bursts” are regulated. Another study in living *Drosophila* embryos revealed that the levels of gene activity depend on the frequency of transcriptional bursts and that this frequency can be regulated by enhancers (Fukaya et al., 2016). The authors evaluated the effect of different developmental enhancers on transcriptional bursting of a reporter gene, and observed that enhancers produced transcriptional bursts with similar amplitudes and duration but generated different bursting frequencies, with strong enhancers producing more bursts than weak enhancers. This result suggests thus that regulation of bursting frequencies by enhancers can be a parameter of gene control. Moreover, it was observed that mutations in the TATA-box led to a decrease in burst size, that is the number of transcribing RNA Pol II molecules per burst (Hornung et al., 2012), suggesting that burst size is a promoter-specific property. These data show that several factors can influence transcription bursts, either in frequency or in amplitude and thereby regulate the transcriptional output. Further research will be needed to fully understand the mechanisms regulating transcriptional bursts.

3. Regulation of gene expression by untranslated regions

Additional sequences within the gene can also provide a regulatory point for transcription. Majority of eukaryotic genes contain introns. The process of removing introns is energy and time-consuming, however, these elements have been evolutionary conserved which is indicative of a biological functional role. The most obvious advantage they confer is increasing the repertoire of proteins through alternative splicing, meaning that a single gene can produce multiple isoforms of a protein depending on cell type and environment (Nilsen and Graveley, 2010). Additional research in the field has associated introns to a variety of functions such as protection against environmental stress, or as the source of non-coding RNAs (Dwyer et al., 2021). A third role for introns, well established in the literature, is their function in regulating gene expression. It has been observed in yeast that in a physiological

context, genes require promoter proximal introns for full transcriptional output (Furger et al., 2002). Additionally, a quantitative analysis of constructs containing a human intron in the open reading frame of a reporter gene, revealed that its presence increased not only the accumulation of mature mRNA but also the efficiency of their translation (Nott, 2003). The capacity of introns to increase gene expression has been observed in many eukaryotes including mammals, plants, yeast, and insects (Shaul, 2017), however the underlying mechanisms are not fully elucidated. Studies in budding yeast have started to provide hints for the mechanisms at play. It was observed that the presence of an intron within a gene results in formation of a multi-looped gene architecture. When looping is defective, these interactions are abolished and there is no enhancement of transcription despite normal splicing (Agarwal and Ansari, 2016). Introns are thus important players in the regulation of gene expression.

Furthermore, the 5' and 3' untranslated regions (UTR) located at the extremities of a given gene have also been involved in the regulation of expression. A study investigated the role of alternative 5'UTRs of the same gene on mRNA translation efficiency with an *in vivo* reporter assay (McClelland et al., 2009). They observed that among the 5'UTRs tested, three enhanced translation while two had repressive effects. Modeling of the mRNA secondary structure in the 5'UTR revealed the presence of compact structures around the start codon in the repressive 5'UTRs. Translational efficiency was also found inversely correlated to 5'UTR length. Additionally, genome-wide studies in human revealed differences in structure and nucleotide content of 5'UTRs of housekeeping and developmental genes (Ganapathi et al., 2005). *In silico* comparisons of genes with low and high levels of protein output showed that 5'UTRs that enable efficient translation are short, have low GC content, are relatively unstructured, and do not contain upstream AUG codons (Kochetov et al., 1999). Although mechanisms describing how 5'UTR sequence can influence transcription are not well understood, these data suggest a correlation between this gene element and regulation of transcription. The 3'UTR, situated downstream of the protein coding sequence, has also been involved in the regulation of gene expression. Indeed, the poly(A) tail at the 3' end mRNAs provide a binding platform for poly(A) binding proteins (PABP) that have roles in mRNA export, stability, decay and translation (Gorgoni, 2004; Mangus et al., 2003). For example, PABP mRNAs can bind poly(A) tracts in their own 5'UTRs, resulting in translational repression. Moreover, in human, it was observed that the generation of alternative 3' UTR isoforms is a characteristic of ubiquitously transcribed genes that are involved in diverse gene

regulatory processes and are distinct from the classical housekeeping genes that generate single UTRs (Lianoglou et al., 2013). These ubiquitously transcribed multi-UTR genes use different 3' UTR isoforms to achieve tissue- and context-dependent expression. Changes in the 3'UTR isoform expression represents thus a component of gene expression programs. Altogether, these data shows that untranslated regions can have an important role in transcriptional regulation but mechanisms are yet to be elucidated.

To conclude, transcription is a complex process that can be modulated by regulatory elements encoded within the DNA molecule and by the binding of specific proteins. However, DNA is not naked in the nucleus but wrapped around histone proteins. This structure represents thus a major point of transcriptional regulation.

II. Part II: Shaping of the chromatin regulates gene transcription

1. The Nucleosome: the basic unit of chromatin

The basic unit of organization of chromatin is the nucleosome, formed of 146 bp of DNA that wrap around an octamer of histones. This octamer is composed of two copies of the four “core histones”: H2A, H2B, H3 and H4. Additionally, linker DNA between two nucleosome cores is bound by the linker histone H1 (Fig4). Histones have positively charged surfaces formed by basic amino acid side chains which allow them to interact with the negatively charged DNA backbone forming thus a highly tight bound.

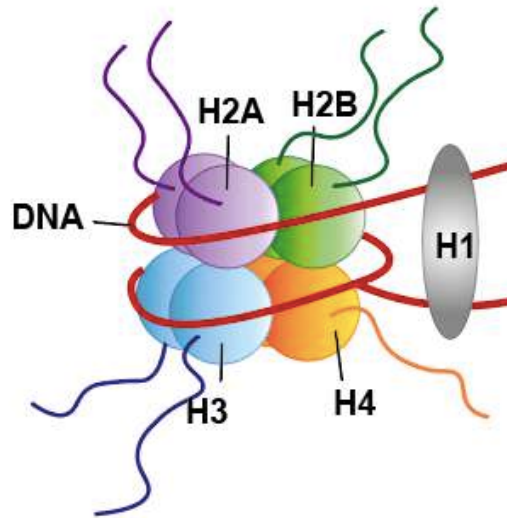


Fig4-Schematic representation of the nucleosome.

The nucleosome core is composed of 146 bp of DNA wrapped around a histone octamer containing two H2A-H2B dimers and two H3-H4 dimers. Linker DNA between two nucleosome cores is bound by the linker histone H1.

The amino acid sequences of the core histones (H2A, H2B, H3 and H4) are extremely well conserved in evolution. Histones are small proteins (10-15 kDa) that share a globular domain with a α -helical arrangement called histone fold, implicated in histone dimerization (Arents et al., 1991; Luger et al., 1997). Importantly, core histones possess flexible N-terminal regions, called histone tails, that associate more loosely with the nucleosome and remain accessible for posttranslational modifications. The most well-studied histone modifications include acetylation, phosphorylation, methylation, and ubiquitylation, although many other modifications have been reported. These modifications play a central role in the regulation of gene expression as well as many other DNA processes such as repair, replication and recombination (Bannister and Kouzarides, 2011). Notably, these histone modifications are reversible by specific enzymes giving the possibility of a dynamic and fine regulation of transcription.

Nucleosome composition can be altered by the insertion of histone variants (Martire and Banaszynski, 2020; Talbert and Henikoff, 2017). These are encoded by a different set of genes and show slight variations in the amino acid sequence. Different variants of histones can be incorporated. Variations of histone H3 and histone H2A are common while histones H2B and H4 appear to be predominantly canonical.

To date, the H2A family contains the greatest sequence diversity of identified variants. The variant H2A.Z, H2Av in *Drosophila*, is present in almost all organisms and has been associated to diverse transcriptional states. Indeed, several studies have shown that H2A.Z nucleosome occupancy at promoters is inversely correlated with transcription (Guillemette et al., 2005; Zhang et al., 2005). However, it has also been observed that H2A.Z is required for activation of heat shock genes during heat shock responses (Zhang et al., 2005) and that it can promote the recruitment of the RNA Pol II (Adam et al., 2001). Additionally, in *Drosophila*, the presence of H2Av, the homolog of H2A.Z (van Daal et al., 1988), correlates with paused RNA Pol II (Mavrich et al., 2008). It was later observed that H2Av occupies the promoter in absence of gene expression but it decreases upon gene induction (Kusch et al., 2014). Together, these results suggest that the role of H2A.Z may be to recruit Pol II and poise genes for activation.

Histone H3 possesses two universal variants. First, CENP-A, who is found at chromosome centromeres and is important for cohesin recruitment at these regions (Santaguida and Musacchio, 2009). Second, H3.3 enriched at gene bodies and regulatory elements such as promoters and enhancers (Ahmad and Henikoff, 2002; Mito et al., 2007), suggesting an association with active transcription. Although, absence of H3.3 in *Drosophila* led to transcriptional defects, the latter could be compensated by increased expression of variant H3.1 (Sakai et al., 2009). This suggests that transcription was affected by lack of histone replacement rather than by the variant of H3. Additionally, H3.3 has also been found at transcriptionally silent regions such as telomeres and centromeres (Goldberg et al., 2010; Szenker et al., 2012). These data suggests thus that the link of H3.3 with transcription is more complex than initially thought. Overall, even though the mechanisms are not fully elucidated, histone variants can influence transcription and have thus functional roles in the genome.

Dynamic histone PTMs and variants contribute to the complexity of epigenetic regulation of the genome. Chromatin dynamics is mediated by epigenomic effectors it is thus crucial to understand how these factors work.

2. Diversity of epigenomic effectors

The chromatin is highly dynamic and can be shaped to respond to the needs of the cells through a wide variety of mechanisms that will be detailed throughout this manuscript. Accordingly, the molecular actors behind chromatin shaping are also very varied. They can essentially be classified in three main categories: histone modifiers, nucleosome remodelers and histone chaperones. During my PhD I also focused on insulator binding proteins, that even though they do not interact with histones, they are also important for the organization of the chromatin landscape. I will thus also describe this category.

Histone modifiers: Histone PTMs play vital roles in regulating both gene activation and repression. These modifications are reversible which helps fine-regulate gene transcription according to specific cues. Over the years a plethora of histone modifiers i.e., enzymes catalyzing changes in histone PTMs have been identified. These epigenetic players have been categorized as writers: that introduce various chemical modifications on histones, readers: the specialized domain containing proteins that identify and interpret those modifications and erasers: the dedicated group of enzymes proficient in removing these chemical tags. Writers and erasers are divided into classes on the basis of the specific PTM they effect (Fig5) (Bannister and Kouzarides, 2011; Biswas and Rao, 2018). The corresponding enzymatic activities are referred to as histone acetyltransferases (HATs), histone methylases (HMTs), histone kinases and histone ubiquitin-transferases. Analogously, erasers comprise histone deacetylases (HDACs) histone demethylases (HDMs, or KDMs for lysine demethylases) phosphatases and deubiquitinating enzymes. In general, HDACs have relatively low substrate specificity by themselves, a single enzyme being capable of deacetylating multiple sites within histones. HKMTs and HKDMs on the contrary possess a high level of substrate specificity with respect to their target lysine.

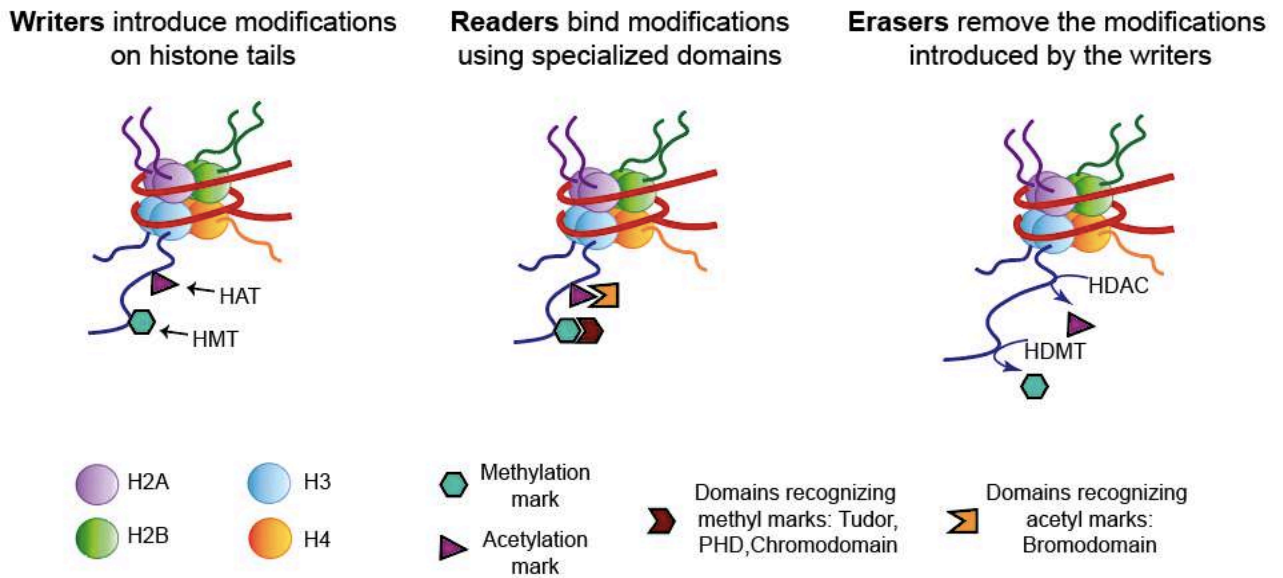


Fig 5-Actions of epigenetic writers, readers and erasers.

Enzymes and proteins capable of adding, binding or removing posttranslational modifications on histone tails are known as writers, readers and erasers respectively. Examples for methylation and acetylation are represented. HAT: histone acetyltransferase, HMT: histone methyltransferase, HDAC: histone deacetylase, HDMT: histone demethylase (modified from Biswas and Rao, 2018).

Nucleosome remodelers are a category of enzymes capable of altering DNA-histone interactions at target nucleosomes, thereby locally and differentially regulating access to DNA. Remodelers can (i) mediate nucleosome sliding i.e., the translational movement of a nucleosome in either direction to expose a region that was previously occluded, (ii) exchange a core histone for a variant histone and (iii) induce eviction of a nucleosome to expose the associated DNA (Fig6). A conserved feature among remodelers is the use of the energy from ATP hydrolysis to achieve these reconfigurations (Becker and Workman, 2013; Clapier and Cairns, 2009).

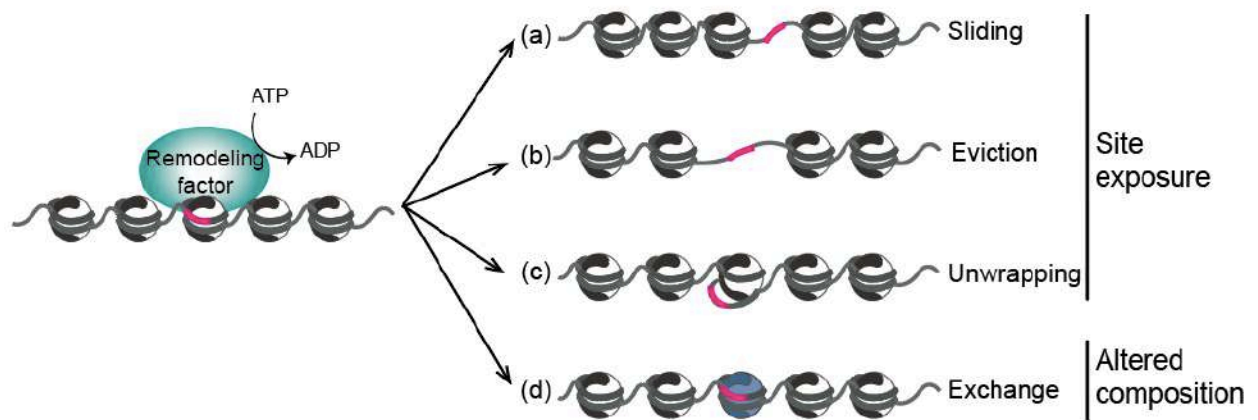


Fig6-Outcomes of nucleosome remodeling

Models for nucleosome remodeling. Remodelers (green) alter nucleosome-DNA interactions in an ATP-dependent manner. They mediate: (a) nucleosome sliding, the translational movement of a nucleosome, (b) nucleosome eviction, (c) localized unwrapping or (d) nucleosome exchange. Processes (a) to (c) lead to DNA site exposure while (d) results in altered composition of the nucleosome (modified from Clapier and Cairns, 2009).

Histone chaperones are proteins that handle non-nucleosomal histones *in vivo*. They accompany and safeguard histones throughout their cellular life. Histone chaperones escort histones, preventing them from aggregating or from spurious interactions with DNA (Hammond, 2017). Importantly, histone chaperones are necessary for deposition of canonical histones as well as of histone variants into particular places in the genome thereby participating in the regulation of chromatin processes. Interestingly, the conserved histone chaperone CAF-1 complex involved in the assembly of H3-H4 histone dimers on newly synthesized DNA, has also been found to participate in the maintenance of heterochromatin through interaction with heterochromatin effectors in fly (Roelens et al., 2017). Overall, this shows the importance of chromatin dynamics regulation by histone chaperones.

Insulators are conserved DNA elements that help organize eukaryotic genomes into physically and functionally regions through diverse mechanisms. They are thought to act by recruiting specific Insulator Binding Proteins (IBPs). The first IBP to be discovered was the CCCTC-binding factor (CTCF). To date, CTCF remains the major protein implicated in insulation in vertebrates (Lobanenkov et al., 1990). On the contrary, in flies, various insulator sequences have been identified and classified according to the proteins they bind. Experiments in *Drosophila* have identified a dozen or more IBPs binding specific DNA

sequences (Melnikova et al., 2020; Özdemir and Gambetta, 2019). They were identified based on their ability to bind to characterized insulators and mediate their function or in genetic screens as being required for the function of a specific insulator. Interestingly, most of these proteins are also classified as transcription factors, such as the conserved dCTCF/CTCF or the GAGA factor (GAF). While most IBPs exhibit direct DNA binding, additional proteins were identified as necessary for insulator activity. This is particularly the case for the cofactors CP190 and Mod(mdg4) which are required along with the DNA-binding protein Su(Hw) at the gypsy transposon for proper insulator function (Georgiev and Gerasimova, 1989; Gerasimova et al., 1995; Pai et al., 2004). Subsequently, studies have shown that CP190 is common to almost all insulators (Ahanger et al., 2013). At least one of the roles of these co-factors is to mediate homotypic and heterotypic protein-protein interactions bridging thus contacts between distant genomic regions. Initially, insulators were defined as having either an enhancer blocker role or a barrier role between euchromatin and heterochromatin. However, extensive research has now associated insulators to broader functions in nuclear biology, such as nuclear organization.

Chromatin dynamics enables the cell to tightly regulate fundamental activities of the genome. It is thus the concerted activity of epigenomic effectors (modifiers, remodelers, chaperones) along with nuclear organizing proteins (insulators), that achieve fine gene regulation. Malfunctioning of these machineries is tightly linked to diseases such as cancer or intellectual disability. It is thus crucial to understand the molecular mechanisms mediated by epigenomic effectors to modulate transcription.

3. Nucleosome dynamics impact transcription

An unavoidable side effect of the structural organization of chromatin is the occlusion of DNA sequences and thus of its regulatory elements and binding sites. Nucleosomes represent a barrier for RNA Pol II progression, therefore, the coordination of histones sliding, leaving, recycling, depositing and positioning together with RNA Pol II passage is crucial in defining transcriptional activity (Fig7).

Nucleosome positioning and composition are regulated processes that can reflect DNA-related activities such as transcription. Genome-wide analyses of nucleosome landscapes have revealed a general feature of eukaryotic promoters. Two well-positioned nucleosomes, designated “+1” and “-1”, separated by a nucleosome-depleted region (NDR) of variable length, demarcate the site of transcription initiation (Jiang and Pugh, 2009). In yeast, a nucleosome is part of the TSS [Albert et al., 2007] while in higher metazoans the TSS does not overlap with a nucleosome, but there is one just downstream (Mavrich et al., 2008).

What determines nucleosome positioning? Some DNA sequences like poly(dA:dT) stretches are unfavorable to DNA wrapping around histones *in vitro* and AT-rich sequences are good predictors of nucleosome-depleted regions *in vivo* (Segal and Widom, 2009). Additionally, DNA-binding activator proteins can recruit nucleosome remodelers and generate nucleosome-depleted regions (Schwabish and Struhl, 2007). Importantly, elongation by RNA Pol II needs cycles of disassembly/reassembly of nucleosomes in coding regions. This complex process implies nucleosome disruption, mobilization and reassembly after the RNA Pol II passage. As a result, a decrease on nucleosomal density can be observed at highly transcribed genes (Lee et al., 2004; Schwabish and Struhl, 2004). Additionally, high rates of histone turnover have been observed at active genes while repressed regions showed low histone turnover in ESCs and *Drosophila* cells (Deal et al., 2010; Deaton et al., 2016). A recent study has shown that transcription is the major cause of old histones eviction, with a more pronounced effect on the variant H3.3 than on the canonical H3.1 (Torné et al., 2020). The authors also described a mechanism ensuring histone recycling and new deposition dependent on the histone chaperone HIRA during transcription. Overall, these data show the complex processes around nucleosome dynamics which are crucial for transcriptional activity.

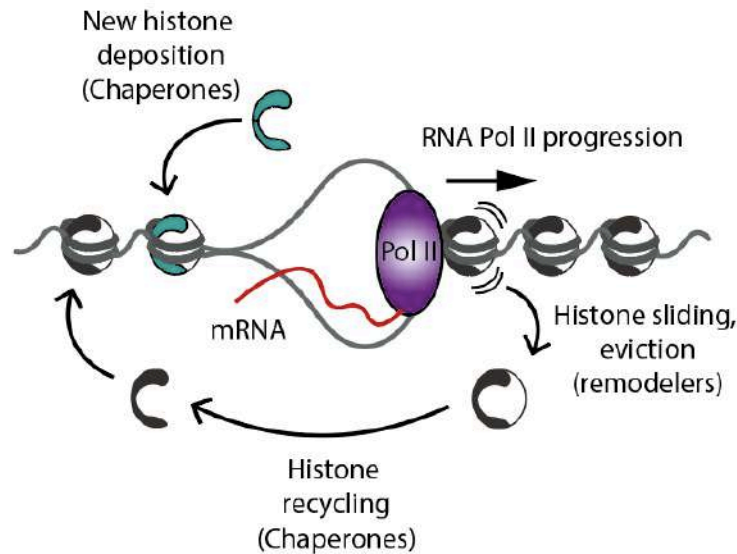


Fig7-Nucleosome dynamics during transcription.

Examples of nucleosomes dynamics during transcription. To allow RNA Pol II progression and reassemble chromatin after, histones are evicted and can then be recycled or a *de novo* histone deposition can take place. These processes are mediated by epigenomic effectors (inspired from Torné et al., 2020).

Despite showing a stereotypical nucleosomal landscape around TSS, not all genes have the same nucleosome stability. At these regions the occupancy of nucleosomes containing the histone variants H3.3/H2A.Z was observed in *Drosophila* and yeast (Henikoff et al., 2008; Xi et al., 2011) and it was shown that these are highly unstable nucleosomes. Although why these nucleosomes are unstable remains unclear, this could provide a mean to achieve NDR regions and promote transcription. Additionally, genome-wide characterization of *in vivo* promoter nucleosome landscapes in yeast revealed two types of promoters. One type is characterized by the presence of dynamic, unstable nucleosomes and is found at highly expressed genes. The other type, contains well-known stable nucleosomes and is found at less frequently expressed genes. Current hypothesis is that the presence of dynamic nucleosomes at highly expressed genes helps to rapidly unwind DNA and as often as necessary, resulting in an increased access of transcriptional machinery to the promoter (Kubik et al., 2015). Accordingly, TSSs of many constitutively expressed housekeeping promoters are usually depleted of nucleosomes and so depend less on nucleosome remodeling (Cairns, 2009; Ganapathi et al., 2005; Rach et al., 2011). In contrast, tightly regulated genes, depend more on remodeling factors to clear their promoters. These data highlight the importance of nucleosome positioning in transcription activity.

To achieve nucleosome depleted regions necessary for transcription, a specific interplay between different categories of epigenomic effectors can take place. Nucleosome remodelers play a crucial role in generating NDRs but they are not the only actors in this process. Indeed, acetylation of histone tails can eliminate positive charges on this residue, decreasing thus their interaction with the negative charged DNA and result in chromatin decompaction (Bannister and Kouzarides, 2011). Many chromatin remodeling enzymes contain protein motifs that recognize modified histones, opening thus the possibility of an interaction between these epigenomic effectors. Interestingly, two *in vitro* studies using budding yeast enzymes showed such interaction. First, Carey and colleagues observed that histone acetylation is sufficient to recruit the remodeler Rsc to nucleosomes (Carey et al., 2006). In turn, Rsc facilitated the passage of RNA Pol II in an ATP-hydrolysis dependent manner, resulting in stimulation of transcription elongation. Second, a study using several yeast nucleosome remodelers showed that specific patterns of histone acetylation resulted in increased rate of nucleosome repositioning or eviction (Ferreira et al., 2007). Notably, a single histone mark combined with different remodelers had different outcomes. This was observed for histone H4 tetra-acetylation, which increased nucleosome transfer by the Rsc complex but reduced the activity of two other remodelers, Chd1 and Isw2. These data show how histone PTMs can recruit remodelers and/or alter their activity impacting thus nucleosome dynamics. In budding yeast, the cooperation between the remodeler complex Swi/Snf and the histone deacetylase complex Saga was found necessary for induction of the stress response transcriptional program (Sanz et al., 2016). Upon cell wall stress, the TF Rlm1 is activated by phosphorylation and interacts with the Swi/Snf complex. It was then observed that Saga subunits are recruited to the promoter of cell wall stress-responsive genes, in a Rlm1 and Swi/Snf-dependent manner, where it acetylates histone H3 at promoters. Interestingly, both Swi/Snf and Saga complexes were necessary for H3 eviction, Rlm1 recruitment and subsequent gene expression. This data suggests thus a cooperation between epigenomic effectors in order to create a favorable chromatin environment for transcriptional activity. Nonetheless, further *in vivo* studies are required to investigate if this is a more general mechanism for the establishment of basal transcriptional programs, to identify the actors at play and assess the impact on transcription

A relationship between nucleosome density and transcription level was also observed in mouse ESCs (Fig8). The comparison of nucleosome density profiles for genes with

varying levels of transcriptional activity showed that patterns emerge from a featureless profile as transcription level elevates. Interestingly, the qualitative trends observed in mouse ESCs are conserved across multi-cellular organisms (Jiang and Zhang, 2021). A correlation between nucleosome positioning and transcription has been observed for a while but how positioning is determined has not been entirely elucidated. Recently, a “tug-of-war” model was proposed based on stochastic simulations. On one hand, enzymes that regulate and reduce inter-nucleosome spacing tend to drive the nucleosome array away from the transcription start site (TSS). On the other hand, positioning enzymes help to align nucleosomes towards the TSS. Competition between these enzymes results in two types of density profiles with well- and ill- positioned +1 nucleosome that qualitatively reproduce in vivo results from both yeast and mouse ESC (Jiang and Zhang, 2021). This results show how the coordinated activity of enzymes could potentially achieve a specific nucleosome positioning.

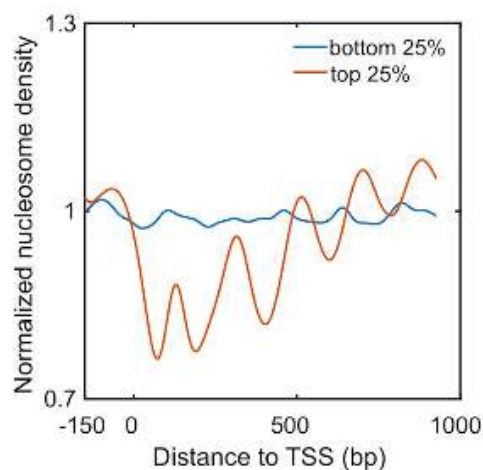


Fig8-Nucleosome density is correlated to transcription level.

Normalized nucleosome density profile in mouse embryonic stem cells near the transcription start site (TSS). Genes were separated depending on levels of transcription activity. Bottom 25% corresponds to the quartile of most inactive genes and top 25% corresponds to the quartile of most active genes (from Jiang & Zhang, 2021).

Overall, these data show the tight link between nucleosome dynamics and the process of transcription. Complex mechanisms and numerous actors determine nucleosome positioning, composition and movements. Further research will be thus necessary to continue to elucidate these mechanisms and their impact on transcription.

4. The heterochromatic H3K9me3/HP1 pathway

Heterochromatin is an architectural feature of eukaryotic chromosomes. It corresponds to tightly packed chromatin and refers to molecular subtypes of transcriptionally repressed domains. One type of heterochromatin has been named “constitutive heterochromatin” since it’s found at structural chromosomal elements such as telomeres and pericentromeres, as well as transposable elements (TEs) and virus-derived sequences (Allshire and Madhani, 2018). This chromatin is distinguished by the formation of heterochromatin blocks characterized by histone hypoacetylation, methylation of histone H3 at lysine 9 (H3K9me) and the presence of the protein HP1 (Heterochromatin Protein 1).

Loss of H3K9me3 following depletion of silencing guides or effectors is associated with transcriptional upregulation of normally silenced regions decorated with this mark (Karimi et al., 2011; Matsui et al., 2010; Smolko et al., 2018). Additionally, artificial recruitment of H3K9 histone methyltransferases or other silencing factors such as HP1 to euchromatin regions result in transcriptional silencing of reporter genes (Ayyanathan, 2003; Ivanov et al., 2007). These data suggest thus, that repressive mechanisms use H3K9me3 and HP1 to achieve this transcriptional state.

Chromatin-based silencing can be achieved by a denser chromatin structuring process that results in efficient exclusion of RNA polymerases or other nuclear enzymes. To create this structure, di- and tri-methylation of H3K9 act as molecular anchors to recruit proteins that either directly modify chromatin or recruit others that do so (see below). Additionally, HP1 proteins can also serve as platforms to establish a repressive chromatin structure. For example, the two HP1 proteins in *S. pombe*, Swi6 and Chp2 are associated with histone deacetylases (HDACs) (Fischer et al., 2009; Yamada et al., 2005). Swi6, Chp2 and their associated HDACs were found to limit RNA Pol II occupancy at centromeric repeats. Furthermore studies *in vitro* and *in vivo* showed that the Swi6 protein can promote repression by capturing transcripts and direct them to the RNA degradation machinery (Keller et al., 2012). Finally, recent work has indicated roles of human and *Drosophila* HP1 in regulating higher-order chromatin structure through the formation of liquid-liquid phase-separated compartments (Larson et al., 2017; Strom et al., 2017). These studies suggest that heterochromatin-mediated gene silencing occurs in part, through sequestration of compacted

chromatin in specific nuclear compartments. These compartments can isolate repressed chromatin from transcription factors and machinery while maintaining a high concentration of factors required for heterochromatin formation, such as HP1 and H3K9 methyltransferases.

How is H3K9me3 chromatin targeted to specific genomic territories? Two main mechanisms have been described relying: (i) on an RNA-based recognition system to appropriately localize histone modification enzymes and HP1 proteins and initiate the heterochromatin formation cascade or (ii) on protein recognition of specific DNA sequences (Fig9).

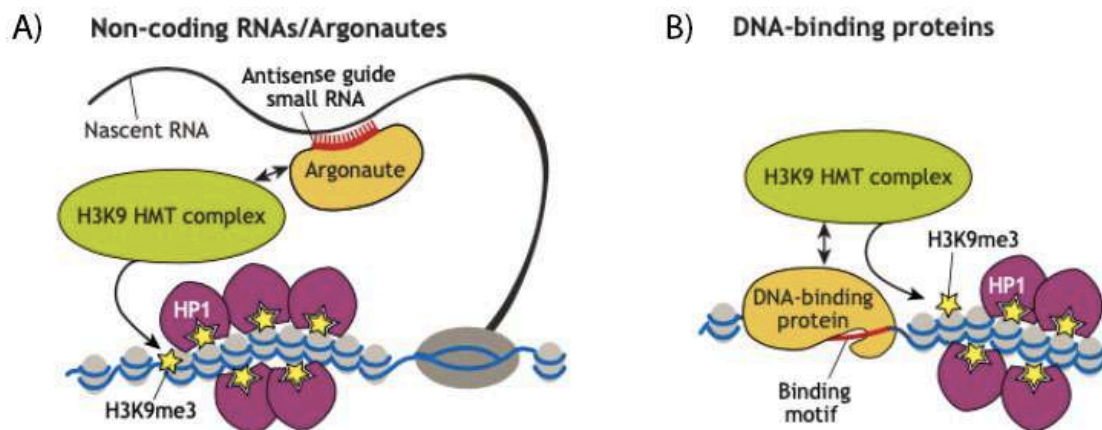


Fig 9-Recruitment of the H3K9me3 machinery

H3K9 methyltransferases can be recruited to specific genomic targets by (A) small non coding RNAs associated with Argonaute proteins or (B) by sequence specific DNA-binding proteins. (Adapted from Ninova et al., 2019)

Studies mainly in *S. pombe* helped elucidate the mechanisms of heterochromatin nucleation. It was observed that the deletion of *argonaute* (*ago*), *dicer* (*dcr*) and *RNA-dependent RNA polymerase* gene homologs resulted in loss of H3K9 methylation and derepression of transgenes at centromeres (Volpe, 2002). The depleted genes belong to the RNA interference (RNAi) machinery hinting thus on a mechanism involving this pathway. Additionally, induced double-stranded RNA (dsRNA) in fission yeast was sufficient to generate synthetic small interfering RNAs (siRNAs) and direct H3K9me-heterochromatin formation at the locus producing the dsRNAs [Simmer et al., 2010]. Verdel and colleagues purified indeed a nuclear Ago complex, the RITS (RNA-induced initiation of transcriptional gene silencing) complex,

which uses Dcr-generated siRNAs to localize to heterochromatic domains (Verdel, 2004). Indeed, when cells lack Dcr, the loss of siRNAs is accompanied by delocalization of RITS from centromeric regions. Further research revealed that the RITS complex can associate with the Clr4 complex, which contains the Clr4 enzyme, the sole H3K9 methyltransferase in fission yeast (Zhang et al., 2008). This interaction promotes thus nucleation of heterochromatin at determined regions. The RITS complex associates also with the RNA-dependent RNA polymerase complex (RDRC). RDRC uses primary transcripts as templates for synthesis of dsRNA, which is subsequently processed into siRNAs, thereby increasing siRNA production and amplifying the system of heterochromatin formation (Allshire and Madhani, 2018). The established model proposes thus that RNA Pol II transcribes RNAs from heterochromatin repeats which are then processed into siRNAs. These siRNAs are next bound by the Ago1 protein and are used to target homologous nascent repeat transcripts, resulting in the recruitment of silencing factors such as H3K9 methyltransferases. In metazoans, RNAi-based targeting and silencing mechanism involve another class of small non coding RNAs, piwi-interacting RNAs (piRNAs). This system of silencing is used for transposable elements in most Metazoa, particularly in the female germline and will be detailed in section III-3-A.

RNA independent mechanisms to target heterochromatin nucleation involve DNA binding proteins. In mammalian systems, the KRAB-containing zinc finger proteins (KRAB-ZFPs) have been found to induce H3K9me3-dependent silencing at endogenous retrovirus targets (Wolf et al., 2015) and TEs (Ecco et al., 2016; Imbeault et al., 2017), presumably by recruitment of H3K9 methyltransferases. Additionally, a high concordance between heterochromatic repeat regions and some transcription factor-binding sites in mouse has been observed. For example, paired box 3 (Pax3) and Pax9 have binding sites in pericentric heterochromatin and depletion of these factors results in derepression of satellite transcripts and impairment of heterochromatic marks (Bulut-Karslioglu et al., 2012). The data points to a model where specific DNA-binding proteins could attract silencing effectors to their DNA-bound region. Interestingly, in some cases, binding of specific transcription factors to euchromatic sites appears to trigger the formation of small blocks of heterochromatin to achieve silencing. In *Drosophila*, phosphorylated Hers (Histone gene-specific Epigenetic Repressor in late S phase) binds to histone gene regulatory regions and anchors HP1a and the H3K9 methyltransferase Su(var)3-9 to induce silencing of this repeated gene cluster (Ito et al., 2012). Similarly, the corepressor TIF1b/KAP1 (Transcription Intermediary Factor 1-

beta/KRAB-associated protein 1) targets HP1 to specific loci in the euchromatic arms to silence those genes (Cammass, 2004; Schultz, 2002). Importantly, in contrast to pericentric heterochromatin, there is little spreading (see below) of the silencing marks in these cases, suggesting a silencing state that differs in some key characteristics from classic heterochromatin formation. Unfortunately, there has been little study of this process to date.

Once established, histone methylation serves as a molecular anchor. In *Drosophila*, di- or tri-methylation of H3K9 by the HMT Su(var)3-9 provides binding sites for HP1a resulting in a stable interaction (Eskeland et al., 2007; Jacobs et al., 2001). Chromatin-bound HP1 in turn mediates the recruitment of more HMTs Su(var)3-9, either directly or through the bifunctional binding partner, Su(var)3-7 (Delattre et al., 2000). Su(var)3-9 is the main producer of H3K9 methylation and as long as it carries out this reaction on an adjacent histone, the heterochromatin assembly process continues. By recognizing both the histone modification and the enzyme responsible for that modification, HP1a provides a mechanism for heterochromatin spreading and maintenance (Fig10). Similarly, in fission yeast, nucleated heterochromatin repeat elements can spread in a Swi6/HP1-dependent manner (Hall, 2002). A key feature of heterochromatin is thus its ability to propagate, thereby influencing gene expression in a region-specific, sequence-independent manner.

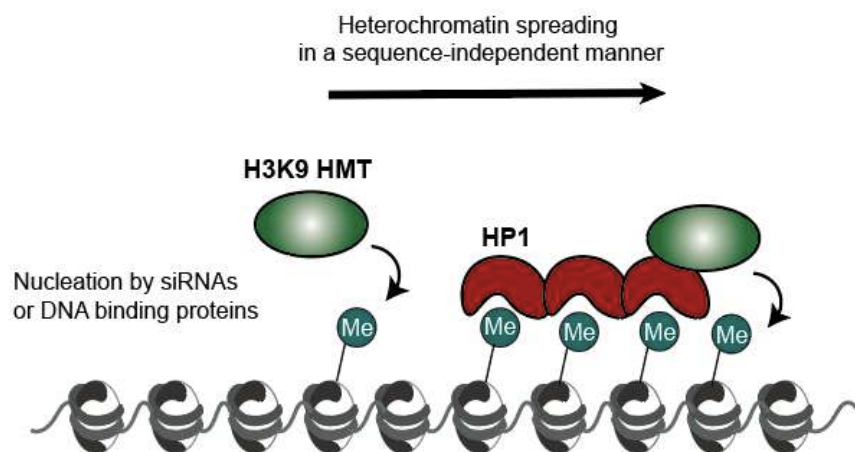


Fig 10-H3K9me3/HP1 heterochromatin spreading

Once nucleated, di- and tri-methylated H3K9 provide a binding platform for HP1 proteins. In turn HP1 can recruit H3K9 methyltransferases and the process is repeated, thereby spreading heterochromatin structure in a region-specific manner. Abbreviations: HMT= histone methyltransferase, Me= methyl mark.

Because heterochromatin can spread, mechanisms to restrict its expansion are necessary to avoid erroneous and potentially deleterious gene silencing. tRNA genes are a class of heterochromatin-spreading barrier conserved from yeast to human (Raab et al., 2012; Scott et al., 2006). Binding sites for transcription factors associated to the RNA Pol III can function as heterochromatin barriers independent of tRNA genes, revealing their role in barrier function. Another mechanism to restrain heterochromatin spreading is the generation of nucleosome depleted regions. This is observed at the silent mating type region in fission yeast which contains large nucleosome free regions that form a ‘gap’ in chromatin over which some reader–writer machineries cannot cross (Garcia et al., 2010). The existence of multiple mechanisms to limit heterochromatin spreading highlights the importance of proper chromatin organization in maintaining genome homeostasis.

Interestingly, HP1a can also have a role in transcription of genes residing in heterochromatin. In *Drosophila*, two genes, *light* and *rolled*, found within a heterochromatin domain show a loss of expression upon depletion of HP1a (Lu et al., 2000; Wakimoto and Hearn, 1990). There is a loss of silencing marks at the TSS of active genes in these domains although the usual heterochromatic marks, including H3K9me_{2/3}, are still present upstream and across the gene body. The TSSs are occupied by RNA Pol II and are flanked downstream by nucleosomes with euchromatic marks. This data shows that the presence of H3K9 methylation on the gene body, but not on the TSS, is compatible with transcription. Further work will be required to determine if the specific localization of these marks has a role on active transcription.

Furthermore, a non-repressive role for H3K9me₃ was described in a recent study by the Torres-Padilla laboratory. Immediately after fertilization, the paternal pronucleus of mammalian embryos acquires *de novo* H3K9me₃ via catalysis by the methyltransferase SUV39H2 which in turn is negatively regulated by satellite RNAs transcribed from the paternal pericentromere. *De novo* H3K9me₃ is initially non-repressive for gene expression, but instead bookmarks promoters for compaction later during development, suggesting that the mark is not repressive per se (Burton et al., 2020). This study provides an illustration of the complex role of heterochromatin and demonstrates that heterochromatin function can vary depending on the cellular and developmental context.

In conclusion, H3K9/HP1 heterochromatin is a potent silencing mechanism. However, the actors involved in this pathway seem to have a more complex role in transcriptional regulation than just promote transcriptional repression.

5. Polycomb effectors mediate gene silencing...and more?

The Polycomb group (PcG) genes were first discovered in *Drosophila* as crucial epigenetic repressors of homeotic (*Hox*) genes. They have since then been discovered to control hundreds of genes in insects, mammals and many other branches. Indeed, PcG genes and targets are conserved in evolution and their study has allowed unveiling regulation of a plethora of cellular processes.

The PcG machinery is mainly composed of two biochemical complexes: Polycomb Repressive Complex 1 (PRC1) and PRC2. PRC1 contains the E3 ligase activity for monoubiquitylation of H2A at Lys 118 in *Drosophila* and Lys 119 in mammals (Wang et al., 2004). PRC2 contains histone methyltransferase activity specifically targeting H3K27 (Cao et al., 2002; Czermin et al., 2002; Kassis et al., 2017; Kuzmichev, 2002; Müller et al., 2002; Schuettengruber et al., 2017). H3K27me₃ is found on many silenced regions in a cell-specific manner (Beuchle et al., 2001; Plath, 2003; Ringrose and Paro, 2004). This mark is deposited by the histone methyltransferase E(z)/EZH2, a subunit of the PRC2 complex and its catalytic activity is required for *Hox* gene repression (Czermin et al., 2002; Müller et al., 2002). Additionally, it was observed that flies carrying a point mutation in lysine 27 of H3 fail to repress PRC2-target genes demonstrating thus the need of this mark for PcG-mediated repression (Pengelly et al., 2013). Contrariwise, it was observed that the mark catalyzed by the PRC1 complex is dispensable for repression of canonical PcG targets during *Drosophila* embryogenesis (Kahn et al., 2016; Pengelly et al., 2015). H2AK118ub is carried out by the PRC1 subunit Sce in *Drosophila*. Flies with catalytically inactive Sce or with pointmutated H2A were generated and it was observed that H2Aub-deficient animals fully maintain repression of PRC1 target genes. These data suggest thus that catalytic activity of PRC2, *i.e.*, H3K27 methylation, is necessary for repression however PRC1 represses canonical targets independently of its catalytic activity.

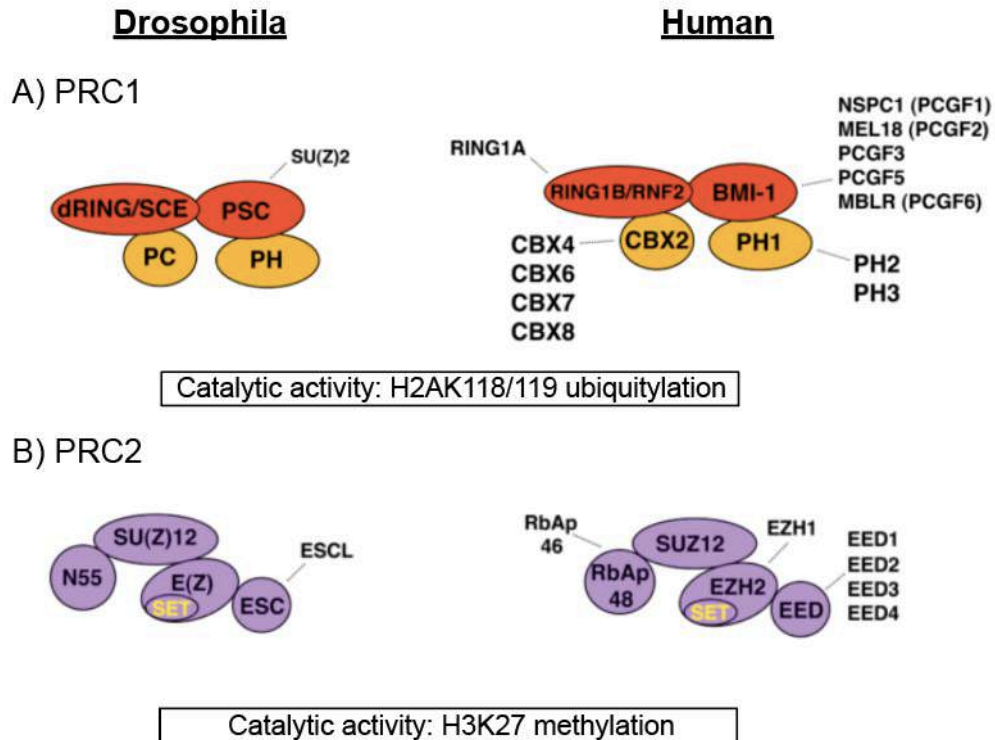


Fig 11-Composition of Polycomb Repressive Complexes (PRC) in *Drosophila* and Human.
 (A) PRC1 complex contains four core subunits. Its catalytic activity is ubiquitylation of H2AK118 in *Drosophila* and H2AK119 in mammals, conferred by the dRING/RING subunit.
 (B) PRC2 complex contains four core subunits. Its catalytic activity is methylation of H3K27 conferred by the subunit E(z)/EZH2. Dashed lines indicate alternative subunits (Simon and Kingston, 2013).

What is then the role of H3K27me3? It has been suggested that H3K27me3 stabilizes interactions between DNA-bound PcG complexes and the surrounding chromatin. It was observed that abolishing H3K27me3 within a promoter or transcriptional unit of a target gene impairs the interactions of DNA-anchored PcG complexes with these gene elements (Kahn et al., 2016). By promoting interaction stability with these gene regions, H3K27me3 would deliver PcG complexes that can interfere with transcription. Stable loops could also contribute to spreading and maintenance of H3K27me3 from its anchor points, thereby reinforcing the system. Additionally, a bulk tri-methylation of H3K27 may contribute to the repression directly by competing with acetylation of H3K27, a mark involved in gene activation. Furthermore, PRC1 subunit Pc/CBX can inhibit the acetyltransferase activity of dCBP/CBP (Tie et al., 2016, 2009). Even if the importance of H3K27me3 for polycomb-

mediated silencing has been shown, the precise mechanism by which it allows gene repression is not fully understood.

Initially, it was observed that PcG proteins and H3K27me3 marked histones accumulate in discrete nuclear foci: “Polycomb bodies” (Buchenau et al., 1998; Messmer et al., 1992), indicating that nuclear compartmentalization could play a functional role in Polycomb-mediated repression. Interestingly, interactions with elements of the nuclear periphery seem to be important for the formation of these compartments and for gene repression. Indeed, the knock-down of lamin A/C causes the dispersion of Polycomb bodies and leads to impaired PcG-mediated transcriptional silencing in both mammalian and *Drosophila* cell lines (Cesarini et al., 2015; Marullo et al., 2016). Nuclear compartmentalization could thus play an important role in PcG-mediated silencing.

Although targets and PcG proteins are well conserved, the mechanisms to recruit them at target sequences vary between species. The *Drosophila* genome is equipped with, Polycomb Response Elements (PREs), these are discrete DNA elements to which PRC1 and PRC2 are targeted. These elements contain binding sites for many different DNA-binding proteins such as Pho, GAF/Psq, Dsp1, Spps, Zeste, Grh, Adf1, Cg and many others (Blastyák et al., 2006; Brown and Kassis, 2010; Déjardin et al., 2005; Orsi et al., 2014; Ray et al., 2016). These TFs play an important role in PcG recruitment however none of them is sufficient to recruit PcG complexes on their own. Initially, a hierarchical recruitment model was proposed in which TFs recruit PRC2, which subsequently recruits PRC1 via the interaction of the Pc subunit with the PRC2-deposited mark, H3K27me3. Accumulating evidence has challenged this model (Dorafshan et al., 2017). For instance, it has been observed that PRC1 can bind PREs in the absence of PRC2 but at many PREs, PRC2 requires PRC1 to be targeted (Kahn et al., 2016). Additionally, PRC1-bound regions devoid of H3K27me3 exist in both *Drosophila* and human cultured cells (Loubiere et al., 2016; Schwartz et al., 2006). Moreover, none of the histone marks deposited by PRCs complexes are required for their targeting at PREs, suggesting that PREs work upstream to histone mark deposition (Kahn et al., 2016). The relationship among TFs, PRC1, and PRC2 seems thus cooperative rather than hierarchical. Whilst PRE characterization in flies is well advanced, their mammalian counterpart has not been clearly identified. In mammals it has been shown that CpG islands can recruit PcG proteins (Lynch et al., 2012; Mendenhall et al., 2010; Reddington et al., 2013), a genomic feature (CpG islands) that has not been found in flies. Nonetheless, similarly to *Drosophila*, *de novo* PRC2 recruitment is independent of

H3K27me3 and H2AK118ub in ESCs (Lavarone et al., 2019). Some mammalian PREs have binding sites for TF important at *Drosophila* PREs (Kassis and Brown, 2013). For example, a homolog of the *Drosophila* protein Polycomb-like, can bind DNA and recruit PRC2 in mouse ESCs. It selectively binds regions with high density of unmethylated CpGs which discriminate target from non-target CpG islands (Li et al., 2017; Perino et al., 2018). The quest for mammalian PREs is still a field of active ongoing research that will shed light into the specific targeting PcG proteins in mammals.

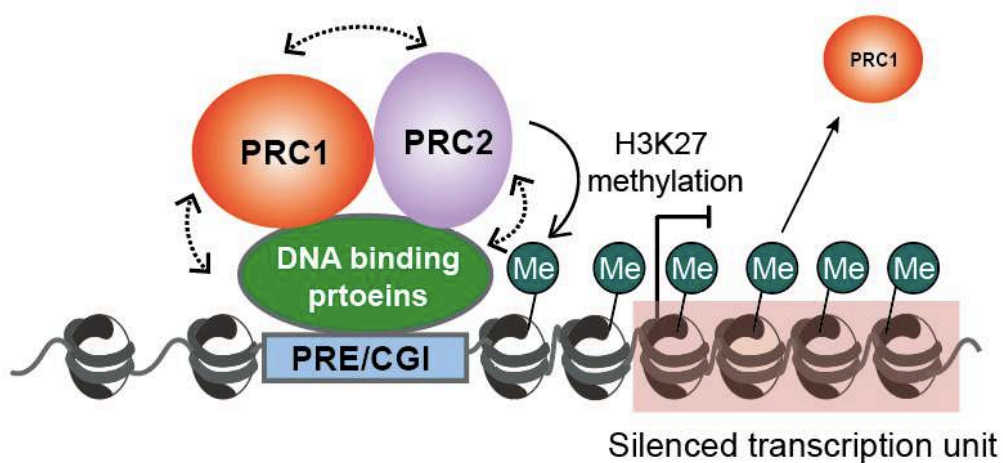


Fig 12-Polycomb Repressive Complex recruitment

PcG proteins are recruited to specific DNA sequences, in flies these are Polycomb Response Element (PREs). In mammals CpG islands have been found capable of recruiting PcG proteins. PREs contain binding sites for numerous DNA binding proteins important for Polycomb Repressive complexes (PRC) recruitment. Dashed arrows represent a cooperative relationship between DNA binding proteins and PRCs to target them at PREs/CGIs. PRC2 catalyzes H3K27me3 necessary for transcriptional silencing and PRC1 is capable of binding this modification thereby stabilizing their recruitment to chromatin.

Although many elements required for targeting and establishment of PcG-mediated repression are known, more and more evidence are revealing other roles for the actors involved. Firstly, fragments of DNA that contain PREs have been shown to mediate gene activation in transgenes under certain conditions and at some chromosomal insertion sites (Kassis and Brown, 2013), suggesting that these regulatory elements can have dual regulatory functions. Indeed, during *Drosophila* embryogenesis, some classic PREs can recruit tissue-

specific TFs and function as developmental enhancers *in vivo*, activating spatio-temporal expression of a reporter gene (Erceg et al., 2017). Conversely, a subset of developmental enhancers binds the PcG complex, Pho-RC, resulting in polycomb-dependent transcriptional silencing. This dual activity of *cis*-regulatory elements may help fine-tune gene expression and ensure the timely maintenance of cell identities. The same *cis*-regulatory element can thus have opposite outcomes on gene transcription in a context-dependent manner. Furthermore, the characterization of the *in vivo* role of the two PREs near the *vestigial* (*vg*) gene, reported a promoter that requires a PRE for expression during development (Ahmad and Spens, 2019). On one hand, the PRE near the *vg* promoter is required for its activation and not for repression. On the other hand, the distal PRE, located in the middle of the chromatin domain, is required for high-level of H3K27me3 in the domain. Surprisingly, removal of both PREs does not completely eliminate H3K27me3 across the *vg* domain and this residual methylation is similar to that in cells where the *vg* gene is active. These data confirms that PREs can also have a role in gene activation and the definition of their roles might depend on the factors they bind.

Interestingly, the binding of factors implicated in PcG silencing are not incompatible with transcription. PRC1 components were found to bind actively transcribed genes in *Drosophila* larval imaginal discs and human cultured cells (Loubiere et al., 2016). Strikingly, despite the proven need of H3K27me3 for polycomb-mediated repression, its presence was not incompatible with transcription of the *vg* gene (Ahmad and Spens, 2019). Normal transcription of the *vg* gene occurs concomitantly with a basal enrichment of H3K27me3 at the endogenous locus. This shows the complexity of the roles of the PcG proteins.

We also mentioned that PRC1 can be recruited independently of H3K27me3, suggesting that it might have functions independently of PRC2. It was observed in *Drosophila* that PRC1 components are targeted to a distinct set of genes that lack H3K27me3 during larval development (Loubiere et al., 2016). The redeployment of PRC1 was also observed in human differentiated cells when compared to embryonic stem cells. In both species PRC1-only gene targets were involved in regulation of cell proliferation, polarity and signaling. Mutations in PRC1, but not PRC2 resulted in the upregulation of a majority of these new target genes, suggesting that PRC1 has a role dampening a specific subset of genes independently of PRC2. Moreover, RNAi depletion of PRC1 subunits in *Drosophila* cell culture alters phosphorylation of RNA Pol II at most active genes and enhancers (Pherson et al., 2017). These effects coincide with changes in nascent RNA density indicating altered transcriptional

elongation and RNA processing. Although the mechanisms by which PRC1 can alter gene transcription are not fully elucidated, these findings describe a role for PRC1 components beyond epigenetic silencing.

Overall, these data reveal the complexity of the Polycomb system showing that factors and DNA regulatory elements involved can play different roles in a context-dependent manner, hence the inability to establish absolute rules for its role in transcriptional regulation.

6. Counteracting repression to promote activation

Gene repression can be accomplished by creating a chromatin configuration that blocks access to the large number of proteins required for transcription. A way to promote gene activation is thus to counteract silencing mechanisms.

There is a regulatory interplay between epigenomic effectors that supposes a balance between their activities in order to ensure the proper transcriptional outcome. One of the most well-studied epigenetic regulatory systems is the Polycomb group (PcG) and Trithorax group (TrxG) genes. These two groups of genes were discovered in *Drosophila* by their opposing effects on homeotic gene (*Hox*) expression. PcG proteins exert a negative effect on transcription while TrxG are positive regulators. Given the complexity of factors required for gene expression, the biochemical nature of TrxG proteins is a very heterogeneous group that includes remodelers and histone modifiers, acting at different levels of gene transcription (Kassis et al., 2017; Kingston and Tamkun, 2014; Schuettengruber et al., 2017).

How can regulators of these groups achieve antagonistic effects at their targets? One proposed mechanism by which TrxG can promote activation is by counteracting Polycomb-mediated repression. This was proposed for the TrxG histone methyltransferases Trx and Ash1. Indeed, genetic studies in *Drosophila* showed that the removal of PcG complexes reactivates genes even in the absence of Trx and Ash1, suggesting that these proteins function as PcG antirepressors rather than direct activators (Klymenko and Müller, 2004). Trx and Ash1 methyltransferases target H3K4 and H3K36 respectively, both of which are considered active marks (see section II-7). In agreement with genetic studies, *in vitro* assays using

engineered methylated histones showed that H3K4me3 or H3K36me2/3 can inhibit H3K27 methylation on the same histone by fly or human PRC2 complex (Schmitges et al., 2011). Therefore, these data suggest mechanistic antagonism between these two groups could be mediated by histone PTMs.

As mentioned, the TrxG encompasses a great diversity of biochemical activities indicating that mechanisms involving other type of epigenomic effectors could also be at play. Several subunits of the nucleosome remodeler SWI/SNF complex, including the catalytic subunit, have been classified as TrxG proteins. The opposition between PRCs and the SWI/SNF complex has been the focus of extensive research especially because of its implication in development and disease such as cancer (Kadoch and Crabtree, 2015). Recent studies have unveiled the mechanisms for the interaction between these two groups of regulators. It was observed in cultured cells, that recruitment of the mammalian SWI/SNF complex leads to a rapid, ATP-dependent eviction of both PRC1 and PRC2 (Kadoch et al., 2017). The reversal of this process results in reassembly of Polycomb-mediated heterochromatin. This study proposes thus a mechanism in which a TrxG remodeler opposes Polycomb complexes by their active removal, resulting in chromatin accessibility (Fig13). Interestingly, a recent study reported that degradation of the SWI/SNF ATPase subunit in mouse ESCs resulted in derepression of genes highly occupied by Polycomb, such as *Hox* genes, suggesting also a role for SWI/SNF in promoting PcG-based repression (Weber et al., 2021). Upon rapid depletion of the catalytic SWI/SNF subunit, PRC1 and PRC2 are redistributed away from sites where they usually accumulate, like *Hox* clusters. Collectively, these findings reveal the complexity of the regulatory interplay of the Polycomb-Trithorax axis to achieve proper gene expression.

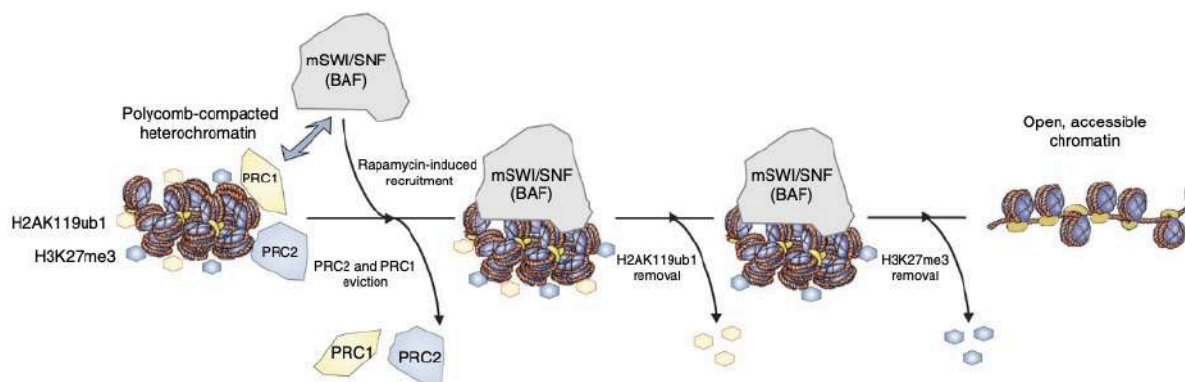


Fig 13-Model for opposition between mSWI/SNF and Polycomb

Recruitment of the mSWI/SNF complex to Polycomb-repressed regions leads to an ATP-dependent eviction of PRC1 and PRC2 complexes, followed by H2AK119 and H3K27me3 removal resulting in an increase of DNA accessibility (modified from Kadoch et al., 2017)

Another mechanism used to counteract repression is to create a boundary between regions of different transcriptional activities. In fly embryos, insulators and their insulator binding proteins (IBPs) can be found at borders between different chromatin landscapes and block the spreading of histone modifications at *Hox* genes (Bowman et al., 2014; Fujioka et al., 2013; Kahn et al., 2006). Additionally, the IBPs, Beaf-32, CP190, dCTCF and Mod(mdg4) can be found enriched at promoter pairs of differentially expressed genes (Nègre et al., 2010) suggesting that insulators and their binding proteins can indeed separate different transcriptional states. However, in *Drosophila* cultured cells, knock-downs of the mentioned IBPs led to changes of the repressive mark H3K27me3 within the chromatin domain rather than spreading outside of it (Van Bortle et al., 2012). IBPs can indeed be found enriched at borders of H3K27me3 domains but only a fraction of these sites can actually restrict the mark's spreading (Schwartz et al., 2012). One possibility is that additional co-factors are necessary to inhibit the repressive mark spreading. Such example was described in *Drosophila* embryonic cells, where the IBP Beaf-32 is capable of recruiting the H3K36 histone methyltransferase dMes-4 to promote transcription of flanking genes by antagonizing the spread of H3K27 methylation from nearby regions (Lhoumaud et al., 2014). More work is needed to identify this type of co-factors and their role at heterochromatin/euchromatin borders. Nonetheless, these data suggest a cooperation between epigenomic effectors to maintain appropriate transcriptional states.

Altogether, this suggests that repression counteracting is the result of an interplay between epigenomic effectors. Although this can contribute to promote gene expression, activation of a gene requires numerous steps and actors and thus many other mechanisms involved.

7. Histone PTMs in active transcription

As for a silenced state of transcription, histone modifications can also be associated to an active state of transcription. For example, genome-wide studies have found H3K27ac and

H3K4me1 at active enhancers, H3K4me3 and H3/H4 acetylation are found at promoters of active genes, H2Bub1 at promoters and gene bodies and finally high levels of H3K79me3 and increasing H3K36me3 towards the 3' end are found at actively transcribing genes (Fig14) (Gates et al., 2017). While these correlations have been widely described, determining which modifications are causal drivers and by which mechanisms they promote transcription is not completely understood. Nonetheless, roles for histone marks, particularly during the elongation step, have been described.

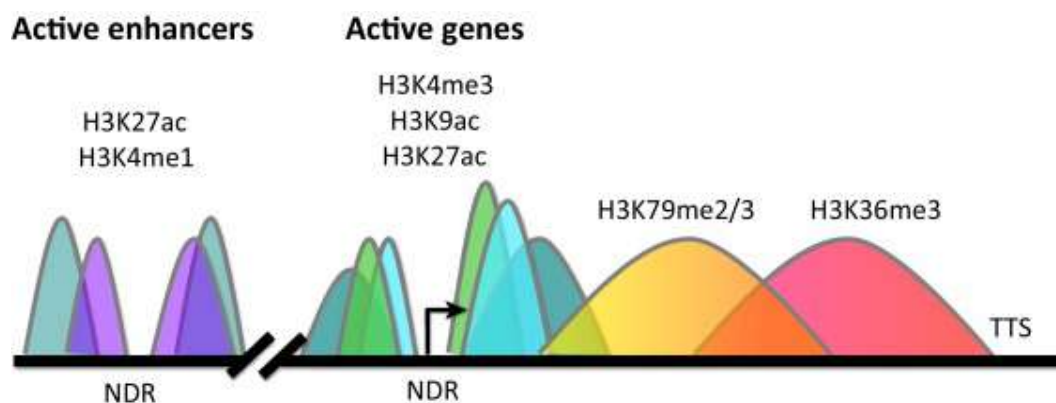


Fig 14-Histone marks at active genes.

Non-exhaustive illustration of the distribution of histone modifications at active genes and enhancers. Regions in color represent ChIP-signal at eukaryotic genes. H3K4me3 and H3K9ac are associated with transcriptionally active gene promoter regions while H3K36me3 and H3K79me3 are localized at gene bodies. H3K27ac localizes to both active gene promoters and enhancer regions, and H3K4me1 is predominantly enriched at enhancers. Abbreviations: NDR= nucleosome depleted region, TTS = transcription termination site, the transcription start site is shown by an arrow (adapted from Gates et al., 2017).

A. Histone PTMs during transcription elongation

Trimethylation of H3K36me3 during elongation has been found important for maintaining transcription fidelity. Studies in budding yeast have indeed investigated the function of H3K36 methylation during transcription. Successive transcription may cause histone hyperacetylation in gene bodies leading to cryptic transcription initiation. The RNA Pol II phosphorylated on the Serine 2 *i.e.*, the elongating Pol II, recruits the histone methyltransferase Set2 to deposit H3K36me3 which serves as a docking site for the Rpd3S

histone deacetylase complex to actively transcribed genes. In turn, Rpd3S catalytic activity, ensures that coding regions remain hypoacetylated, which is important to suppress cryptic transcription initiation (Carrozza et al., 2005; Drouin et al., 2010). Additionally, it was later observed that longer genes and genes transcribed at lower frequency had a stronger dependency on the Set2-Rpd3S pathway to suppress spurious transcription (Li et al., 2007). Transcription elongation-coupled SETD2 recruitment and H3K36me3 deposition are conserved in mammals, as well as the role of H3K36me3 to prevent aberrant transcription initiation [Huang and Zhu, 2018]. However, in mouse embryonic cells it was observed that this depended on the recruitment of DNA methyltransferases to the gene body by H3K36me3 (Neri et al., 2017). Although mechanisms to prevent spurious transcription may be different, these data highlight the importance of H3K36me3 in maintaining transcription fidelity.

Another mark associated with transcription elongation is the monoubiquitination on histone H2B (H2Bub1). This mark has been found enriched at promoters and open reading frames of actively transcribed genes (Kao, 2004; Minsky et al., 2008), suggesting thus a role in gene activation. *In vitro* experiments showed that H2B monoubiquitination by the ubiquitin ligase RNF20/40 jointly with the ubiquitin-conjugating enzyme UbcH6 facilitates displacement of the H2A/H2B dimer from the core nucleosome, facilitating RNA Pol II passage and enhancing transcript elongation rates on chromatin templates (Pavri et al., 2006). This effect was observed while H2ub1 remained in the chromatin but *in vivo* observations in yeast showed that both ubiquitination and deubiquitination are important for full transcription activation at the *Gall* promoter (Henry, 2003). Further research is needed to understand the full extent of this histone mark on transcription initiation and elongation.

These data show the implication of histone PTMs in elongation and reflects how the epigenome can regulate gene expression at different points of the transcription process.

B. Is H3K4me3 instructive for transcription activation?

A connection between H3K4 methylation and transcriptional activity was first evoked when this mark was observed to decorate the transcriptionally active macronucleus but not the transcriptionally inert micronucleus in the single-celled eukaryote *Tetrahymena*. (Strahl et al., 1999). When mapped in genome-wide experiments, the histone mark H3K4me3 was

found at TSSs of transcriptionally active genes in various organisms and the levels of this mark were strongly correlated to nascent transcripts (Bernstein et al., 2005; Guenther et al., 2007; Heintzman et al., 2007; Howe et al., 2017; Nojima et al., 2015). This led to the classification of H3K4me3 as a hallmark of active promoters resulting in the use of it as a marker of active transcription. Supporting this view, studies in human cultured cells showed that H3K4me3 interacts with TAF3, a subunit of the general transcription factor TFIID, facilitating thereby the formation of the transcriptional pre-initiation complex (Lauberth et al., 2013). This resulted in the enhancement of specific p53-dependent transcription in response to genotoxic stress. A more recent study, again in human cells, employed epigenome editing to investigate the role of H3K4me3 in transcriptional activation (Cano-Rodriguez et al., 2016). They used the methyltransferase PRDM9 to locally induce H3K4me3 and observed re-expression of silenced targets. However, maintenance of this re-expression was dependent on chromatin environment, particularly on hypomethylated DNA and H3K79me. Indeed, when H3K4me was targeted to a hypermethylated locus, re-activation was only transient. Yet, the precise contribution of these genomic features are to be determined. Although this study links H3K4me3 to gene (re-)activation, the chosen HMT possesses other targets on histones linked to transcription, for example H3K36 *in vivo*, and H3K9 at least *in vitro* [Koh-Stenta et al., 2014, Powers et al., 2016], none of which were examined. Further studies are required to elucidate if these observations are reproducible at other targets and organisms, particularly in those without DNA methylation like *Drosophila*.

Why is it difficult to study the impact of H3K4me3 on transcription? In yeast the Set1 complex is in charge of mono-, di- and trimethylation of H3K4, therefore altering this enzyme would also impact histone marks other than H3K4me3 (Gu and Lee, 2013). In contrast, human possesses several enzymes catalyzing H3K4 methylation, six Set1 homologs (SET1A, SET1B, MLL1 to 4) but also other enzymes non-related to Set1 (MLL5, SET7 (also called SET9), SMYD1-3, SETMAR, and PRDM9). In *Drosophila*, three Set1 homologs have been identified dSet1, Trithorax (Trx), and Trithorax-related (Trr). Deletion of any of their genes results in lethality in flies, indicating that their target genes may not be redundant. Additionally, loss of *dSet1*, but not *Trx* or *Trr*, leads to a global reduction of H3K4me2/3, suggesting that *Trx* and *Trr* have more specialized functions (Ardehali et al., 2011; Mohan et al., 2011). This diversity complexifies the study of the direct implication of only the H3K4me3 mark due to possible overlapping functions and/or indirect effects on other marks. Indeed, the field is still lacking a conserved model mechanism to support causality between

H3K4me3 and transcriptional activity and accumulating evidence has challenged this causal link.

Several studies in yeast revealed that absence of H3K4me3 has little effect on gene transcription even at genes where H3K4me3 is highly enriched (Margaritis et al., 2012; Ramakrishnan et al., 2016; Weiner et al., 2012). Accumulating evidence in *Drosophila* has also questioned the relevance of H3K4 methylation in transcriptional activation. One of *Drosophila*'s H3K4 HMT, Trx can be cleaved into two proteins Trx-C and Trx-N. Trx-N lacks methyltransferase activity and is associated with broad regions of active genes (Schuettengruber et al., 2009; Schwartz et al., 2010), suggesting that it activates transcription via mechanisms other than histone methylation. Furthermore, the replacement of H3K4 with a non-methylable arginine, revealed that despite the absence of H3K4 methylation, transcriptional activation was still possible and only minimal changes in developmental gene expression were observed (Hödl and Basler, 2012). Methylation of H3K4 is thus dispensable for transcriptional activation and this mark is likely to have a more complex role in chromatin regulation. Furthermore, a recent study in mouse ESCs showed that gene reactivation can occur without reacquisition of H3K4me3 (Douillet et al., 2020). These findings showed thus an uncoupling between transcriptional activation and the histone mark H3K4me3.

Overall, these data show that the role of H3K4me3 cannot be exclusively reduced to that of transcriptional activation.

8. Bivalent promoters

A peculiar class of promoters, known as bivalent promoters, were originally identified in ES cells and are characterized by the simultaneous enrichment of both H3K27me3, associated to gene repression and H3K4me3, commonly associated to gene activation (Berstein et al., 2006, Mikkelsen et al., 2007). The enrichment of these opposing modifications correlates with a low-level expression or no expression (Bernstein et al., 2006). During cell differentiation, these bivalent regions are resolved as they undergo either full transcriptional activation, in which case they preserve H3K4me3 and lose H3K27me3, or stable silencing where they preserve H3K27me3 and lose H3K4me3 (Gaertner et al., 2012; Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007). Interestingly, in more committed cells, neural progenitors and embryonic fibroblasts, bivalent domains were still present in

different proportions, 8% and 43% respectively (Mikkelsen et al., 2007). To explain this difference the authors proposed that the high number of bivalent domains in embryonic fibroblasts may reflect a less differentiated state and/or heterogeneity in the population analyzed. Nevertheless, at least in neural progenitors, genes where bivalent domains were not resolved, continued to be repressed. These data suggest thus a functional role for bivalent domains.

Genome-wide mapping of bivalent chromatin revealed it is frequently found within promoter regions of developmentally important genes (Bernstein et al., 2006; Lesch et al., 2013), providing a silencing system for these genes while keeping them poised for activation. Like this, during development, certain genes become active in a tissue-specific manner leading to cell lineage specification. These regions have been suggested to “safeguard differentiation” and their malfunction might have a profound impact on the cell (Voigt et al., 2013). Indeed, bivalent domains identified in human tumors, such as ovarian cancer, colon cancer, and glioblastomas, coincide with genomic regions decorated with H3K4me3/H3K27me3 in ESCs (Curry et al., 2018; Hall et al., 2018; Rodriguez et al., 2008). However, the extent of this overlap and the biological significance of these recovered bivalent domains in cancerous cells remains to be investigated.

Despite its apparent importance, bivalency is functionally and mechanistically not well understood. A study in ESCs revealed however that nucleosomes carrying H3K4me3 along with H3K27me3 did so on opposite H3 tails. PRC2-mediated methylation of H3K27 was inhibited when nucleosomes contained symmetrically, but not asymmetrically, positioned H3K4me3, showing that the location of the “active” mark can be decisive for establishment of bivalency (Voigt et al., 2012). Further studies should elucidate the requirements for establishment of these chromatin domains.

Until recently, bivalent domains had not been identified in *Drosophila*. Akmammedov and colleagues used re-ChIP to confirm the co-occurrence of H3K27me3 and H3K4me3 in fly embryos, leading to the uncovering of such bivalent domains (Akmammedov et al., 2019). However, only a handful of endogenous sites, all members of the *Hox* genes, were tested. More studies will be needed to see if these domains are present across the genome of *Drosophila* at key developmental genes and if their state of bivalency changes during differentiation.

9. Genome folding and transcription

DNA and its associated proteins are confined within the nucleus. This is a topologically constrained and a highly crowded environment. The 3D folding of the eukaryotic genome compatible with all the activities described above is thus a major requirement.

The study of a functional link between higher-order chromatin arrangements and transcription has greatly increased thanks to the rapid development on new techniques. In the mid-2000s, FISH experiments in human fibroblasts, showed the existence of chromosome territories (Bolzer et al., 2005). These territories segregate in regions rich in active genes, typically located in the interior of the nucleus, and regions rich in inactive genes, found at the nuclear periphery. The rapid development of chromosome capture techniques, namely C-based techniques, now allow to assay contact frequency at a genome-wide level. At large scales, Hi-C confirmed two major types of structural domains: A and B compartments. The A compartment corresponds to active chromatin, presents transcriptional activity, higher chromatin accessibility and H3K36me3 deposition. Compartment B contains repressed chromatin in a more compacted state, with low transcriptional activity, associated with the nuclear lamina and presenting H3K27me3 deposition (Lieberman-Aiden et al., 2009; Rao et al., 2014). At a lower scale, chromosomes fold into domains of 100kb-1Mb, with preferential intradomain interactions compared to interdomain interactions. These contact domains are partitioned by boundaries between them and are referred to as topologically associating domains (TADs) (Fig15) (Dixon et al., 2012). The presence of TADs has been confirmed across cell lines and species, indicating that they may represent a conserved feature of genome organization. Importantly, a conserved characteristic across species, is the relationship between gene activity and genome folding (Szabo et al., 2019). Overall, these observations point towards a functional implication of chromosomal organization within the nucleus.

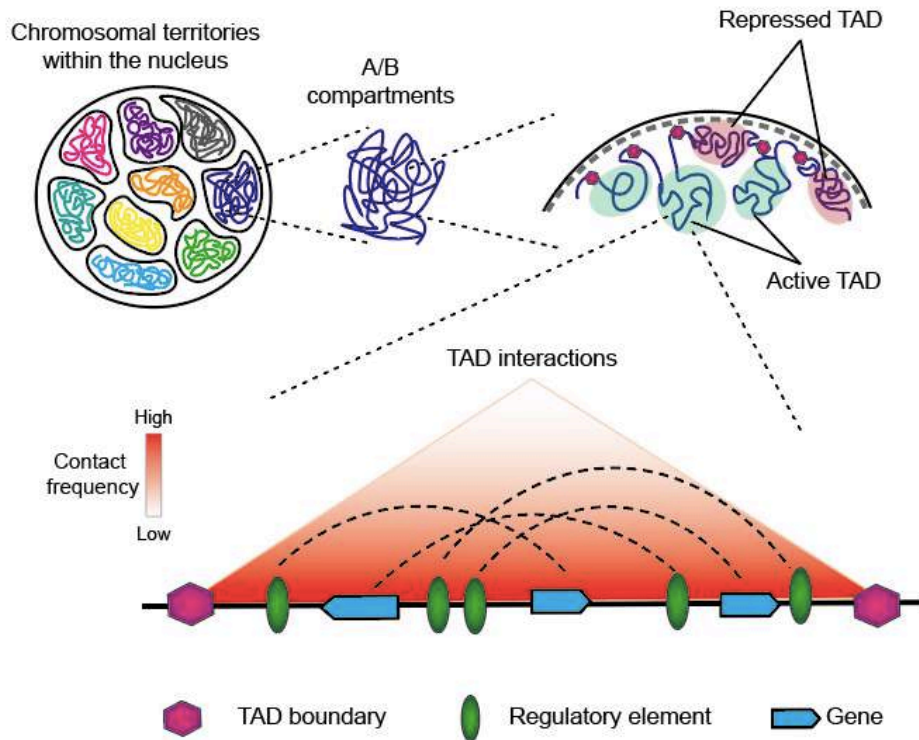


Fig15-Hierarchical organization of the eukaryotic genome

Individual chromosomes occupy specific nuclear spaces, forming chromosomes territories, shown schematically in different colors. At the chromosomal scale, chromatin is segregated into active “A” and repressed “B” compartments. At a finer scale the chromatin fibers are partitioned into higher-order domains of preferential internal interactions defined by boundaries referred to as Topologically Associating Domains (TADs). An example of an active TAD with several interactions between distal regulatory elements and genes within it is shown (Matharu and Ahituv, 2015).

Studies in mammalian cells have shown that disruption of TADs can lead to *de novo* interactions between regulatory elements, such as enhancers, and promoters, resulting in gene misexpression and disease (Flavahan et al., 2016; Franke et al., 2016; Lupiáñez et al., 2015). To fully understand the role of TADs in transcriptional regulation it is thus necessary to understand their formation and maintenance. Notably, TAD boundaries in mammals are frequently enriched in both the transcription factor and insulator CTCF and the structural maintenance of chromosomes (SMC) cohesin complex (Dixon et al., 2012; Phillips-Cremins et al., 2013; Rao et al., 2014) Hi-C maps indicate a strong contact between the CTCF- and cohesin-bound TAD boundaries *i.e.*, “corner peaks”, suggesting a model wherein CTCF binds its cognate sites and recruits cohesin, which then folds the in-between chromatin into a loop structure. Interestingly, removal or change in orientation of a single CTCF site can abolish or

shift the position of the TAD boundary (de Wit et al., 2015; Guo et al., 2015; Lupiáñez et al., 2015), suggesting that not only the presence of CTCF but also a convergent orientation is important for TAD boundaries in mammals. A loop extrusion model has been proposed (Fig16), in which extruding factors, the engaged cohesin SMC complex, progressively forms larger chromatin loops until it encounters boundary proteins, including CTCF, or until the complex is dissociated (Fudenberg et al., 2017, 2016). This model has been reinforced by recent studies that show that depletion of CTCF, cohesin or its loading factor disrupt loop domains while depletion of the cohesin release factor reinforces the strength of the loops at TAD borders (Nora et al., 2017; Rao et al., 2017; Wutz et al., 2017), suggesting a determinant role for cohesion in loop formation.

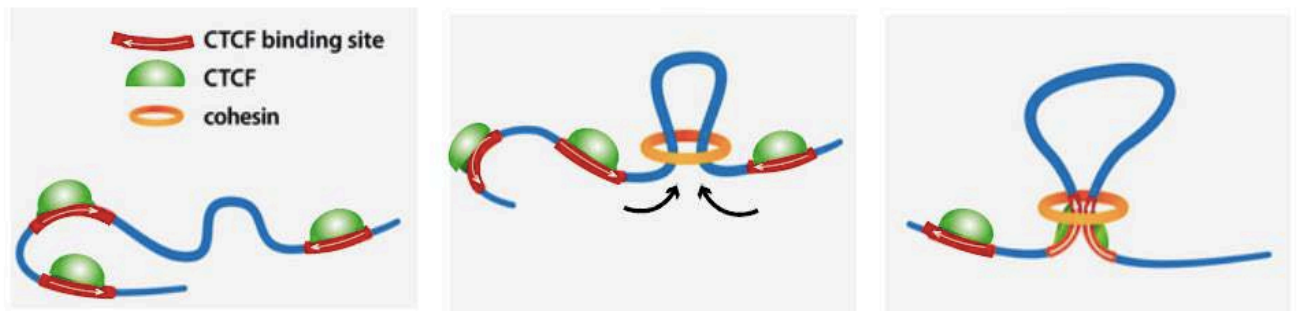


Fig16-Model for loop extrusion

Cohesin binds to chromatin, extrudes a chromatin loop and stalls upon encountering CTCF bound at convergently oriented CTCF sites (Wutz et al., 2017).

TADs are also present in the *Drosophila* genome; however, differences have been observed compared to mammalian TADs. Cohesin is also enriched at TAD borders but there are no interaction loops at these borders and enrichment of dCTCF is minor (Van Bortle et al., 2014). Nonetheless, various insulator proteins have been found enriched at boundaries, including Beaf-32, CP190, Pita, M1BP and Chromator (Cubeñas-Potts et al., 2017; Ramírez et al., 2018; Sexton et al., 2012). Also, combinations of these proteins such as Beaf-32/Chromator or Beaf-32/CP190 are good predictors of boundaries (Wang et al., 2018). Surprisingly, siRNA-mediated depletion of Beaf-32 does not abolish boundaries nor has a significant effect on chromatin interactions (Ramírez et al., 2018). A possibility evoked by the authors is that the transcription factor Dref, who binds almost the exact motif as Beaf-32,

could also play a role in chromatin organization. However, no data to support this hypothesis was presented. It is also possible that depletion of a combination of factors is required to disrupt TAD boundaries in *Drosophila*. Further work is thus needed to determine the role of insulator proteins and of other proteins, such as Dref, at *Drosophila* TAD boundaries.

Other mechanisms could also be responsible for 3D organization of chromosomes. Indeed, in mammals even though a majority of TAD boundaries were associated to CTCF, a fraction turned out to be resistant to CTCF loss (Nora et al., 2017). It was observed in mammals and *Drosophila* that TAD borders were enriched for housekeeping genes (Cubebñas-Potts et al., 2017; Dixon et al., 2012; Rennie et al., 2018; Ulianov et al., 2016), those are zones of active transcription hinting thus on a possible role for this process at TAD boundaries. Furthermore, modeling of chromatin fibers proposes that transcription-associated supercoiling could be involved in driving loop extrusion (Racko et al., 2018), suggesting a role for transcription in TAD formation. Based on Hi-C and RNA-seq data in four fly cell lines of various origins, Ulianov and colleagues proposed that inactive TADs are separated by active chromatin regions (Ulianov et al., 2016). In agreement, super-resolution analysis of immunolabeled repressive H3K27me3 or “active” H3K4me3 marks showed active domains at the borders of repressed ones (Boettiger et al., 2016; Cattoni et al., 2017). Moreover, a recent study observed that the degree of transcriptional activity correlates with the strength of TAD insulation (Luzhin et al., 2019). However, inhibition of transcription in mammals and *Drosophila*, does not abolish TAD boundaries (Du et al., 2017; Hug et al., 2017; Ke et al., 2017) nor its induction is sufficient to create TAD boundaries *de novo* (Bonev et al., 2017). It is therefore likely that additional factors are necessary for TAD formation.

Interestingly, polymer simulations using as single input the experimentally derived epigenome from *Drosophila* embryonic cells, agreed with the folding patterns observed in chromosome conformation capture experiments (Ghosh and Jost, 2018; Jost et al., 2014). Additionally, Ulianov and colleagues proposed a “self-assembly” model where nonacetylated nucleosomes from inactive chromatin aggregate whereas acetylated nucleosomes in inter-TADs and TAD boundaries are less prone to interact (Ulianov et al., 2016). These data suggests that the epigenome is a primary driver of chromosome folding in *Drosophila*. Various factors including, transcription, insulator/architectural proteins, epigenetic marks seem to have a role in TAD formation. Further study is thus necessary to determine which

elements are required to form and maintain these compartments separating transcriptional states. This will also help to understand the transition from one domain to another.

As mentioned before, at a larger scale, chromatin in the nucleus is also segregated according to its transcriptional state. Recent research suggests that liquid-liquid phase separation can result in these non-membrane bound compartments in cells. These studies propose that the nucleus is a phase separated compartment containing several different immiscible liquid-like sub-compartments. Segregation of heterochromatin is driven by phase separation mediated, at least in part, by multivalent hydrophobic interactions of HP1a (Larson et al., 2017; Strom et al., 2017). Additionally, clusters of enhancers, regulating cooperatively gene expression can undergo phase separation by transcriptional coactivators suggesting that active domains may also generate phase-separated compartments (Sabari et al., 2018). Together, these physical forces may account at least in part for the compartmentalization of the nucleus.

Importantly, most data gathered from C-based techniques come from a large number of cells and therefore, only averages are observed. To overcome this, chromatin conformation capture-based techniques have been extended to single cell analysis. These studies have revealed heterogeneity in contacts at the TAD scale from cell to cell, with domains appearing as tendencies that become more visible when averaged over a population of cells (Nagano et al., 2013; Stevens et al., 2017; Szabo et al., 2019). This raises the question of the physical reality of TADs rather than just the result of statistical averages. Also, of their functional relevance if such heterogeneity between cells is confirmed. Further single-cell studies are required to elucidate the conservation of TADs at this scale.

There is still much we do not know about nuclear compartments formation and their role in genomic functions. Nonetheless, data gathered in the last few years indicate that these compartments correspond to a functional subdivision of the genome. However, the need of partition the genome into domains to ensure proper gene regulation is still a mystery that hopefully the combination and continuous improvement of new technologies (single-cell omics, super resolution microscopy, modeling of the chromatin fiber...) will be able to elucidate.

In conclusion, eukaryotic transcription is a complex multistep process subject to numerous modes of regulation. This regulation involves the controlled interaction between regulatory elements in the DNA sequence and proteins with a variety of biochemical activities. Interestingly, regulatory elements in the DNA sequence seem to correlate with gene function, however the molecular mechanisms underlying this apparent specificity are yet to be elucidated. In the nucleus, DNA is organized as a chromatin fiber. Therefore, the structure, composition and folding of chromatin represent a major point of control of gene expression. These chromatin features can be dynamically regulated by the concerted action of epigenomic effectors to achieve specific transcriptional outcomes. Indeed, the establishment of specific and highly regulated transcriptional programs determines cell identity and function. During my PhD I particularly focused in the epigenetic control of the *Drosophila melanogaster* ovarian transcriptome. In the next part I will thus describe this particular system.

III. Part III Epigenomic regulation during *D. melanogaster* oogenesis

Fertilization involves the union of two highly different gametes followed by the formation of a totipotent embryo. This implicates a series of complex nuclear and cellular events (Loppin et al., 2015). Remarkably, this occurs in the absence of zygotic transcription which means that these processes are almost entirely controlled by factors already present in the mature oocyte (Avilés-Pagán and Orr-Weaver, 2018; Stitzel and Seydoux, 2007). This developmental strategy is used by a nearly every animal and requires thus the accumulation and deposition of maternal stores during oogenesis. Maternal stockpiles include mRNAs, proteins, and nutrients which permit early embryogenesis to occur in the absence of zygotic transcription.

Preparation of the fertilized oocyte includes profound changes such as completion of female meiosis, formation of male and female pronuclei and the selective translation of

maternal RNAs. The ensemble of events required for the transition from mature oocyte to developing embryo are collectively known as egg activation. Great efforts have been deployed to identify the factors needed for the oocyte-to-zygote transition. (Avilés-Pagán et al., 2020; Avilés-Pagán and Orr-Weaver, 2018; Horner and Wolfner, 2008; Stitzel and Seydoux, 2007). For example, maternal mRNAs synthesized during oogenesis and loaded into the oocyte need to be stably maintained for prolonged periods yet not translated until the right time. Dynamic changes of the poly(A) tail have been involved in the translational control of maternal mRNAs (Eichhorn et al., 2016; Lim et al., 2016; Tadros et al., 2007). These changes depend at least on the kinase Png, the RNA binding protein Smaug and the noncanonical poly(A) polymerase Wispy. However, translational control has not been fully elucidated indicating that more factors are involved. Overall, the oocyte-to-zygote transition requires complex regulation to link developmental signals with profound changes in mRNA translation, cell cycle control, and metabolism. These are complex processes that depend therefore on a wide range of factors with different biochemical properties.

The control of the oocyte-to-embryo transition in *Drosophila* parallels that of other animals, but *Drosophila* offers experimental advantages as a model. In addition to the numerous genetic tools available, the oocyte is the single largest cell and a single ovary contains every stage of oocyte maturation, from stem cell to mature oocyte, and each stage is morphologically distinct (Bastock and St Johnston, 2008; McLaughlin and Bratu, 2015).

1. Main steps of *Drosophila* oogenesis

Oogenesis denominates the process of female gamete formation. In insects, this process occurs in the ovarioles of an ovary. *Drosophila melanogaster* females have two ovaries, each containing 16 to 20 autonomous ovarioles, each composed of their own stem cell populations and egg chambers at varying developmental stages (Fig17-A). Each egg chamber gives rise to a single egg. The process of oogenesis has been arbitrarily divided into 14 stages based on morphological criteria (Fig17-B) (King, RC, 1970; Spradling, 1993).

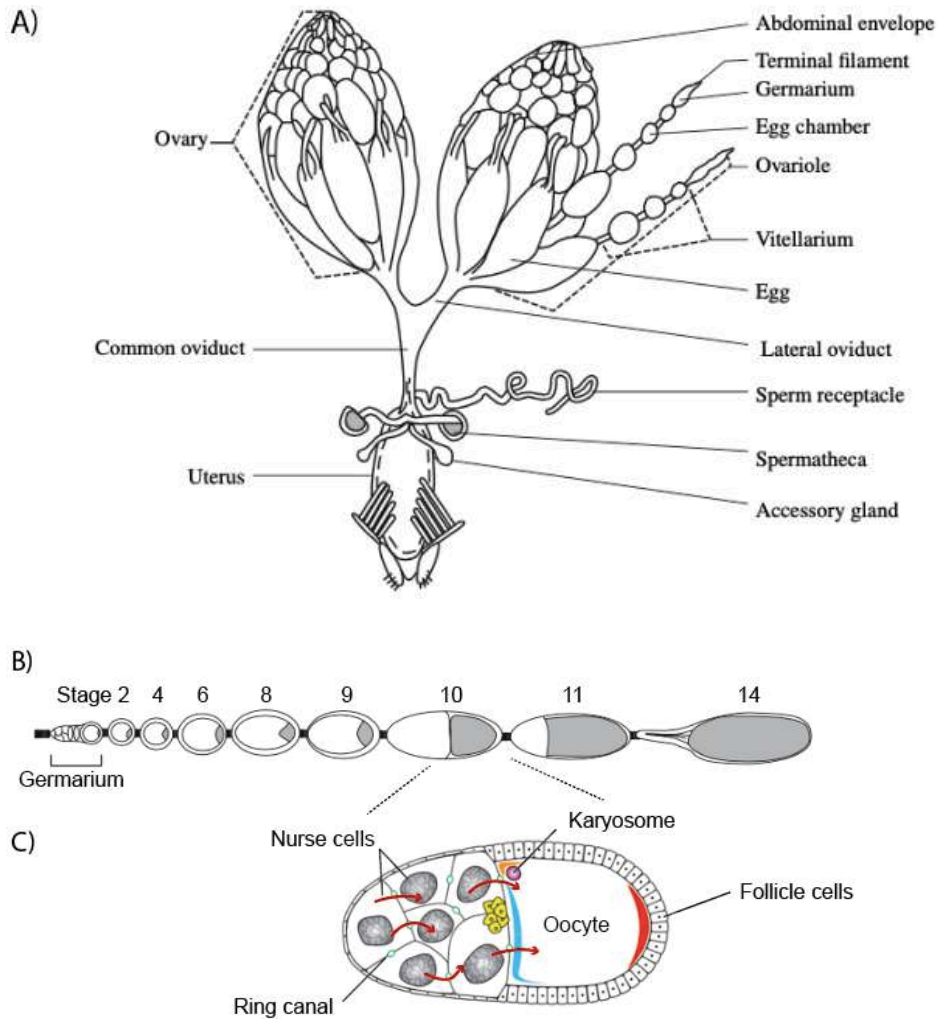


Fig 17-Oogenesis process in *D. melanogaster*

(A) Organization of the reproductive system in a mature *D. melanogaster* female. The two ovaries are composed of ovarioles each composed of the germarium and egg chambers at varying developmental stages. (B) Scheme of the ovariole with all stages of oogenesis. Egg chambers are formed at the anterior tip, in the germarium and bud at stage 2. The egg chamber grows until stage 10 where the nurse cells empty their content into the oocyte (dumping). Finally at stage 14, the mature egg is enveloped by the vitelline membrane and the chorion and is ready for fertilization. The oocyte is in gray. (C) Detail of an egg chamber at stage 10 where the process of dumping occurs indicated by red arrows (McLaughlin and Bratu, 2015; Ogienko et al., 2007).

Oogenesis progresses from the anterior to the posterior of the ovariole (McLaughlin and Bratu, 2015; Ogienko et al., 2007). The latter can be divided into three regions: a terminal filament, a germarium and a vitellarium. The germarium contains somatic and

germline stem cells (GSC). From there, egg chambers bud off and mature as they pass down the ovariole. Processes like vitellogenesis and choriogenesis are completed in the vitellarium. These constitute the synthesis of protective layers around the egg: the vitelline membrane and the chorion to jointly form the eggshell (Pascucci et al., 1996). Finally, mature eggs reach the posterior part of the ovariole competent for fertilization.

At the anterior tip of the germarium GSCs divide asymmetrically to produce a cystoblast and a new stem cell. Next, the cystoblast undergoes a total of four divisions, producing a 16-cell cyst. Each mitotic division is accompanied by incomplete cytokinesis, forming inter-cellular cytoplasmic bridges known as ring canals. One out of the 16 cells of the cyst becomes the oocyte and the other 15 cells differentiate into nurse cells (Fig18). The oocyte and nurse cells enveloped by somatic follicle cells constitute an egg chamber. Before the egg chamber leaves the germarium, DNA in the oocyte condenses into a compact structure called karyosome. Additionally, meiosis is arrested in prophase I and it will not be continued until late oogenesis, at which point meiosis progresses to metaphase I and is arrested again until egg activation.

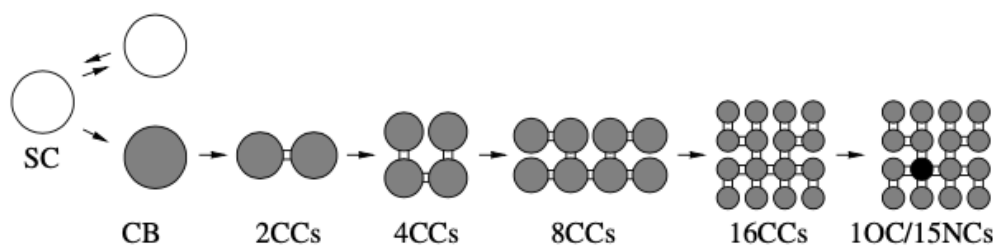


Fig18-Scheme of germline cell division from a stem cell to a 16-cell cyst.

Abbreviations: SC= stem cell, CB= cystoblast; CC= cystocyte; OC= oocyte; NC= nurse cell (Ogienko et al., 2007).

The egg chamber buds from the germarium to the vitellarium, marking stage 1 of oogenesis. Developing egg chambers move along the ovariole to the posterior end. The oocyte and nurse cells grow while follicle cells undergo active mitotic divisions. Both nurse cells and follicle cells undergo several rounds of endocycles to synthesize nutrients, mRNAs and proteins essential for oocyte growth and development (see next section). By the end of stage 10, nurse cells empty their content into the oocyte (Fig17-B,C), only nuclei, some actin filaments and a minor amount of cytoplasm remains in the nurse cell. This process is known as “dumping”. It ends with the formation of mature egg, ready for fertilization. Simultaneously, follicle cells

secrete the chorion and the vitelline membrane to protect the mature oocyte. Finally, both nurse cells and follicle cells experience apoptosis at the end of egg chamber development.

2. Transcriptional programs of nurse and follicle cells are essential for oocyte nutrition and maturation

In *Drosophila*, germ line transcriptional activity is ensured by the nurse cells, directly connected to the oocyte by cytoplasmic junctions to provide it with large amounts of mRNAs, proteins and other cellular material (Bastock and St Johnston, 2008; Spradling, 1993). This strategy ensures gene expression while maintaining chromosome condensation in the oocyte (Davidson, E, 1986). During early development of egg chamber, certain mRNAs, proteins, and organelles are preferentially transported from nurse cells into the oocyte, in a process known as selective transport. This is a slow and highly selective transport and is essential for oocyte determination and polarity. This process is dependent on the microtubule network and leads to the asymmetric distribution of proteins and mRNAs (Ogienko et al., 2007). After stage 10, rapid transport starts and nonselective dumping of the nurse cell content takes place. This substantial supply is possible thanks to the massive RNA synthesis occurring in the nurse cells. In these cells DNA undergoes 10 to 12 rounds of endoreplication cycles (Dej and Spradling, 1999) and as a result, polyploidy reaches 2048C in cells adjacent to the oocyte. Interestingly, genomic intervals are differentially replicated during the endocycle S phase such that some regions are under-replicated, while others can be amplified (Royzman and Orr-Weaver, 1998). For example, during polyploidization of the nurse cell nuclei, satellite DNA is differentially lost and ribosomal DNA increases in content in their genome (Hammond and Laird, 1985). This bias certainly reflects the needs of the mature oocyte, where great amounts of translation of maternal mRNAs will be needed.

Follicle cells surround the developing oocyte and are also essential for oocyte maturation. Until stage 6 of oogenesis, these cells proliferate by mitosis giving rise to a maximum number of ~1000 cells surrounding the egg chamber (Deng et al., 2001). At stage 6, follicle cells stop a normal mitotic cycle and enter several rounds of endocycles (Nordman and Orr-Weaver, 2012). At stage 10, follicle cells exit the endocycle and begin gene amplification cycles. this event is known as E/A switch. During this amplification, four specific genomic loci, encoding genes involved in chorion and vitelline membrane synthesis,

are amplified from 4 to 80-fold. The four amplified loci are known as a *Drosophila* amplicon in the follicle cells (Claycomb and Orrweaver, 2005). Amplification of two clusters of chorion protein genes allows the production of high-levels of chorion-related proteins, required for egg maturation. Egg chambers are composed of morphologically, genetically and functionally different cells and transcriptional regulation of each of them is important for correct oogenesis.

In conclusion, DNA endoreplication and gene amplification constitute an effective strategy to supply the high demand that represents nourishment and maturation of the oocyte.

3. Mechanisms of epigenomic regulation in the female germline

A. Silencing of bulk chromatin by the piRNA system

Additional to the silencing of bulk chromatin (described in section II-4) mediated by the H3K9me3/HP1 pathway, the metazoan germline developed the piRNA (PIWI-interacting RNA) system. This is a small RNA silencing system that acts in animal gonads and protects the genome against the deleterious influence of transposable elements (TEs). Indeed, TEs are DNA pieces that can move within the genome potentially compromising faithful transmission of the genetic information in the germline. Loss of piRNAs is associated with significant over expression of retrotransposons (Aravin et al., 2001; Malone et al., 2009), suggesting thus that the mechanism silencing them is piRNA-dependent.

Two types of silencing can be achieved with this pathway, transcriptional silencing mediated by the nuclear Piwi protein in *Drosophila*, or post-transcriptional silencing mediated by cytoplasmic proteins such as Aubergine and Argonaute3. One of the prevalent models for Piwi-mediated transcriptional silencing proposes that the Piwi-piRNA complex binds to the nascent transposon transcript and recruits several proteins ultimately tethering the H3K9 histone methyltransferase Eggless/dSETDB1. This results in the establishment of a repressive chromatin state, silencing thus transposon expression (Fig19). Alternatively, it has also been proposed that Piwi can directly recruit HP1a to initiate the heterochromatinization process. These models are not mutually exclusive and share two features, the binding of a Piwi-piRNA complex at the target site to recruit chromatin factors and transcriptional

repression mediated by chromatin structure modification. Post-transcriptional silence takes place in the cytoplasm where a Piwi-piRNA complex binds the complementary transposon RNA and the latter is cleaved by the Piwi protein. This also promotes piRNA biogenesis and amplifies the system (Ozata et al., 2019; Wang and Lin, 2021), maintaining thereby a solid repression.

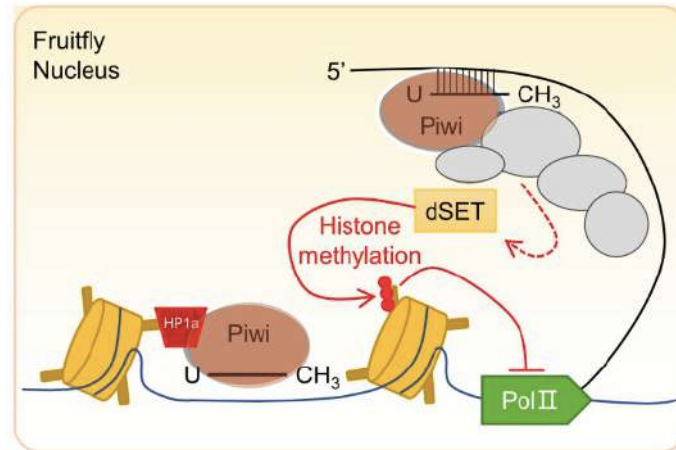


Fig19-Model for Piwi-mediated transcriptional silencing of transposons in *D. melanogaster* nucleus.

Piwi-piRNA complex binds to a nascent transposon transcript ultimately recruiting to the vicinity of the target chromatin region the H3K9 methyltransferase Eggless/dSetDB1 which establishes a repressive chromatin state to suppress transposon expression. Alternative model proposes that Piwi can directly recruit HP1a (Wang and Lin, 2021).

piRNA biogenesis can be divided into two stages. First, long RNA precursors are transcribed in the nucleus and exported into the cytoplasm. In the cytoplasm, piRNA precursors are further processed to generate mature piRNAs that get loaded into Piwi proteins. In flies, piRNA precursors come from heterochromatic loci. In the germline, majority of piRNAs are produced in “dual-strand” clusters. They produce sense and antisense piRNAs regardless of transposon orientation and their transcription requires both the repressive chromatin mark H3K9me3 and the transcriptional silencing protein Piwi (Akkouche et al., 2017; Rangan et al., 2011). Dual-strand clusters give rise to piRNA precursor RNAs via non-canonical transcription facilitated by the Rhino protein, a germline-specific HP1 variant (Klattenhoff et al., 2009). Rhino binds H3K9me3 and tethers a specific transcription factor initiator, Moonshiner, on both strands of DNA. In turn, Moonshiner forms an alternative pre-initiation complex, allowing RNA Pol II to initiate transcription from many

sites on both DNA strands. Rhino is thus able to bypass the need of specific DNA regulatory sequences such as promoters (Andersen et al., 2017). To conclude, the piRNA-Piwi silencing system provides a potent chromatin-based defense against potential deleterious effects from TEs, thereby guarding genome integrity in the future gametes.

B. The importance of fine transcriptional regulation during oogenesis

The products contained in the oocyte ensure a successful oocyte-to-zygote transition and early embryogenesis. However, most of our knowledge of this process is centered on the post-transcriptional regulation of gene expression (Kronja et al., 2014). Polyploid nurse cells are capable of mass-producing the necessary factors for the onset of development. Yet, transcriptional regulation during oogenesis has been largely overlooked. Focus on the fundamental importance of proper regulation of transcriptional programs during oogenesis has just recently emerged.

a. Targeted silencing is essential for female germline determination

In part II of this introduction, I described a series of mechanisms to ensure fine gene regulation based on chromatin-related mechanisms. Nonetheless, knowledge on how these mechanisms are used in the female germline is very limited. As said before, a unique germ stem cell (GSC) gives rise to different lineages of the germline, the nurse cells and the oocyte. This unique cell must thus undergo major changes in gene expression and chromatin organization.

New insights underlying epigenomic changes during oogenesis came from a recent study, from the Spradling lab, where they used *Drosophila* oogenesis to study Polycomb repression. Implication of Polycomb proteins in gene regulation during *Drosophila* oogenesis had already been observed. A mutation in E(z), the catalytic subunit of the PRC2 complex, impairs gene silencing of *Cyclin E* and *dacapo*, a cyclin-dependent kinase inhibitor. This results in the oocyte-to-be entering a nurse cell-like endoreplicative program and failing to be determined as the oocyte (Iovino et al., 2013). In the new study by DeLuca *et al.*, authors unveiled a role for Polycomb silencing in the transition from GSC to nurse cell (DeLuca et

al., 2020). They first noticed that the patterns of H3K27me3 between these two cell types were different. Indeed, nurse cell progenitors lack silencing and H3K27me3 shows a broad distribution, referred to by the authors as non-canonical H3K27me3 pattern. As nurse cells differentiate, the H3K27me3 pattern becomes focused on common PcG domains. The authors proposed a model where association of PRC2 with the PcG protein Pcl, prevents it from sampling at many sites, resulting in infrequent and stochastic silencing. As differentiation occurs, Pcl levels drop and core-PRC2 is freed to sample and silence more sites. This study shows how Polycomb-mediated reshaping of chromatin is essential for female germline development in *Drosophila*.

Interestingly, it was also observed that female germ cell fate is maintained by an epigenetic regulatory pathway depending on H3K9me3/HP1a silencing of key spermatogenesis genes (Smolko et al., 2018). Female germ line specific knockdown of the H3K9 methyltransferase *eggless/dSETDB1* and its partners *HP1a* and *windei* results in ectopic expression of testis-specific genes. Mapping of H3K9me3 revealed the accumulation of this mark on 21 of the ectopically expressed genes. Remarkably, and contrary to the general vision of H3K9me3 as broad heterochromatin blocks, the mark was highly localized and did not spread into neighboring loci. A striking example the authors described was the *phf7* gene, where in ovaries, H3K9me3 is restricted to the region surrounding the silent testis-specific TSS. The mechanisms through which H3K9me3 is targeted and restricted in such a specific manner are not fully elucidated. Nevertheless, this study provides a non-common usage of the H3K9me3/HP1a silencing pathway in the female germline.

Overall, these data emphasize the importance of fine chromatin-based regulation of transcription in the female germline for correct gamete production.

b. Histone modifiers play an important role in transcription activation during oogenesis

Different studies found an important role for histone modifiers during oogenesis, pointing out the importance of an epigenetic mode of regulation during this process. The importance of the H3K4me3 histone demethylase *Lid/dKDM5* during oogenesis was observed by two different groups in addition to ours (see Results). Zhaunova and colleagues observed that absence of *Lid/dKDM5* in the female germline leads to a series of defects in meiotic chromatin organization in oocytes, including instability of the recombination

machinery (Zhaunova et al., 2016). Oocytes remain arrested in prophase I of meiosis for a significant amount of time. They are transcriptionally silent throughout this arrest but they reactivate transcription prior to the resumption of meiosis (Mahowald and Tiefert, 1970). Navarro-Costa and colleagues characterized the epigenome of the prophase I-arrested oocyte and found that it is highly dynamic and contains both euchromatic and heterochromatic marks that vary during oocyte quiescence and reactivation (Navarro-Costa et al., 2016). Female germ-line depletion of Lid/dKDM5, led to a significant increase in the levels of H3K4me3 and to precocious transcriptional reactivation of the quiescent prophase I-arrested oocytes. On the contrary, no changes in H3K27me3 levels were detected. Importantly, the demethylase activity of Lid/dKDM5 was required for correct transcriptional reactivation of the oocyte and meiotic progression, suggesting a role of this histone demethylase in the regulation of the oocyte epigenome. These data support the hypothesis that correct reactivation of the dormant primary oocyte is epigenetically regulated. They also observed that loss of Lid/dKDM5 severely affects meiotic completion and accordingly, most fertilized eggs fail to initiate mitotic divisions.

Furthermore, a recent study identified the H3K4 methyltransferase Trr as essential for the oocyte-to-zygote transition. Interestingly, in the absence of Trr, oocytes develop normally but fail to complete maternal meiosis and to form the paternal pronucleus. The proposed model is that during oocyte development, Trr promotes the expression of a subset of genes that are not required during oogenesis but their presence in the mature oocyte is critical for proper zygote formation (Fig20). One of the identified genes under direct control of Trr and necessary for the oocyte-to-zygote transition coded for the IDGF4 glycoprotein (Prudêncio et al., 2018) but its specific role at zygote formation is yet to be determined. Overall, these data emphasize the important role of chromatin regulation during female gametogenesis.

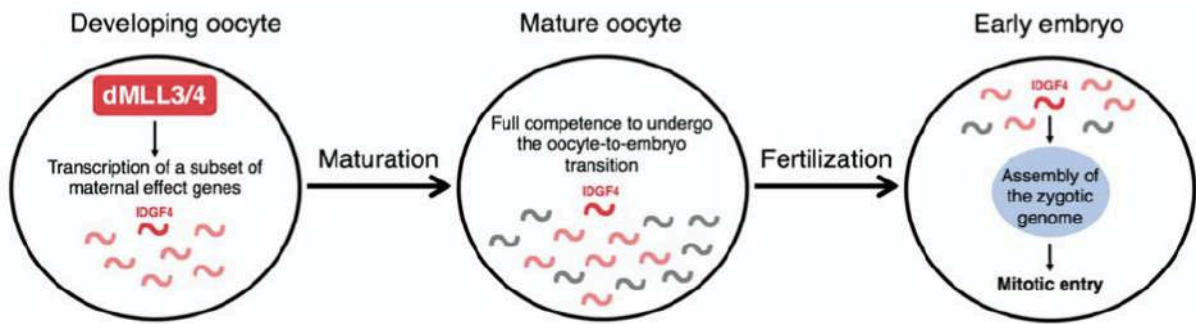


Fig20-Proposed model for Trr-regulated acquisition of embryo fate at fertilization.

During oocyte maturation, Trr promotes the expression of a subset of genes, which products are not required for normal oogenesis but will be indispensable for zygotic genome assembly at fertilization (Prudêncio et al., 2018).

Emergent evidence suggests that transcriptional regulation during oogenesis is mediated, at least in part, by epigenetic mechanisms that define specific gene expression modules. This allows the establishment of the molecular basis of the crucial and complex oocyte-to-zygote transition.

IV. Part IV The peculiar case of *dhd* regulation

My PhD project was based on the discovery of a specific functional connection between a series of epigenomic effectors and the highly regulated terminal effector of zygote formation, the maternal thioredoxin Deadhead (Dhd). The different effectors involved are: the histone demethylase Lid/dKDM5, the histone deacetylase complex scaffold Sin3A, the Brahma chromatin remodeler sub-unit Snr1 and the insulator component Mod(mdg4). The case of *dhd* represented thus an opportunity to study the molecular mechanisms underneath transcriptional regulation by chromatin factors. I will first introduce the role of Dhd at fertilization and the specific features of this protein and of the gene that encodes it. This will be followed by a state-of-the art on the epigenomic effectors that regulate this singular gene.

1. The essential role of Deadhead at fertilization

The *dhd* gene encodes an egg specific thioredoxin required for female fertility and development of viable embryos (Emelyanov and Fyodorov, 2016; Salz et al., 1994; Tirmarche et al., 2016). Thioredoxins are small, highly conserved redox proteins that catalyze the reduction of disulfide bonds on target proteins (Arnér and Holmgren, 2000). At fertilization, Dhd is critically required for sperm chromatin remodeling. DNA in the sperm nucleus is highly compacted due to the almost total replacement of histones by sperm nuclear basic proteins (SNBPs), including protamines (Miller et al., 2010). This level of compaction is incompatible with basic nuclear activities such as transcription, replication or repair. One of the first key events at fertilization is thus remodeling of this nucleus so that paternal chromosomes can be integrated into the zygote (Fig21-A) (Loppin et al., 2015). This process depends on Dhd reducing disulfide bonds between protamines thus allowing sperm nuclear decondensation. Indeed, in eggs laid by *dhd* null mutant females, the sperm nucleus retains protamines and remains needle-shaped, reminiscent of its ultra-compacted DNA (Fig21-B,C) (Emelyanov and Fyodorov, 2016; Tirmarche et al., 2016). A catalytic mutant is unable to rescue the *dhd* mutant phenotype showing that this process depends on Dhd catalytic reducing activity (Emelyanov and Fyodorov, 2016; Tirmarche et al., 2016). Dhd has also been involved in the redox balance at the oocyte-to-embryo transition (Petrova et al., 2018). The study of redox state changes during this process revealed that early embryos have a more oxidized state than mature oocytes. It was observed that *dhd* mutant oocytes are prematurely oxidized and exhibit meiotic delay. A highly specific list of Dhd substrates was established and a major fraction of Dhd's interactors are ribosomes or ribosome-associated. These data show that Dhd has crucial roles during the oocyte-to-zygote transition.

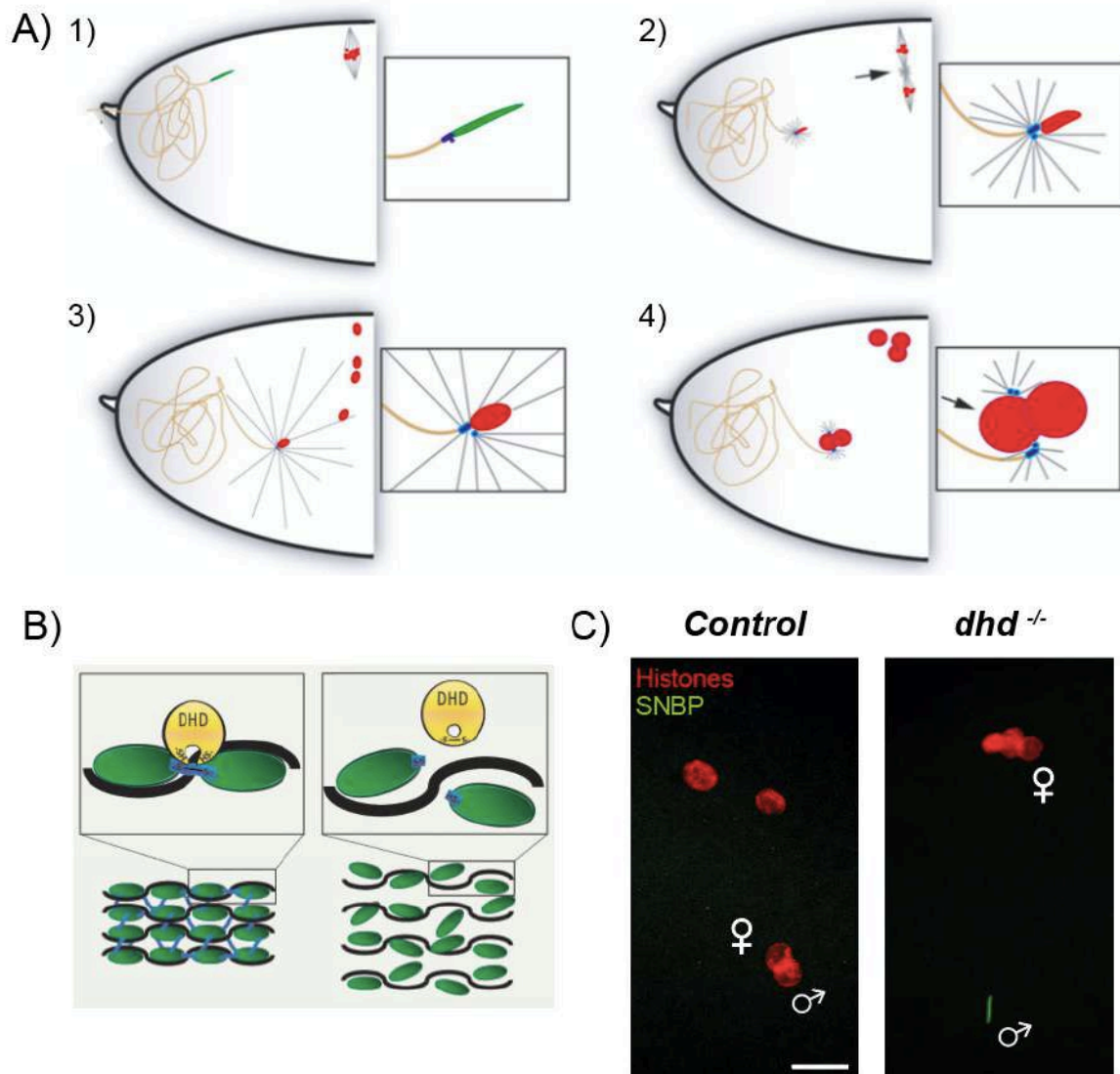


Fig 21: The essential role of Deadhead at fertilization.

(A) First steps of zygote formation in *Drosophila melanogaster*. (1) Sperm enters the egg. At fertilization, maternal chromosomes are arrested in metaphase I of meiosis I. Sperm nucleus is ultra-compact and packed with sperm nuclear basic proteins (SNBPs) (green). (2) Maternal chromosomes progressed to metaphase of meiosis II. Remodeling of the sperm chromatin took place: SNBPs were replaced by maternally provided histones (red) and the nucleus is decondensed. (3) Pronuclei migration. The inner most product of female meiosis migrates towards its male counterpart. (4) Apposition of female and male pronuclei. The first zygotic replication begins (Loppin et al., 2015). (B) Schematic representation of the role of Deadhead at fertilization. The thioredoxin DHD reduces disulfide bonds between SNBPs (left) allowing their eviction and chromatin decondensation (right) (Horard and Loppin, 2017). (C) Confocal images of eggs at the apposition stage, laid by control (left) and *dhd* null (right) mutant females. In control eggs both male and female pronuclei are round and

contain only histones. In *dhd* mutant eggs, the sperm nucleus retains SNBPs and is needle-shaped, reminiscent of compacted DNA. Scale bar 10 μ m.

Dhd is exclusively found in ovaries, its levels increase during oocyte maturation and start decreasing at egg activation (Kronja et al., 2014; Tirmarche et al., 2016). At fertilization, Dhd is abundant and homogeneously distributed throughout the egg cytoplasm. Strikingly, it becomes rapidly undetectable after completion of the first zygotic cycle (Fig22) (Tirmarche et al., 2016). This indicates that likely Dhd plays no other role after zygote formation and it is possible that its continued presence is detrimental to embryogenesis.

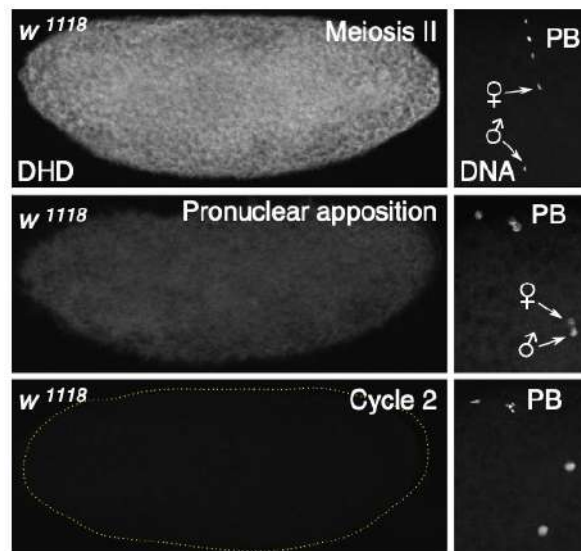


Fig22-Deadhead is rapidly degraded after fertilization

Confocal images of control eggs stained for DHD (left) and DNA (right). DHD is abundant and homogeneously distributed at fertilization but is rapidly undetectable (Tirmarche et al., 2016).

The role of Dhd is thus essential for the oocyte-to-zygote transition and the onset of embryonic development. Interestingly, *dhd* stands out for a series of unusual features at the protein level as well as at its genomic locus.

2. Dhd, a thioredoxin not like the others

As mentioned before, Dhd is a thioredoxin. Thioredoxins constitute a family of small thiol proteins that are present in all organisms studied so far, and are characterized by the sequence of their conserved active site (WCGPC). In addition to their role maintaining redox homeostasis in the cell, they have been implicated in DNA synthesis, regulation of transcription factors or programmed cell death (Arnér and Holmgren, 2000), indicating a wide variety of roles for these proteins.

Drosophila melanogaster possesses three true thioredoxins, the testis-specific TrxT, the ovary-specific Dhd and the ubiquitous Trx2. Alignment of the protein sequences revealed that Dhd is less related to Trx2 than TrxT (J. Svensson and Larsson, 2007). Accordingly, *in vitro* Trx2 is not able to substitute Dhd's role in reducing protamine disulfide bonds (Emelyanov and Fyodorov, 2016; Tirmarche et al., 2016). Interestingly, a recent characterization of the structure of the *Drosophila melanogaster* thioredoxins revealed that Dhd has an atypical structure for a thioredoxin (Freier et al., 2021). In contrast to the negatively charged surfaces commonly found in most thioredoxins, it was observed that Dhd has positively charged patches on its surface. Dhd is in charge of reducing protamine disulfide bonds in sperm chromatin and was found associated with ribosomes (Petrova et al., 2018; Tirmarche et al., 2016). The unusual positive patches on its surface might thus help Dhd in selecting proteins and DNA/RNA partners by complementarity with their negatively charged backbone. This distinctive charge distribution helps to define the initial encounter with DNA/RNA complexes that will lead to final specific interactions with cofactors to promote chromatin remodeling (Freier et al., 2021). Specific and unusual features of the Dhd thioredoxin are thus important for its role at fertilization.

3. The intricate *dhd* locus

Dhd protein is present exclusively in ovaries (Salz et al., 1994; Tirmarche et al., 2016). Accordingly, RNA *in situ* hybridization experiments show that the gene is transcribed in nurse cells and the transcript is highly present at stage 10 and then deposited in the oocyte (Salz et al., 1994; Svensson et al., 2003; Tirmarche et al., 2016).

Data from Flybase (<http://flybase.org>), and our own RNA-seq analyses, indicate that *dhd* is among the most highly expressed genes in ovaries (Fig23-A). This is all the more

surprising due to the genomic context of this gene. *dhd* is a small, intronless gene located in the middle of a cluster of fifteen densely packed genes spanning about 40 kb of genomic DNA (Fig23-B). Commonly, *Drosophila* genes with tissue-, temporally-biased expression patterns have been shown to be concentrated in neighborhoods of contiguous genes (Parisi et al., 2004; Spellman and Rubin, 2002). However, the *dhd* gene lies within a 1.4 kb region that is immediately flanked by two genes with testis-specific expression, the thioredoxin *Trx-T* and the *CG4198* gene of unknown function. *TrxT* and *dhd* are arranged as a gene pair, transcribed in opposite directions separated by 282 bp. These two genes have thus closely spaced promoters and are differentially regulated by a short common control region. A transgene including the *TrxT-dhd* region and an additional 2 kb and 1 kb downstream of *TrxT* and *dhd* respectively ($P[w^+ snf^+ TrxT^+ dhd^+]$) (Fig23-B), inserted at the X chromosome or at an autosome was able to recreate the expression pattern of *TrxT* and *dhd*. This shows that the endogenous X chromosome location of *TrxT-dhd* is not essential for tissue-specific transcription (Svensson et al., 2007). Furthermore, an even smaller transgene spanning only *Trx-T*, *dhd* and part of *CG4198* ($pW8-attB-dhd^{WT}$) (Fig23-B), fully rescues *dhd* maternal effect embryonic lethal phenotype (Tirmarche et al., 2016), showing that this 4.3 kb region is capable of recapitulating *dhd* expression. The necessary regulatory signals for *dhd* activation are thus contained within this restricted region.

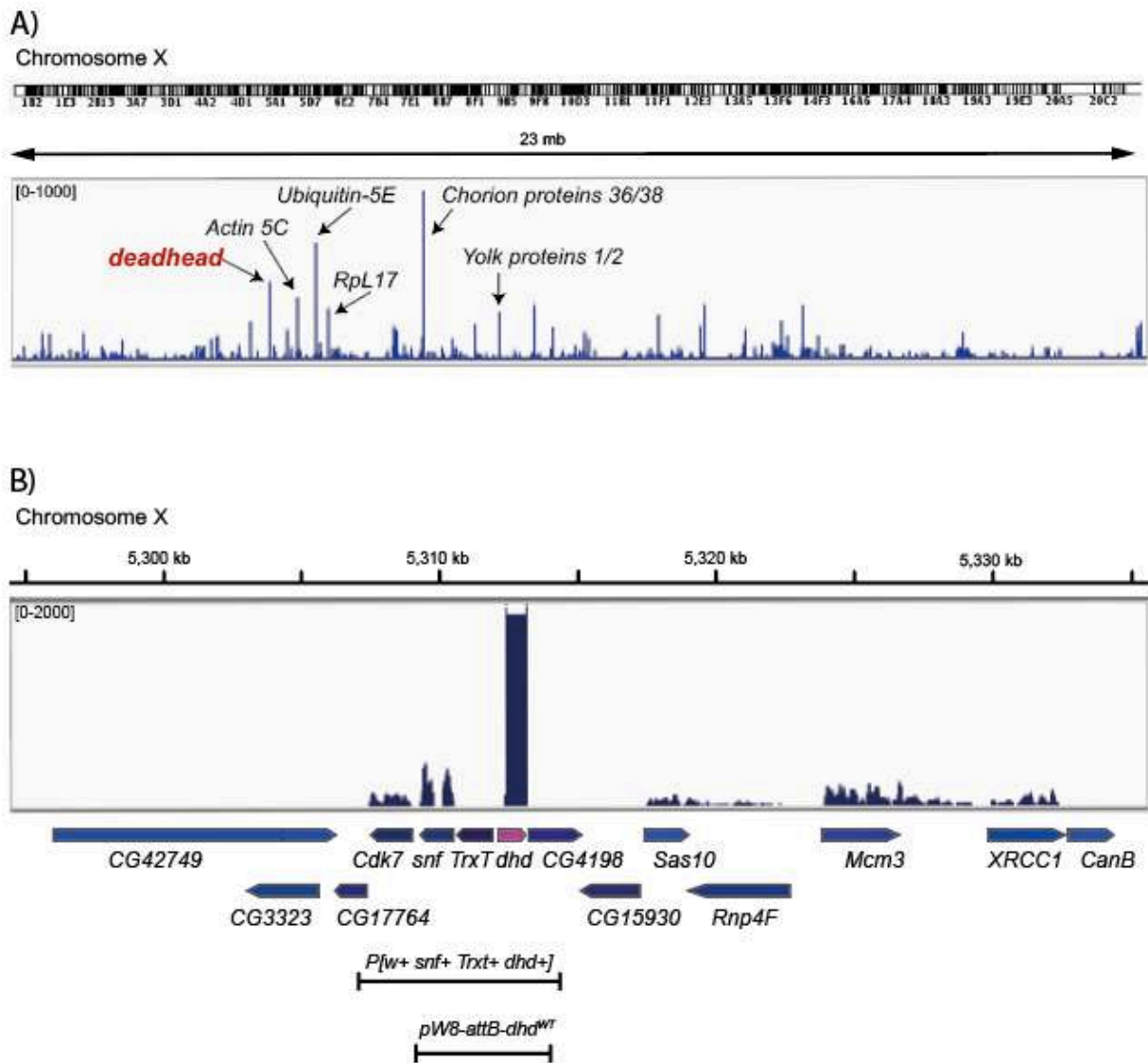


Fig 23-*deadhead* is massively expressed in ovaries

dhd is amongst the highest expressed genes in ovaries. (A) Genome browser view of control ovarian RNA-seq signal of the X-chromosome showing the high level of transcription of *dhd*. Examples of other highly expressed genes are indicated by arrows. (B) Genome browser view of control ovarian RNA-seq signal at the *dhd* region. Signal coming from the *dhd* locus was truncated for readability. The genomic region covered by two different transgenes are shown: (i) the $P[w+ snf+ Trxt+ dhd+]$ transgene, capable of recreating *dhd* expression pattern (Svensson et al., 2003) and (ii) the pWB-attB-*dhd*^{WT} transgene, capable of restoring *dhd* expression and rescue female sterility [Tirmarche et al., 2016].

Svensson and colleagues did a thorough analysis of the organization of the *TrxT* and *dhd* locus across several *Drosophilid* species (Svensson et al., 2007). The intriguing gene organization and regulation of *TrxT* and *dhd* is remarkably well conserved. For both of these

genes, the lack of introns in the coding region is conserved across species. Additionally, predictions of TSSs for *dhd* and *TrxT* in different *Drosophilid* species revealed that the closer to the transcription start sites, the higher the conservation level of the sequence. Conserved motif analysis within 150 bp upstream of the predicted TSSs revealed 9 different motifs however, none of them had obvious similarities with known common target sequences for transcription factors. The motifs identified were also compared to the 5' flanking regions of ovary- and testis-specific genes. They were detected in less than 1% of the genes in both ovary and testes sets indicating that they are neither ovary- nor testis-specific. Finally, the search for gene pairs including a testis-specific gene and an ovary- specific gene transcribed in opposite directions and separated by no more than 1000 bp, showed only 8 additional gene pairs in this configuration. None of the nine motifs previously identified were found in the 5' flanking regions of any of the other gene pairs. Importantly, none had such extreme ovary- or testes-specific expression as *dhd* and *TrxT* when expressed as a ratio, showing thus the rareness of this organization.

In conclusion, the genomic organization and regulation of *TrxT* and *dhd* is unique and well conserved in evolution. Strikingly, the *dhd* locus accumulates a series of counterintuitive features when considering its high expression.

4. State-of-the art of *dhd* regulators in transcription regulation

My PhD project focused on the epigenetic regulation of the *dhd* singular gene which led me to focus on the role of particular epigenomic effectors so in this section I will more precisely lay out the current literature on their roles in transcription regulation.

A. The histone demethylase Lid/dKDM5

Lid/dKDM5 is the sole member of the KDM5 histone demethylases in *Drosophila*. It specifically targets H3K4me3 *in vivo* (Secombe et al., 2007). Lid/dKDM5 has been found enriched at promoter regions in embryonic cells, whole adults and wing discs (Gajan et al., 2016; Liu and Secombe, 2015; Lloret-Llinares et al., 2012; Zamurrad et al., 2018), hinting about a role in transcriptional regulation.

Since H3K4me3 has long been associated with transcriptional activation and Lid/dKDM5 removes this mark, this enzyme could have been expected to have a role in global repression rather than activation. Nonetheless, RNA-seq analyses in mutants and Lid/dKDM5-depleted cultured cells, show equivalent up-regulated and down-regulated genes (Drelon et al., 2018; Gajan et al., 2016; Liu and Secombe, 2015). A global increase in H3K4me3 is consistently observed upon Lid/dKDM5 depletion or mutation, however the impact on transcription has been found mild (Drelon et al., 2018; Liu et al., 2014; Lloret-Llinares et al., 2012; Zamurrad et al., 2018). This reinforces the enigmatic role of H3K4me3 in transcriptional activation discussed in section II-7-B.

The impact of H3K4me3 on transcription is not clear however Lid/dKDM5 possesses additional domains with other functions. Indeed, the family of KDM5 proteins contain the following domains: JmjN domain (unknown function), JmjC domain (H3K4me3 demethylase activity), ARID (implicated in DNA binding), C5HC2 zinc finger and two or three PHD domains (protein-protein interaction). Interestingly, null *lid* mutants are lethal but can be rescued with a demethylase-dead transgene (Drelon et al., 2018; Li et al., 2010; Liu and Secombe, 2015), showing thus that Lid/dKDM5 has catalytic-independent crucial roles. Indeed, in whole adults, Lid/dKDM5 proved to be a critical regulator of genes associated with mitochondrial structure and function. This regulation was independent from its demethylase JmjC domain however the C-terminal PHD motif, capable of binding di- and trimethylated H3K4 was essential (Liu and Secombe, 2015). Nonetheless, in a study where only fly brains were analyzed, it was observed that Lid/dKDM5 can directly activate or repress transcription in a demethylase-dependent manner (Zamurrad et al., 2018). These data indicates thus that Lid/dKDM5 regulates genes involved in different cellular processes and that this regulation can depend or not on its demethylase activity.

An extra layer of complexity comes from the fact that Lid/dKDM5 can act with different partners, probably due to its different protein domains. These interactions can lead to different outcomes on gene transcription. In *Drosophila* embryos, Lid/dKDM5 was found in a complex with the histone deacetylase Rpd3, the scaffold Sin3A, the histone chaperones (Asf1 or Nap-1) and other proteins. The resulting complexes acted as repressors of Notch target genes (Moshkin et al., 2009). However, in embryonic fly cells, Lid/dKDM5 interaction with the transcription factor Foxo and the histone deacetylase dHDAC4 leads to the activation of a subset of Foxo target genes (Liu et al., 2014), showing the versatility of Lid/dKDM5 on transcriptional outcomes. It has been also observed that Lid/dKDM5

genetically interacts with the H3K4me1/2 demethylase dLsd1 (Di Stefano et al., 2011). Lid/dKDM5 opposes the functions of dLsd1 and the H3K9 methyltransferase Su(var)3–9 in promoting heterochromatin spreading at heterochromatin–euchromatin boundaries. However, Lid/dKDM5 cooperates with dLsd1 in regulating certain Notch target genes in euchromatic contexts, illustrating that the activity of histone demethylases is context-dependent. Lid/dKDM5 has thus a complex and intricate role in gene transcription regulation that depends on several factors such as cell-type and interactors.

B. The Sin3-HDAC complex

The Sin3 protein is highly conserved from yeast to mammals and has different isoforms in metazoans (Chaubal and Pile, 2018). Sin3, through its interaction with DNA-binding factors, acts as a scaffold protein that recruits histone deacetylases (HDACs) and other chromatin-modifying enzymes onto target promoters (Silverstein and Ekwall, 2005).

Classically, Sin3-HDAC complexes have been associated with transcriptional repression via direct recruitment of transcriptional repressors mediated by Sin3 and nucleosome deacetylation (Kadamb et al., 2013; McDonel et al., 2009; van Oevelen et al., 2008; Sahu et al., 2008). Accumulating evidence, however, points to a dual role of the SIN3 complex in transcriptional regulation. The Sin3A protein in *Drosophila* was found to bind promoters at euchromatic regions, preferentially at TSSs (Das et al., 2013; Saha et al., 2016). The transcriptional profile of a Sin3-deleted yeast strain showed both gene upregulation (173) and downregulation (269), suggesting thus a dual role of this protein in transcription (Bernstein et al., 2000). This tendency was also observed in *Drosophila* cultured cells upon depletion of Sin3A (Gajan et al., 2016; Saha et al., 2016). Sin3 can then positively or negatively influence transcription.

Studies integrating transcriptome and genome-wide binding data in human kidney and fly embryonic cells provided evidence of a direct role of SIN3A in gene transcription. For human SIN3A, 42% of activated genes and 61% of repressed genes were directly bound by this protein (Williams et al., 2011). In *Drosophila* embryonic cells, 92 % and 46 % of the genes repressed and activated by Sin3A, respectively, are direct targets (Saha et al., 2016).

These results suggest thus a direct role of SIN3A both in activation and repression and at least for *Drosophila*, repressed genes are more likely to be direct targets. The corepressor activity of Sin3A is commonly attributed to its association with histone deacetylases. However, the mechanisms through which Sin3A could mediate gene activation are not known. One possibility is that the deacetylase component of the SIN3 complex could act on transcription factors thereby altering DNA-binding capability or protein interactions. For example, it has been observed that acetylation of the transcription factor FoxO1 reduces DNA-binding affinity (Brent et al., 2008). These data show that at least a subset of genes can be direct targets of Sin3A, however the Sin3A-dependent molecular mechanisms regulating transcription are fully understood. Much work remains to determine the network of interactors of Sin3 proteins which will help elucidate how a single complex can have opposite effects on transcription.

C. The nucleosome remodeler subunit Snr1

The conserved SWI/SNF family contains multi-subunit chromatin-remodeling complexes. These are large protein complexes usually formed around Brm (Brahma) in *Drosophila* or BRG1 in Humans, which confer the ATP-hydrolysis catalytic activity. In *Drosophila*, two Swi/Snf complexes exist: the BRM-associated protein complex (BAP) and the polybromo-containing BAP complex (PBAP). These complexes share a common core and differ by signature subunits (Mohrmann et al., 2004). Snr1 is part of the common core of the BAP/PBAP complexes and was found essential for fly viability (Dingwall et al., 1995).

First indications that *snr1* might have a role in gene expression came from its genetic interaction with the TrxG genes *brahma* and *trithorax* (Dingwall et al., 1995). Furthermore, staining in polytene chromosomes showed that a subset of Snr1 overlaps with a fraction of RNA Pol II, suggesting a wide array of potential *in vivo* targets for gene regulation (Zrally et al., 2003). Additionally, an upregulation of the clustered *Ecdysone induced genes* (*Eig*) was observed in loss of function *snr1* mutants and *snr1*-depleted cultured cells. The proposed model by the authors was that upon Snr1 loss, nucleosome accessibility increases and minimizes barriers to transcription. Like this, RNA Pol II proceeds elongation with little or no stalling (Zrally et al., 2006; Zrally and Dingwall, 2012). This suggests thus a function for Snr1 within remodeler activities.

Moreover, it was observed that Snr1 is not required or expressed in all tissues dependent on BAP/PBAP complex activities (Zraly et al., 2003). Snr1 is thus not necessary for all BAP/PBAP functions. Since Snr1 is essential for viability it likely has regulatory functions independent from the BAP/PBAP complexes. RNA-seq analysis in control and Snr1-depleted tumorigenic wing imaginal discs revealed hundreds of misregulated genes upon *snr1* loss and most of them (336 vs 57) were upregulated (Xie et al., 2017). In contrast to *Eig* genes, in this case Snr1, mostly had a repressive role. Strikingly, this study also described a cytoplasmic localization of Snr1 in wing disc cells and salivary gland cells contrary to other core subunits of the BAP/PBAP complexes. This location was required for its tumor-suppressor role. Overall, these data show that Snr1 has gene regulatory functions within and outside the BAP/PBAP complexes. These functions can favor or repress transcription depending on the targets.

D. The multifaceted protein Mod(mdg4)

Between 1993 and 1997, the *mod(mdg4)* gene was cloned three separate times, each time by a different laboratory and was associated to different functions. First, as an enhancer of position-effect variegation, a protein involved in establishing and/or maintaining an open chromatin conformation (Dorn et al., 1993). Next, as a chromatin insulator by directing the repressive effect of Su(Hw) (Gerasimova et al., 1995). Finally, as a protein that induces apoptosis (Harvey et al., 1997). Around the same time, it was also observed that in *mod(mdg4)* mutant larvae, expression of some homeotic genes was decreased, suggesting a positive role for Mod(mdg4) in their regulation (Gerasimova and Corces, 1998). The multifaceted character of Mod(mdg4) could be due to the wide variety of isoforms that come from the single gene. Indeed, at least 31 isoforms sharing a common N-terminal region have been identified from the *mod(mdg4)* gene (Büchner et al., 2000). To date, the most well-studied isoform is the 67.2 for its essential role along with Su(Hw) and CP190 at the *gypsy* insulator (Gerasimova et al., 1995; Melnikova et al., 2017, 2004). However, polytene chromosome staining and ChIP experiments of the 67.2 isoform do not account for all Mod(mdg4) recruitment to DNA (Büchner et al., 2000; Melnikova et al., 2019; Van Bortle et al., 2012). Additionally, a null allele of *mod(mdg4)* showed that this gene is essential but mutations disrupting only the 67.2 isoform are viable (Savitsky et al., 2016) indicating that the roles of Mod(mdg4) go beyond its insulator function.

Little is known about the mechanisms of action of Mod(mdg4) and its influence on gene transcription but mapping of the protein revealed its enrichment at gene promoters in fly embryos (Nègre et al., 2010). Mod(mdg4) does not have sequence specificity and may not be able to bind DNA directly but can mediate homotypic and heterotypic protein–protein interactions via its BTB/POZ domain. This in turn could support pairing between distant regions in the chromosomes (Gurudatta and Corces, 2009; Kyrchanova and Georgiev, 2014). Studies of the role of Mod(mdg4) at the BX-C complex revealed that it promotes *Abd-B* expression in the posterior abdominal segments (Büchner et al., 2000; Dorn et al., 1993; Savitsky et al., 2016) but no mechanism has been proposed. It was observed that *mod(mdg4)* mutations enhance transformations caused by dCTCF deficiency. Also, dCTCF and Mod(mdg4) share many sites at the bithorax complex and genome-wide (Nègre et al., 2010; Savitsky et al., 2016; Van Bortle et al., 2012), favoring a scenario where these two proteins are part of the same molecular pathway. One possibility is that Mod(mdg4) and dCTCF promote the interaction of *Abd-B* promoter with the corresponding enhancers. Since both proteins bind the vicinity of some of the *Abd-B* TSSs, a direct role promoting transcription is also worth considering (Savitsky et al., 2016). Mod(mdg4) has also been found at the borders of H3K27me3 domains. Its depletion led to a decrease of the mark within the domain. Surprisingly, there was no significant effect on expression at genes flanking the domain nor at those embedded in the domain (Schwartz et al., 2012; Van Bortle et al., 2012). However, this was only assessed at a limited number of genes, a genome-wide study would help to evaluate the impact of Mod(mdg4) depletion on transcription. It is well-documented that Mod(mdg4) can affect gene expression when it is part of an insulator complex. However, there is limited data concerning the roles that other isoforms may have in gene transcriptional regulation. Further work is needed to clarify this potential role and to unveil the molecular mechanisms underneath it.

The accumulation of specific features makes Deadhead truly unique. The *dhd* locus is located in an unfavorable genomic environment but achieves high ovary-specific expression, which makes its regulation even more intriguing. Additionally, the epigenomic effectors at play have been shown capable of favoring repression as well as activation of transcription in a context-dependent manner. The *dhd* locus represents thus a paradigmatic gene exquisitely

regulated by chromatin factors and offering a unique opportunity to study the molecular mechanisms at play in epigenetic transcriptional regulation.

RESULTS & DISCUSSION

I. AN SHRNA SCREEN IDENTIFIES MATERNAL CHROMATIN FACTORS REQUIRED FOR THE OOCYTE-TO-ZYGOTE TRANSITION	84
1. ARTICLE PRESENTATION	84
II. THE CONCERTED ACTIVITY OF EPIGENOMIC EFFECTORS IS ESSENTIAL TO ESTABLISH TRANSCRIPTIONAL PROGRAMS DURING OOGENESIS.....	137
1. ARTICLE PRESENTATION	137
III. ADDITIONAL RESULTS AND DISCUSSION	200
1. IS THERE AN “OVARIAN HYPERACTIVATION CODE”?.....	200
2. AN UNUSUAL HETEROCHROMATIC DOMAIN.....	205
3. DIFFERENT ROLES FOR EPIGENOMIC REGULATORS.....	207
4. AN APPROACH TO PROFILE REGULATORY ARCHITECTURE OF CHROMATIN DOMAINS.	209
GENERAL CONCLUSION	211
REFERENCES:.....	212

I. An shRNA screen identifies maternal chromatin factors required for the oocyte-to-zygote transition

1. Article presentation

This first article published in PLOS Genetics presents an shRNA-based genetic screen designed to identify maternal chromatin factors required for the integration of paternal chromosomes into the zygote. This consisted on expressing shRNAs specifically in the female germline and identify embryos that lost paternal chromosomes and developed as gynohaploids. Like this, we identified the histone demethylase Lid/dKDM5 and the members of the HDAC complex Sin3A and Rpd3 as essential for the activation of a critical effector of the oocyte-to-zygote transition: the maternal thioredoxin Deadhead.

To investigate the mechanism underneath this regulation, we focused mainly on Lid/dKDM5. As mentioned in the introduction, H3K4me3 is a hallmark of active transcription and is also the target of the demethylase Lid/dKDM5, therefore we assessed the impact of depletion of *lid* on this mark in ovaries.

Finally, we tested to what extent the phenotype observed in eggs laid by Lid-depleted females was caused by *dhd* loss.

RESEARCH ARTICLE

The Lid/KDM5 histone demethylase complex activates a critical effector of the oocyte-to-zygote transition

Daniela Torres-Campana¹*, Shuhei Kimura²*, Guillermo A. Orsi¹, Béatrice Horard¹, Gérard Benoit¹, Benjamin Loppin¹*

1 Laboratoire de Biologie et de Modélisation de la Cellule, CNRS UMR5239, Ecole Normale Supérieure de Lyon, University of Lyon, France, **2** Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, CNRS, UMR 5558, Villeurbanne F-69622, France

* These authors contributed equally to this work.

† Current address: Seto Hospital, Kanayama-cho, Tokorozawa-shi, Saitama-ken, Japan

* benjamin.loppin@ens-lyon.fr



OPEN ACCESS

Citation: Torres-Campana D, Kimura S, Orsi GA, Horard B, Benoit G, Loppin B (2020) The Lid/KDM5 histone demethylase complex activates a critical effector of the oocyte-to-zygote transition. *PLoS Genet* 16(3): e1008543. <https://doi.org/10.1371/journal.pgen.1008543>

Editor: Giovanni Bosco, Geisel School of Medicine at Dartmouth, UNITED STATES

Received: June 28, 2019

Accepted: November 26, 2019

Published: March 5, 2020

Copyright: © 2020 Torres-Campana et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The DNA sequencing data from this publication have been deposited to the Gene Expression Omnibus database [<https://www.ncbi.nlm.nih.gov/geo/>] and assigned the identifier GSE133064 (RNA-Seq) and GSE133202 (ChIP-Seq).

Funding: SK was supported by a fellowship from the Agence Nationale de la Recherche (<https://anr.fr/>, I) (ZygePal ANR-12-BSV5-0014). The funders had no role in study design, data collection and

Abstract

Following fertilization of a mature oocyte, the formation of a diploid zygote involves a series of coordinated cellular events that ends with the first embryonic mitosis. In animals, this complex developmental transition is almost entirely controlled by maternal gene products. How such a crucial transcriptional program is established during oogenesis remains poorly understood. Here, we have performed an shRNA-based genetic screen in *Drosophila* to identify genes required to form a diploid zygote. We found that the Lid/KDM5 histone demethylase and its partner, the Sin3A-HDAC1 deacetylase complex, are necessary for sperm nuclear decompaction and karyogamy. Surprisingly, transcriptomic analyses revealed that these histone modifiers are required for the massive transcriptional activation of *deadhead* (*dhd*), which encodes a maternal thioredoxin involved in sperm chromatin remodeling. Unexpectedly, while *lid* knock-down tends to slightly favor the accumulation of its target, H3K4me3, on the genome, this mark was lost at the *dhd* locus. We propose that Lid/KDM5 and Sin3A cooperate to establish a local chromatin environment facilitating the unusually high expression of *dhd*, a key effector of the oocyte-to-zygote transition.

Author summary

Nuclear enzymes that add or remove epigenetic marks on histone tails potentially control gene expression by affecting chromatin structure and DNA accessibility. For instance, members of the KDM5 family of histone demethylases specifically remove methyl groups on the lysine 4 of histone H3, a mark generally correlated with gene expression. Lid (Little imaginal discs), the *Drosophila* KDM5, is essential for viability but is also required for female fertility. In this paper, we have found that the specific removal of Lid in developing oocytes perturbs the decompaction of the sperm nucleus at fertilization and the integration of paternal chromosomes in the zygote. Sperm nuclear decompaction normally requires the presence of a small redox protein called Deadhead (Dhd), which is massively

analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

expressed at the end of oogenesis. Strikingly, our analyses of ovarian transcriptomes revealed that the absence of Lid completely abolishes the expression of *dhd*. This direct functional link between a general histone modifier and the expression of an essential terminal effector gene represents a rare finding. We hope that our work will help understanding how histone demethylases function in controlling complex developmental transitions as well as cancer progression.

Introduction

In sexually reproducing animals, fertilization allows the formation of a diploid zygote through the association of two haploid gametes of highly different origins and structures. Generally, the spermatozoon delivers its compact nucleus within the egg cytoplasm, along with a pair of centrioles, while the egg provides one haploid meiotic product and all resources to sustain zygote formation [1]. In some species, this maternal control extends to early embryo development, as in *Drosophila melanogaster*, where the initial amplification of embryo cleavage nuclei occurs without significant zygotic transcription [2]. Instead, the bulk of transcriptional activity takes place in the fifteen interconnected large polyploid germline nurse cells that deposit gene products in the cytoplasm of the growing oocyte [3]. The developmental potential of the egg is thus initially dependent on the establishment of a highly complex transcriptional program in female germ cells.

One of the earliest events of the oocyte-to-zygote transition is the rapid transformation of the fertilizing sperm nucleus into a functional male pronucleus. In *Drosophila*, the needle-shaped, highly compact sperm nucleus is indeed almost entirely organized with non-histone, Sperm Nuclear Basic Proteins (SNBPs) of the protamine like type [4,5]. Male pronucleus formation begins with the genome-wide replacement of SNBPs with maternally supplied histones, a process called sperm chromatin remodeling, which is followed by extensive pronuclear decondensation [1]. Finally, zygote formation involves the coordinated migration and apposition of male and female pronuclei and the switch from meiotic to mitotic division within the same cytoplasm.

Here, we report the results of a genetic screen specifically designed to find new genes required for the oocyte-to-zygote transition in *Drosophila*. Our screen identified two histone modifiers, the Lid/KDM5 histone H3K4me3 demethylase and the Sin3A-HDAC1 histone deacetylase complex, which are both required for the integration of paternal chromosomes into the zygote. These interacting epigenetic factors are known to regulate the expression of hundreds of genes in somatic tissues but their role in the establishment of the ovarian transcriptome is unknown. Strikingly, RNA-Sequencing analyses revealed that, despite the modest impact of their depletion on ovarian transcripts, Lid and Sin3A are critically required for the massive expression of *deadhead* (*dhd*), a key effector of the oocyte-to-zygote transition [6,7]. Furthermore, we demonstrate that germline knock-down of these histone modifiers specifically prevent sperm chromatin remodeling through a mechanism that depends on the DHD thioredoxin.

Results & discussion

A maternal germline genetic screen for gynohaploid embryo development

We performed an *in vivo* RNA interference screen in the female germline to identify genes required for the integration of paternal chromosomes in the zygote. In *Drosophila*, failure to

form a male pronucleus following fertilization is generally associated with the development of haploid embryos that possess only maternally-derived chromosomes (gynohaploid embryos) and that never hatch [8]. We chose to screen transgenic lines from the TRiP collection that express small hairpin RNAs (shRNAs) under the control of the Gal4 activator [9]. We selected shRNA lines that targeted genes with known or predicted chromatin-related function and that show adult ovarian expression (Flybase). Among the 374 tested TRiP lines, 157 (41.9%) induced female sterility or severely reduced fertility when induced with the germline-specific *P*{*GAL4:VP16-nos.UTR*}^{MVD1} Gal4 driver (*nos*-Gal4), thus underlying the importance of chromatin regulation for oogenesis and early embryo development (S1 Table). We then specifically searched for shRNAs that induced a maternal effect embryonic lethal phenotype associated with gynohaploid development (Fig 1A). Gynohaploid embryos can be efficiently identified by scoring the zygotic expression of a paternally-transmitted *P*{*g-GFP::cid*} transgene at the gastrulation stage or beyond [10]. Among the sterile lines with late developing embryos (class 4 in Fig 1A and S1 Table) that were identified, four shRNA lines (GLV21071, GL00612, HMS00359 and HMS00607) produced embryos that were negative for GFP::CID (Fig 1B). Note that none of these shRNAs induced complete female sterility and about 1 to 4% of embryos hatched and were thus diploid (Table 1).

Two of the identified lines (GLV21071 and GL00612) express the same shRNA against the *little imaginal disc (lid)* gene, which encodes the single fly member of the KDM5/JARID1A family of histone demethylases [11,12]. KDM5 demethylases specifically target the trimethylation of lysine 4 of histone H3 (H3K4me3), a mark typically enriched around the Transcriptional Start Site (TSS) of transcriptionally active genes [13,14]. The two other shRNAs (HMS00359 and HMS00607) target the *Sm3A* and *HDAC1/rpd3* genes, respectively. The conserved Sin3A protein scaffold interacts with the histone lysine deacetylase HDAC1 to form the core SIN3 histone deacetylase complex, which is generally considered as a transcriptional repressor [15]. The SIN3 complex regulates the expression of genes involved in a number of metabolic and developmental processes [16–19]. Interestingly, Lid and the largest Sin3A isoform were previously shown to physically and functionally interact [16,17,20,21], thus opening the possibility that these histone modifiers could control the same pathway leading to the formation of a diploid zygote.

Lid and SIN3 are required for sperm chromatin remodeling at fertilization

The Lid demethylase has been previously shown to be required in the female germline for embryo viability [22,23]. Both studies reported a dual phenotype for eggs produced by *lid* KD females (hereafter called *lid* KD eggs): while a majority of eggs fail to initiate development, a variable but significant fraction developed but died at later stages. Our own observations confirmed that about 15% of *lid* KD embryos reach or develop beyond the blastoderm stage (S1 Fig). Furthermore, our analysis of *P*{*g-GFP::cid*} expression (Fig 1B) indicates that most of these late, non viable KD embryos are haploid and develop with maternal chromosomes. To follow the fate of paternal chromosomes in *lid* KD eggs, we crossed *lid* KD females with males expressing the sperm chromatin marker *Mst35Ba::GFP* (*ProtA::GFP*) [24]. In *Drosophila*, protamine-like proteins such as *Mst35Ba* are rapidly removed from sperm chromatin at fertilization [1] and, accordingly, *ProtA::GFP* is never observed in the male nucleus of control eggs. In striking contrast, the vast majority of fertilized *lid* KD eggs contained a needle-shaped sperm nucleus that was still positive for *ProtA::GFP*, indicating that sperm chromatin remodeling was compromised (Fig 1C and 1D). Anti-histone immunostaining indeed revealed that the replacement of SNBPs with maternally supplied histones was variable in KD eggs, ranging from complete absence of histones in the sperm nucleus to the coexistence of variable amounts

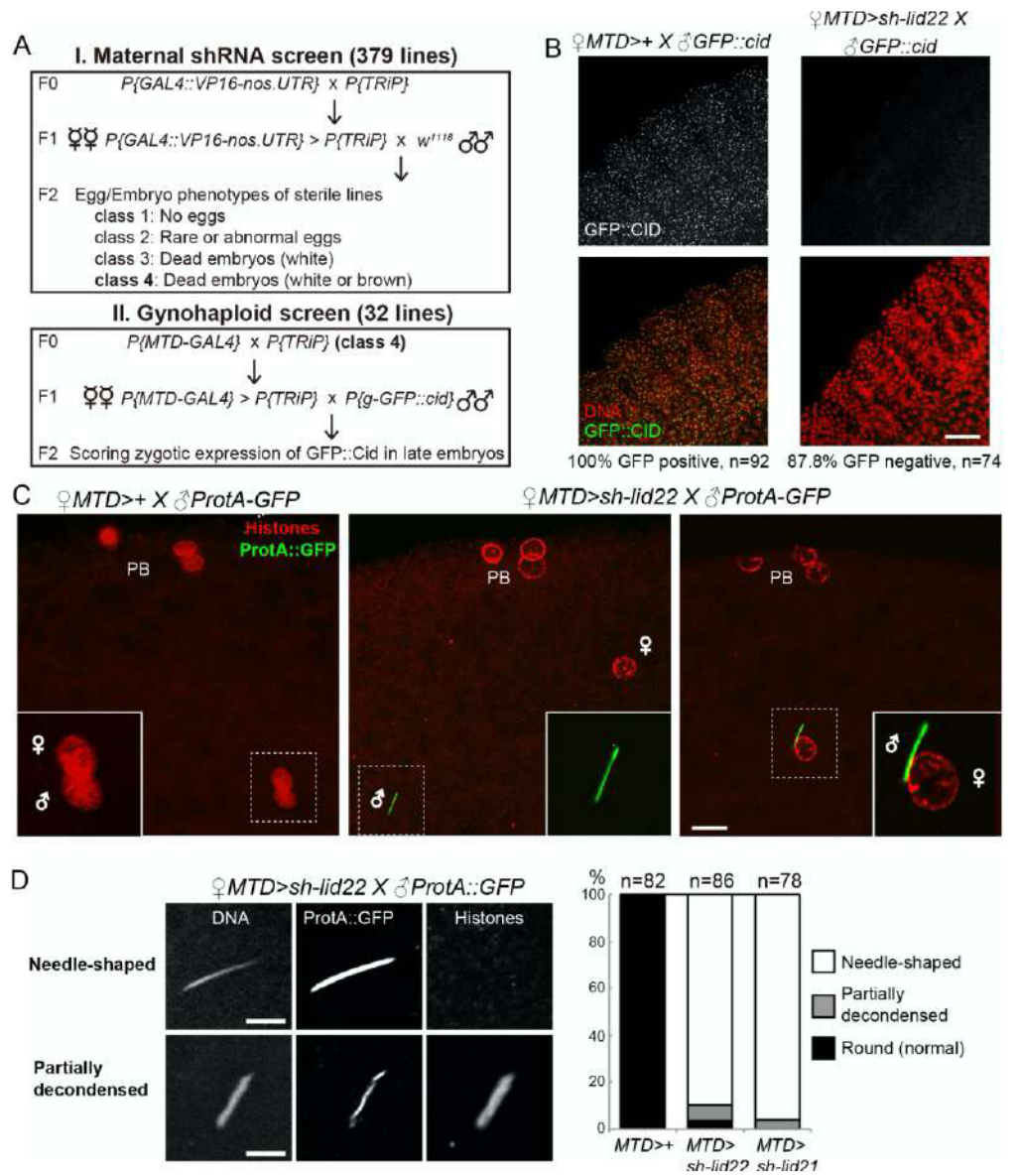


Fig 1. Lid is required for sperm nuclear decompaction at fertilization. A—Scheme of shRNA screen in the female germline. B—Left: Zygotic expression of a paternally-derived $P\{g-GFP::cid\}$ transgene in a control embryo. GFP::Cid marks centromeric chromatin and is visible as nuclear dots. Right: *lid* KD females produce embryos that fail to express the paternally-inherited transgene. Scale bar: 25 μ m. C—Maternal Lid is required for SNBP removal and sperm nuclear decompaction at fertilization. Left: a control egg at pronuclear apposition. Both pronuclei (inset) appear similar in size and shape and the SNBP marker ProtA::GFP is not detected. Middle: A representative *lid* KD egg containing a needle-shaped sperm nucleus (inset) still packaged with ProtA::GFP. Right: A fertilized *lid* KD egg with the sperm nucleus apposed to the female pronucleus. Scale bar: 10 μ m. D—SNBP replacement with histones is impaired in *lid* KD eggs. Left: Confocal images of representative sperm nuclei in *lid* KD eggs. Partially decondensed nuclei are positive for histones. Scale bar: 5 μ m. Right: Quantification of sperm nuclear phenotype in control and *lid* KD eggs.

<https://doi.org/10.1371/journal.pgen.1008543.g001>

Table 1. Embryo hatching rates.

Female genotype	Male genotype	Number of eggs	Hatch. rate (%)
<i>w¹¹¹⁸</i>	<i>w¹¹¹⁸</i>	344	97.67
<i>dhd^{Δ5}</i>	<i>w¹¹¹⁸</i>	375	0.00
<i>dhd^{Δ5}/TM2</i>	<i>w¹¹¹⁸</i>	573	78.36
<i>MTD>+</i>	<i>w¹¹¹⁸</i>	1561	98.27
<i>MTD>sh-lid22</i>	<i>w¹¹¹⁸</i>	1403	2.14
<i>MTD>sh-lid21</i>	<i>w¹¹¹⁸</i>	1144	1.05
<i>MTD>sh-Sin3A</i>	<i>w¹¹¹⁸</i>	1221	0.25
<i>MTD>sh-Rpd3</i>	<i>w¹¹¹⁸</i>	521	4.03
<i>MTD>sh-Hira</i>	<i>w¹¹¹⁸</i>	315	0.00
Rescue with P[<i>dhd^{WT}</i>] genomic transgene			
<i>dhd^{Δ5}; P[<i>dhd</i>]</i>	<i>w¹¹¹⁸</i>	663	85.67
<i>dhd^{Δ5}; P[<i>dhd</i>]/TM2</i>	<i>w¹¹¹⁸</i>	884	94.57
<i>MTD>sh-lid22; P[<i>dhd</i>]</i>	<i>w¹¹¹⁸</i>	1159	4.57
Rescue with P[<i>gnu-dhd^{WT}</i>] transgene			
<i>dhd^{Δ5}; P[<i>gnu-dhd</i>]</i>	<i>w¹¹¹⁸</i>	351	7.25
<i>dhd^{Δ5}; P[<i>gnu-dhd</i>]/TM2</i>	<i>w¹¹¹⁸</i>	1059	3.40
<i>MTD>sh-lid22; P[<i>gnu-dhd</i>]/+</i>	<i>w¹¹¹⁸</i>	695	7.48
Rescue with P[UASP-<i>dhd</i>] inducible transgenes			
<i>nos>+</i>	<i>w¹¹¹⁸</i>	711	95.36
<i>nos> P[UASP-<i>dhd</i>]/+</i>	<i>w¹¹¹⁸</i>	450	97.33
<i>dhd^{Δ5}; nos>P[UASP-<i>dhd</i>]</i>	<i>w¹¹¹⁸</i>	1163	99.28
<i>dhd^{Δ5}; nos>P[UASP-<i>dhd</i>^{XXX3}]</i>	<i>w¹¹¹⁸</i>	702	0.00
<i>MTD>sh-lid22; P[UASP-<i>dhd</i>]</i>	<i>w¹¹¹⁸</i>	1343	28.44
<i>MTD>sh-lid22; P[UASP-<i>dhd</i>^{XXX3}]</i>	<i>w¹¹¹⁸</i>	315	14.60
<i>nos>sh-lid21</i>	<i>w¹¹¹⁸</i>	1508	1.92
<i>nos> P[UASP-<i>dhd</i>]/sh-lid21</i>	<i>w¹¹¹⁸</i>	385	19.32
<i>nos>sh-Sin3a</i>	<i>w¹¹¹⁸</i>	542	2.03
<i>nos> P[UASP-<i>dhd</i>]/sh-Sin3a</i>	<i>w¹¹¹⁸</i>	621	8.86
<i>nos>sh-Rpd3</i>	<i>w¹¹¹⁸</i>	919	2.94
<i>nos> P[UASP-<i>dhd</i>]/sh-Rpd3</i>	<i>w¹¹¹⁸</i>	395	7.59
<i>nos>sh-Hira</i>	<i>w¹¹¹⁸</i>	514	0.00
<i>nos> P[UASP-<i>dhd</i>]/sh-Hira</i>	<i>w¹¹¹⁸</i>	380	0.00

<https://doi.org/10.1371/journal.pgen.1008543.t001>

of histones and ProtA::GFP (Fig 1D). Notably, we noticed that partially decondensed sperm nuclei in *lid* KD eggs were systematically positive for histones. Taken together, our observations indicate that sperm chromatin remodeling is severely impaired in *lid* KD eggs, thus explaining the absence of paternal chromosomes in most developing embryos.

Germline depletion of Lid was previously shown to affect karyosome morphology and chromosome positioning in metaphase I oocytes [25]. In contrast, another study had previously reported that meiotic progression was not affected in *lid* KD oocytes [22]. Interestingly, although our own observations indeed confirmed the aberrant karyosome structure in *lid* KD oocytes, we observed that the second meiotic division appeared to resume normally in a vast majority of *lid* KD eggs (S2 Fig). We noted, however, that, following meiosis, the female pronucleus frequently (62%, n = 76) failed to appose to the sperm nucleus in *lid* KD eggs (Fig 1C). Thus, we propose that defective pronuclear migration largely accounts for the previously reported failure of *lid* KD embryos to initiate cleavage divisions [23]. Like Navarro-Costa

et al., we noted that the rosette of polar body chromosomes was frequently abnormal in morphology, but this phenotype appeared independent of meiosis *per se*.

Remarkably, we found that germline KD of *Sin3A* also induced a highly penetrant sperm nuclear phenotype with all scored fertilizing sperm nuclei retaining a needle-like shape (100%, $n = 19$) (S3 Fig). Finally, a similar but less penetrant phenotype was also observed in *ripd3* KD eggs (S3 Fig). As this low penetrance could result from less efficient gene knock-down, we chose to mainly focus on *lid* and *Sin3A* in the rest of this study.

Transcriptomic analysis of *lid* KD and *Sin3A* KD ovaries identifies *deadhead* as a common and major target gene

In eggs from wild-type females, anti-Lid immunostaining failed to detect Lid protein in the male or female pronucleus, thus suggesting that its implication in sperm chromatin remodeling was indirect. In fact, Lid was not detected in embryos before the blastoderm stage (S4 Fig). We thus turned to RNA sequencing (RNA Seq) to analyze the respective impact of *lid* KD and *Sin3A* KD on the ovarian transcriptome.

We compared transcriptomes obtained from *lid* KD or *Sin3A* KD ovaries (using the maternal triple driver MTD-Gal4) with control transcriptomes (MTD-Gal4 only). Globally, our analyses revealed a relatively modest impact of *lid* KD and *Sin3A* KD on ovarian gene expression, with more genes downregulated in both cases (Fig 2B and S5 Fig; S2 Table). Note that these changes are expected to reflect the activity of Lid and Sin3A in germ cells, as ovarian somatic cells (see Fig 2A) do not express the targeting shRNAs. In their transcriptome analysis (based on microarrays) of *lid* KD wing imaginal discs, Azorin and colleagues found a similar number of differentially-expressed genes, most of them being downregulated [14]. However, RNA Seq analyses recently published by Drelon *et al.* in contrast found 1630 genes (FDR<0.05) dysregulated in wing discs from a null *lid* mutant [26]. Moreover, Liu and Secombe [27] found 8,056 genes differentially expressed (60% were down-regulated) in *lid* adult mutant flies (FDR<0.05), a number which could reflect the greater cell type complexity involved in this analysis.

We found that only 29% (139) of the 473 differentially-expressed genes in *lid* KD were also differentially-expressed in *Sin3A* KD, and only 100 genes (21%) were dysregulated in the same direction in both KD (Fig 2B). As a matter of comparison, Gajan *et al.* found a 65% overlap in *Drosophila* S2 cells [17]. As the *Sin3A* shRNA that we used targets all predicted alternatively spliced mRNAs, this suggests that the knock-down affects Sin3A isoforms with Lid-independent functions [16].

Remarkably, however, we noticed that the *deadhead* (*dhd*) gene was by far the most severely impacted gene, downregulated by more than two orders of magnitude in both *lid* KD and *Sin3A* KD transcriptomes (Fig 2C, S6 Fig). The implication of *dhd* appeared particularly interesting because we and others have recently established that this germline specific gene is critically required for sperm nuclear decompaction at fertilization [6,28]. *dhd* indeed encodes a specialized thioredoxin that cleaves disulfide bonds on SNBPs, thus facilitating their removal from sperm chromatin [6,28]. RT-PCR and Western-blot analyses confirmed the severe down-regulation of *dhd* in *lid* KD and *Sin3A* KD (S5 Fig).

dhd is a small, intronless gene located in the middle of a cluster of fifteen densely packed genes spanning about 40 kb of genomic DNA. Interestingly, the *dhd* gene lies within a 1.4 kb region that is immediately flanked by two genes with testis-specific expression (*Trx-T* and *CG4198*) (Fig 2D). Despite this apparently unfavorable genomic environment, *dhd* is one of the most highly-expressed genes in ovaries [29], as confirmed by our RNA Seq profiles (Fig 2C and S6 Fig). Interestingly, although this 40 kb region contains six additional genes expressed

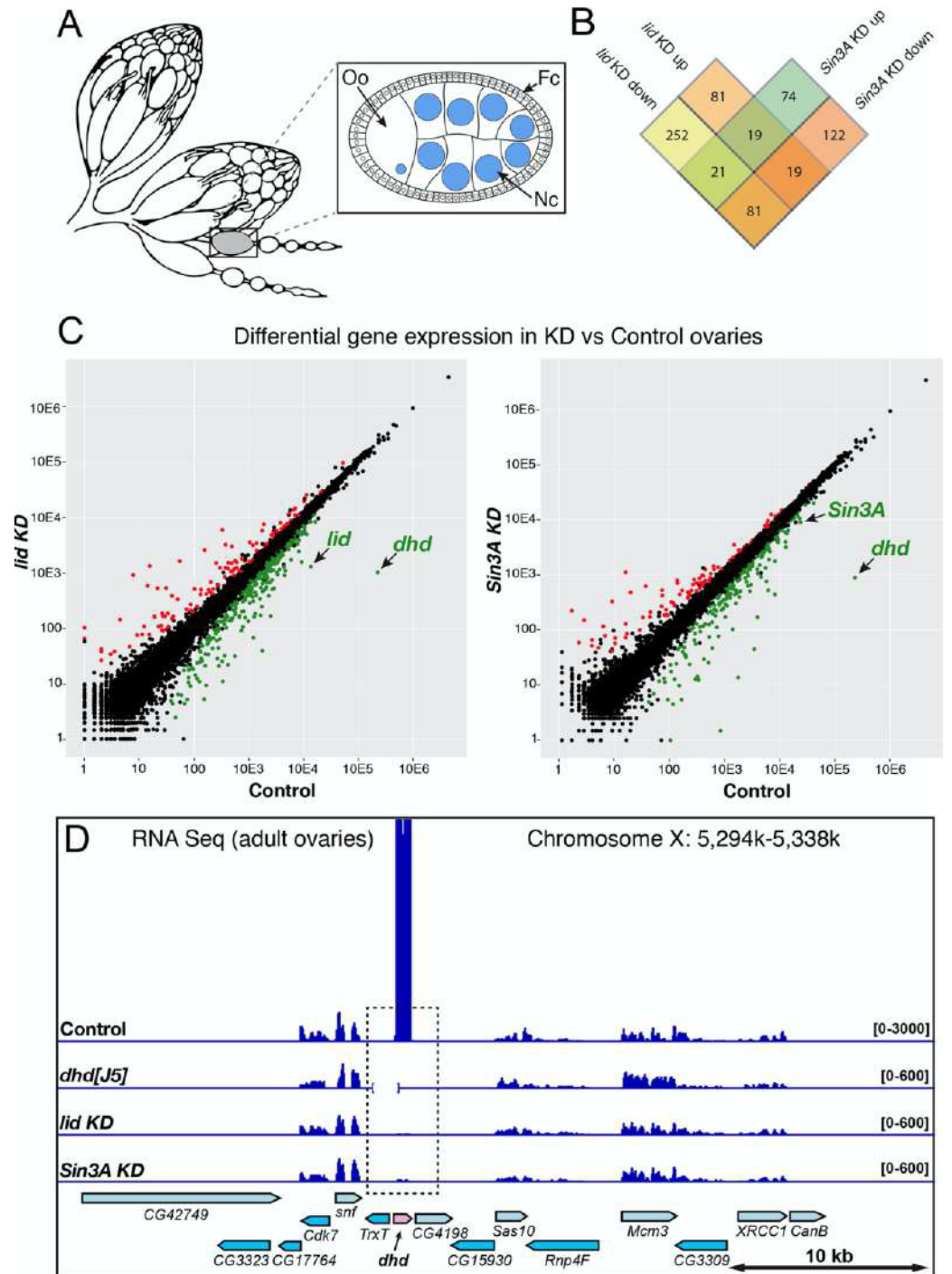


Fig 2. *deadhead* is strongly downregulated in *lid* KD and *Sin3A* KD ovaries. A—Scheme of a pair of adult ovaries with two isolated ovarioles and an egg chamber (inset). Germline nuclei are in blue. Oo: Oocyte, Ne: Nurse cells, Fc: Follicle cells. B—Venn diagram showing the number of differentially expressed genes in *lid* KD and *Sin3A* KD ovarian transcriptomes (FDR<0.001). C—Comparison of RNA seq normalized reads per gene (DESeq2) are shown for *lid* KD vs Control (left) and *Sin3A* KD vs Control (right). Genes with a negative foldchange (downregulated in KD) are in green (FDR<0.001). Genes with a positive foldchange (upregulated in KD) are in red (FDR<0.001). D—Integrative Genomics Viewer (igv) view of Control, *dhd*^{+/+}, *lid* KD and *Sin3A* KD ovarian RNA Seq signal on the *dhd* region. The *Df(1)15* deficiency is indicated as an interrupted baseline on the *dhd*^{+/+} track.

<https://doi.org/10.1371/journal.pgen.1008543.g002>

in ovaries, *dhd* is the only one affected by *lid* KD or *Sin3A* KD (Fig 2D). Thus, Lid and the SIN3 complex exert a critical and surprisingly specific control on the transcriptional activation of *dhd* in female germ cells.

Impact of Lid depletion on the distribution of H3K4me3 in ovaries

To evaluate the impact of *lid* KD on the distribution of its target histone mark in the female germline, we performed chromatin immunoprecipitation and sequencing (ChIP-Seq) analyses of H3K4me3 in ovaries from control and *lid* KD females. Consistent with earlier reports of a global increase of H3K4me3 in *lid* mutant tissues [14,23,25,30], we observed that H3K4me3 ChIP peaks in *lid* KD ovaries were globally more pronounced compared to control ovaries (Fig 3A). Our analysis actually revealed that about 10% (1528) of H3K4me3 identified peaks in control ovaries were significantly increased in *lid* KD ovaries (S3 Table; FDR<0.05). For those peaks that were associated with genes, the relative enrichment of H3K4me3 in *lid* KD ovaries mainly affected the promoter region and gene body (Fig 3B). A similar effect was previously observed in *lid* depleted wing imaginal discs, with H3K4me3 abundance specifically increased at the TSS of Lid direct target genes [14]. We nevertheless found 46 H3K4me3 peaks that were significantly decreased in *lid* KD ovaries compared to control ovaries (S3 Table; FDR<0.05). Among these, the H3K4me3 peak on the *dhd* gene was the second most severely affected (S3 Table and Fig 3C). Furthermore, only ten of the negatively affected peaks covered genes that were downregulated in *lid* KD ovaries, including *dhd*. Remarkably, the prominent H3K4me3 peak on *dhd* was almost completely lost in *lid* KD ovaries while other peaks within the *dhd* region remained essentially unchanged. *CG4198*, which lies immediately downstream of *dhd* is a notable exception, as this gene also shows a decrease of H3K4me3 (Fig 3C). In this case, however, it is interesting to note that the presence of the mark is not correlated with transcriptional activity.

At first, the paradoxical loss of H3K4me3 enrichment on the *dhd* gene upon Lid depletion suggests that the demethylase activity of Lid is not locally responsible for this regulation. In fact, it has been established that *lid* mutant females with a catalytic dead JmjC⁻ *lid* rescue transgene are viable and at least partially fertile [23,26,31], thus suggesting that the catalytic activity of Lid is not absolutely required to form a viable zygote. To directly test the implication of the Lid demethylase domain in *dhd* regulation, we measured *dhd* expression in JmjC⁻ rescued females. Surprisingly, *dhd* transcripts were severely reduced in these females compared to control females rescued with a wild-type *lid* transgene (S7 Fig). We thus conclude that the demethylase activity of Lid is important for *dhd* expression. At least two hypotheses could reconcile this conclusion with our ChIP-Seq results. First, it is possible that the demethylase activity of Lid is only transiently required at the *dhd* locus, perhaps to switch on its transcription at mid oogenesis, while later, massive *dhd* expression would no longer require this activity. Alternatively, Lid could exert a control on *dhd* transcription by restricting the level of H3K4me3 at an enhancer element, similarly to what was previously reported for KDM5 demethylases in different model organisms [32–34]. Although our ChIP-Seq analyses failed to identify any

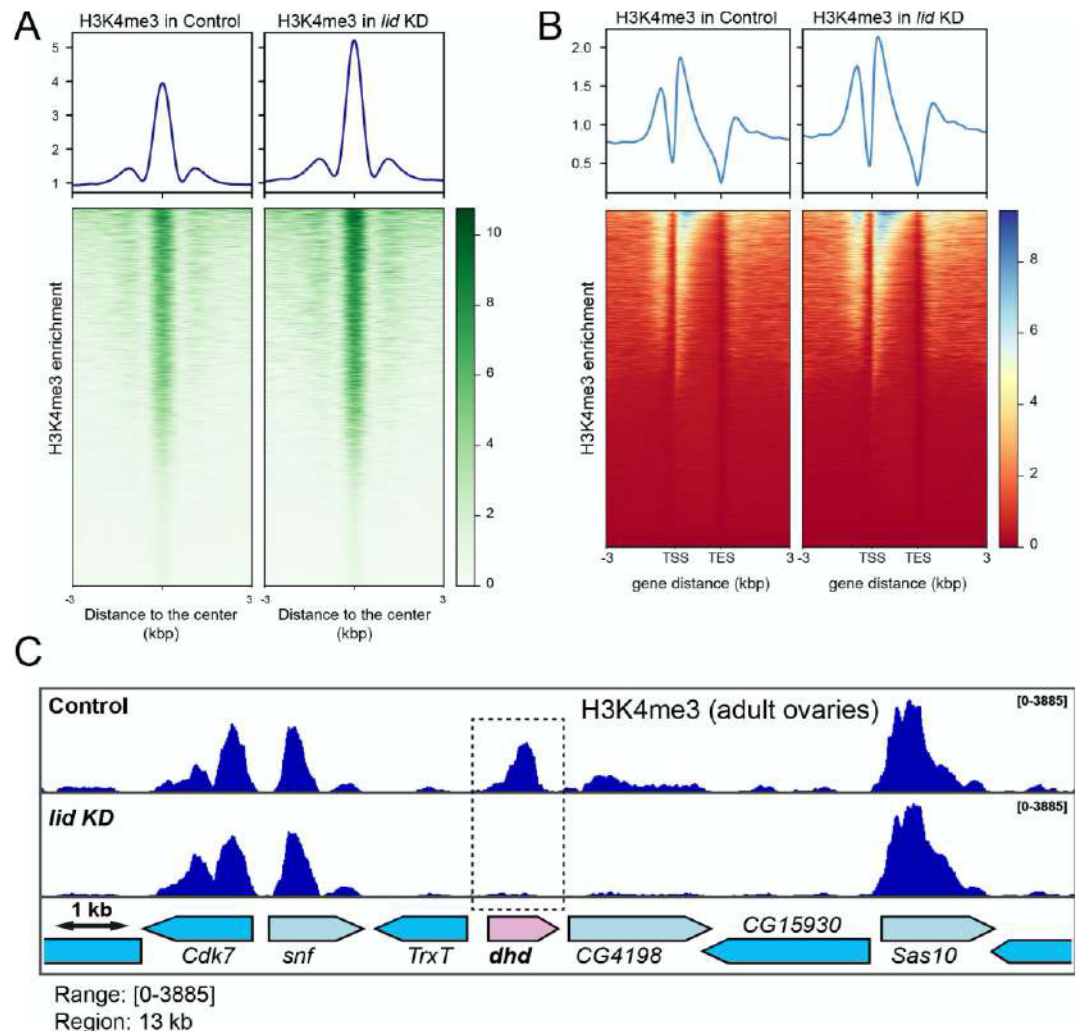


Fig 3. H3K4me3 ovarian ChIP-Seq analysis. A—H3K4me3 enrichment around peak center for Control and *lid* KD ovaries. Upper panels show the average profile around detected peak centers. Lower panels show read density heatmaps around the detected peak centers. B—H3K4me3 enrichment around gene loci for Control and *lid* KD ovaries. Upper panels show the average signal profile on genomic loci defined as 3kb upstream of annotated TSS to 3kb downstream of annotated TES. Lower panels show read density heatmaps around the same genomic loci. C—igv view of H3K4me3 occupancy on the *dhd* genomic region in Control and *lid* KD ovaries.

<https://doi.org/10.1371/journal.pgen.1008543.g003>

obvious candidate enhancer element in the vicinity of the *dhd* gene region, we cannot exclude this possibility.

Besides its JmjC demethylase domain, Lid/KDM5 possesses a conserved C-terminal PHD motif capable of binding H3K4me2/3. This binding motif is required for the recruitment of Lid at the promoter of target genes, where it could promote their activation [27]. The local recruitment of Lid, either through its C-terminal PHD motif or through its DNA binding ARID (AT-rich interaction domain) motif, or both, could thus establish a chromatin

environment permissive to *dhd* massive expression in late oogenesis. In this context, the role of the Sin3A/HDAC1 complex also remains to be clarified. The Sin3A histone deacetylase complex is generally considered as a transcriptional repressor [15], but it also functions as a transcriptional activator in *Drosophila* S2 cells [17]. Lacking an intrinsic DNA binding ability [15], the recruitment of this complex to chromatin requires an additional factor. It is thus tempting to propose that Lid itself could recruit Sin3A/HDAC1 locally to activate *dhd* expression in female germ cells.

Forced *dhd* expression in *lid* KD ovaries partially restores sperm chromatin remodeling at fertilization

Taken together, our cytological and transcriptomic analyses strongly suggest that the loss of *dhd* expression in *lid* KD ovaries at least contributes to the observed fertilization phenotype. To directly test this possibility, we attempted to restore *dhd* expression in *lid* KD female germ cells through the use of transgenic constructs. A genomic transgene (P[*dhd*]) that fully rescued the fertility of *dhd* mutant females [6] only had a very limited impact on the hatching rate of *lid* KD embryos (Table 1) but quantitative RT-PCR analyses revealed that P[*dhd*] remained essentially silent in *lid* KD ovaries (Fig 4A). This result indicates that the 4.3 kb genomic region present in this transgene is sufficient to recapitulate the endogenous control exerted by Lid on *dhd* transcription. We then designed another transgene expressing *dhd* under the control of the *giant nuclei* (*gnu*) regulatory sequences. Like *dhd*, *gnu* is specifically expressed during oogenesis and is functionally required during zygote formation [35]. In addition, our RNA Seq data indicated that its expression is not controlled by Lid. We observed that the P[*gnu-dhd*] transgene indeed restored fertility to *dhd* homozygous mutant females albeit to modest level (about 7% embryo hatching rate; Table 1). In fact, rescued females only produced about 10% of the normal amount of *dhd* mRNA in their ovaries and the DHD protein remained almost undetectable in Western-blot (Fig 4A). Interestingly, when introduced into *lid* KD females, the P[*gnu-dhd*] transgene also slightly increased embryo hatching rate (Table 1). Furthermore, cytological examination of eggs laid by these females revealed a limited but clear improvement of sperm nuclear decondensation (Fig 4B and 4C). These results thus indicate that forced expression of *dhd* can improve the survival of *lid* KD eggs through its positive impact on sperm chromatin remodeling. Finally, we tried to further increase the level of expression of *dhd* by using a Gal4 inducible transgene, P[UAS-*dhd*^{WT}]. Indeed, induction of this transgene in the germline of *dhd*^Δ mutant females fully restored their fertility (Table 1). We also observed a strong effect on the hatching rate of embryos laid by *lid* KD, P[UAS-*dhd*^{WT}] females (about 28%; Table 1). However, a P[UAS-*dhd*^Δ] transgene expressing a catalytic mutant DHD with no rescuing potential also improved the fertility of *lid* KD females. This effect suggests that Gal4 becomes limiting in the presence of two UAS transgenes, with a negative impact on knock-down efficiency. The fertility of P[UAS-*dhd*^{WT}] rescued females was nevertheless doubled compared to P[UAS-*dhd*^Δ] control females (Table 1), thus supporting the idea that partial *dhd* re-expression in *lid* KD ovaries significantly improved the probability of these eggs to form a viable, diploid zygote. Besides its already established roles in controlling the oocyte epigenome and the architecture of meiotic chromosomes, we show that the transcriptional regulation of *dhd* (and possibly additional early acting genes) is indeed a critical function of Lid and associated factors in female germ cells.

Trr controls sperm chromatin remodeling through a DHD-independent pathway

Intriguingly, germline KD of the Trithorax group protein Trithorax-related (Trr, also known as dMLL3/4), a histone methyltransferase responsible for monomethylation of H3K4 [36], was

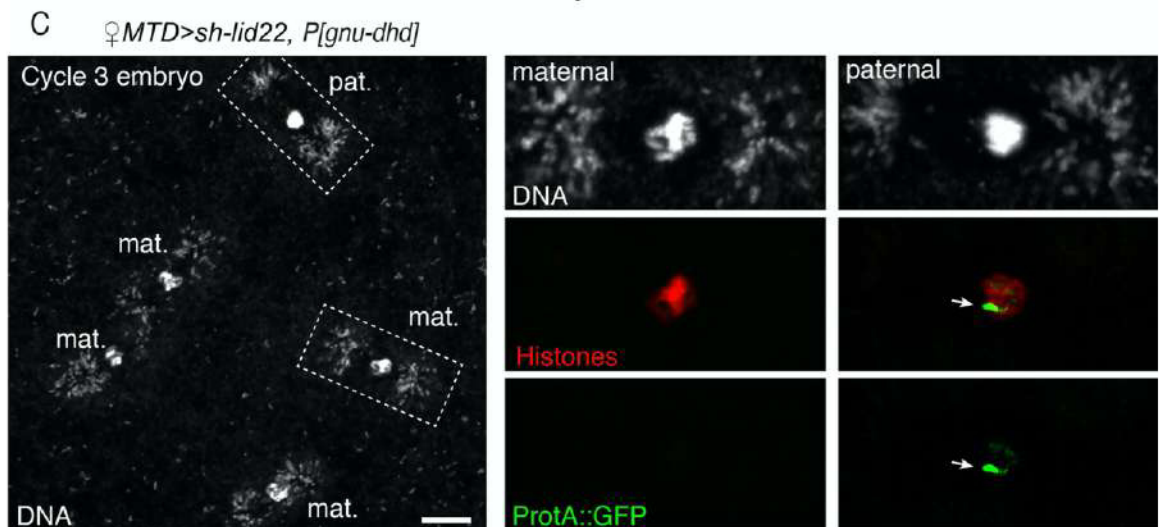
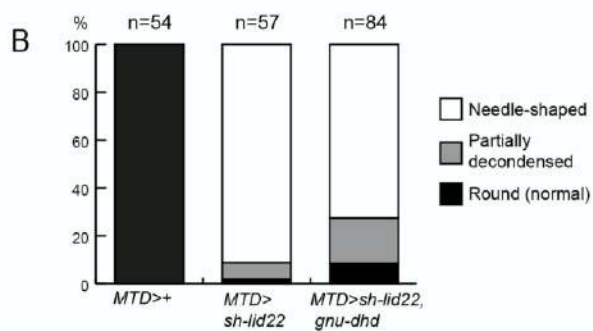
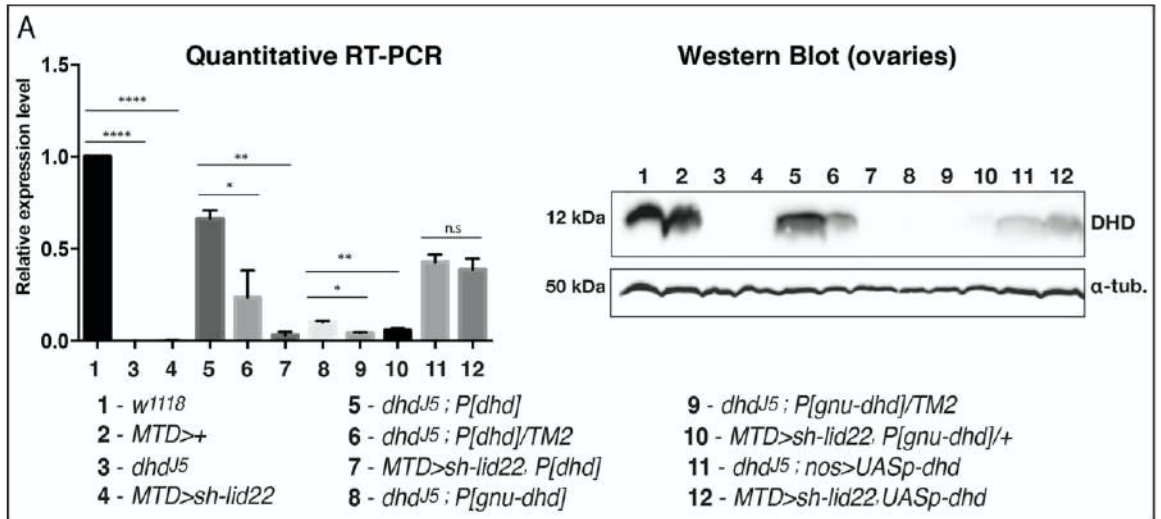


Fig 4. Forced expression of *dhd* partially rescues the *lid* KD phenotype. A—Left: RT-qPCR quantification of *dhd* mRNA levels in ovaries of indicated genotypes (normalized to *rp49* and relative to expression in w^{1118}). Data are presented as mean \pm SD of 2 biological replicates. P values indicate one-way ANOVA with Dunnett's multiple comparisons test to a control (** $P < 0.0001$; * $P < 0.01$; † $P < 0.05$; n.s. = not significant). Right: Western blot analysis of DHD expression in ovaries of indicated genotypes. Alpha-tubulin detection is used as a loading control in Western blotting. B—Quantification of sperm nuclear phenotype in eggs laid by females of indicated genotypes. C—Confocal images of a *MTD>hid22; P[*gou-dhd*]* haploid embryo during the third nuclear division with incompletely remodeled male nucleus. Karyogamy has failed and the embryo contains four haploid nuclei of presumably maternal origin. The paternal nucleus (inset) still contains a region packaged with ProTA::GFP (arrow). Note that DNA positive dots at the spindle poles are *Wolbachia* endosymbionts. Scale bar: 10 μ m.

<https://doi.org/10.1371/journal.pgen.1008543.g004>

recently shown to induce a sperm decondensation defect at fertilization similar to the one reported here for *lid*, *Sin3a* and *tpd3* KD [37]. In their study, however, Prudêncio *et al.* did not find any significant change in *dhd* mRNA level in *trr* KD early embryos. To more directly exclude any implication of DHD in the *trr* KD phenotypes, we stained control and KD eggs with an anti-DHD antibody. At fertilization, maternally-expressed DHD is abundant throughout the egg cytoplasm (100%, $n = 41$) but is rapidly degraded after pronuclear apposition. As expected, DHD protein remained undetectable in most *lid* KD eggs (92%, $n = 50$), including those that were fixed before the end of meiosis II (S8 Fig). In sharp contrast, DHD protein was normally detected in a majority of *trr* KD eggs, even though these eggs indeed contained a needle-shaped sperm nucleus still packaged with SNBPs (S8 Fig). This result thus confirms that *Trr/dMLL3/4* controls sperm nuclear remodeling through a yet unknown, DHD-independent mechanism. Conversely, *Trr* was shown to control meiosis progression through the activation of *Ldgf4* [37], a gene that is not affected by *lid* or *Sin3a* KD (this study). Thus *Trr* and *Lid/Sin3A* respectively activate a distinct repertoire of genes important for the oocyte-to-zygote transition and sperm chromatin remodeling.

Conclusion

Our maternal germline genetic screen has unveiled a complex and remarkably specific transcriptional regulation of the *dhd* gene by *Lid/KDM5* and the *Sin3A/HDAC1* complex. In addition to its crucial role in sperm protamine removal at fertilization, DHD was recently involved in the establishment of a redox state balance at the oocyte-to-zygote transition with a number of identified target proteins [7]. This important DHD-dependent thiol proteome remodeling is thus ultimately controlled by *Lid* and the *SIN3* complex, underlying the critical contribution of these transcriptional regulators to this delicate developmental transition. Future work will aim at dissecting the chromatin mechanisms at play in setting up *dhd* specific activation in female germ cells.

Materials & methods

Drosophila strains

Flies were raised at 25°C on standard medium. The w^{1118} strain was used as a wild-type control. shRNAs lines used in this study (see S1 Table) were established by the Transgenic RNAi Project (TRiP) at Harvard Medical School and were obtained from the Bloomington Drosophila Stock Center at Indiana University. The *hid* and *Sin3A* shRNA lines target all predicted isoforms of their respective target genes. Additional stocks were *EGFP-Cid* [38], *P[otu-GAL4::VP16.R]1*; *P[GAL4-nos.NGT]40*; *P[GAL4::VP16-nos.UTR]MVD1* ("MTD-Gal4"), *P[GAL4::VP16-nos.UTR]MVD1* ("nos-Gal4"), *P[Mst35Ba-EGFP]* [24] and *Df(1)J5/EM7c* [39]. The *pUASP-dhd[WI]* and *pUASP-dhd[SXXS]* transgenic flies were a gift from C. Gonzalez and C. Molnar. The *PatB[gLid-WT-HA]* and *PatB[gLid-JMJC-HA]* transgenic flies were kindly provided by P. Navarro-Costa [23].

Germline knock-down and fertility tests

To obtain *KD* females, virgin shRNA transgenic females were mass crossed with transgenic Gal4 males at 25°C and females of the desired genotype were recovered in the F1 progeny. To measure fertility, virgin females of different genotypes were aged for 2 days at 25°C in the presence of males and were then allowed to lay eggs on standard medium for 24 hours. Embryos were counted and then let to develop for at least 36 hours at 25°C. Unhatched embryos were counted to determine hatching rates.

Immunofluorescence and imaging

Early (0–30 min) and late (about 6 hours) embryos laid by randomly selected females were collected on agar plates. Embryos were dechorionated in bleach, fixed in a 1:1 heptane:methanol mixture and stored at -20°C. Embryos were washed three times (10 min each) with PBS1X 0.1%, Triton X-100 and were then incubated with primary antibodies in the same buffer on a wheel overnight at 4°C. They were then washed three times (20 min each) with PBS 0.1%, Triton X-100. Incubations with secondary antibodies were performed identically. Embryos were mounted in Dako mounting medium containing DAPI.

Ovaries were dissected in PBS-Triton 0.1% and fixed at room temperature in 4% formaldehyde in PBS for 25 minutes. Immunofluorescence was performed as for embryos except for secondary antibodies that were incubated four hours at room temperature. Ovaries were then mounted as described above.

Primary antibodies used were mouse monoclonal anti-histones (Sigma #MABE71; 1:1000), rabbit polyclonal anti-DHD (1:1000) [6], rat polyclonal anti-Lid (1:500) [14], mouse monoclonal anti-GFP (Roche #11814460001; 1:200) and Rat monoclonal anti α -tubulin (Abcam #ab6160; 1:50). Secondary antibodies were goat anti-rabbit antibodies (ThermoFisher Scientific, 1:500), goat anti-mouse or anti-rat antibodies (Jackson ImmunoResearch, 1:500) conjugated to AlexaFluor. Images were acquired on an LSM 800 confocal microscope (Carl Zeiss). Images were processed with Zen imaging software (Carl Zeiss) and Photoshop (Adobe).

Western blotting

Ovaries from 30 females were collected and homogenized in lysis buffer (20mM Hepes pH7.9, 100mM KCl, 0.1mM EDTA, 0.1mM EGTA, 5% Glycerol, 0.05% Igepal and protease inhibitors (Roche)). The protein extracts were cleared by centrifugation and stored at -80°C.

Eggs were collected every 30 min, dechorionated in bleach and quickly frozen in liquid nitrogen. Protein extracts were prepared from ca. 10 μ l of embryos. Protein samples were run on 15% SDS polyacrylamide gel and transferred to Immun-Blot[®] PVDF membrane (Bio-Rad) for 1h at 60V. Membranes were blocked for 1h at room temperature in 5% non-fat milk in PBS 1X-Tween20 0.05%, followed by an overnight incubation with the primary antibody at 4°C in 5% non-fat milk in PBS1X-Tween20 0.05%. Secondary antibodies used were added and incubated for 2 hours at room temperature. Protein detection was performed using ECL solution according manufacturer's instruction (GE Healthcare). Antibodies used were: rabbit polyclonal anti-DHD (1/1000) [6], mouse monoclonal anti- α -Tubulin (Sigma Aldrich #T9026, 1:500), HRP-conjugated goat anti-mouse (Biorad #170-5047; 1:50 000) and peroxidase-conjugated goat anti-rabbit (Thermoscientific #32460; 1:20 000).

Gene expression analysis by RT-QPCR

Total RNA was extracted from ovaries of 3-day-old females using the NucleoSpin[®] RNA isolation kit (Macherey-Nagel), following the instructions of the manufacturer. Duplicates were

processed for each genotype. cDNAs were generated from 1µg of purified RNA with oligo (dT) primers using the SuperScript™ II Reverse Transcriptase kit (Invitrogen).

Generated cDNAs were diluted to 1/5 and were used as template in a real time quantitative PCR assay using SYBR® Premix Ex Taq™ II (Tli RNaseH Plus) (Takara). All qRT-PCR reactions were performed in duplicate using Bio-Rad CFX-96 Connect system with the following conditions: 95°C for 1 min followed by 40 cycles of denaturation at 95°C for 10 s, annealing at 59°C for 30 s and extension at 72°C for 30 s. Relative fold change in gene expression was determined by the comparative quantification ΔΔCT method of analysis [40]. The housekeeping gene *rp49* was used to normalize cDNA amounts in the comparative analysis. The primer sets used in the PCR reactions were: *dhd*-forward 5'-TCTATGCGACATGGTGTGGT-3' and *dhd*-reverse 5'-TCCACATCGATCTTGAGCAC-3'; *lid*-forward 5'-ATTGGTTTCACGAGGATTGC-3' and *lid*-reverse 5'-CATAGCCACTTGGGTCGATT-3'; *Rp49*-forward 5'-AAGATCGTGAAGAA GCGCAC-3' and *Rp49*-reverse 5'-GATACTGTCCCTTGAAGCGG-3'. Statistical tests were performed using GraphPad Prism version 6.00 for Mac OS X (GraphPad Software).

Ovarian RNA sequencing and analysis

For each samples, 8 pairs of ovaries were dissected from 6 day old virgin females and total RNA were extracted using the NucleoSpin® RNA isolation kit (Macherey-Nagel), following the instructions of the manufacturer. Extracted RNAs were treated with Turbo™ DNase (Ambion #AM2238). After DNase inactivation, RNAs were purified using the NucleoSpin® RNA Clean-up XS kit (Macherey-Nagel) according to manufacturer's instructions. Sequencing was completed on two biological replicates of each genotype:

Control KD (*MTD-Gal4>+*)

$P\{w[+mC] = \text{otu-GAL4:VP16.R}1, w[\]/y[1] v[1]; P\{w[+mC] = \text{GAL4-nos.NGT}1/40/+; P\{w[+mC] = \text{GAL4:VP16-nos.UTR}CG6325[MVD1]/P\{y[+7.7] = \text{CaryP}attP2$

lid KD (*MTD-Gal4>shRNA lid*)

$P\{w[+mC] = \text{otu-GAL4:VP16.R}1, w[\]/y[1] sc[\] v[1]; P\{w[+mC] = \text{GAL4-nos.NGT}1/40/+; P\{w[+mC] = \text{GAL4:VP16-nos.UTR}CG6325[MVD1]/P\{y[+7.7] v[+1.8] = \text{TRiP.GLV21071}attP2$

Sin3A KD (*MTD-Gal4>shRNA Sin3A*)

$P\{w[+mC] = \text{otu-GAL4:VP16.R}1, w[\]/y[1] sc[\] v[1]; P\{w[+mC] = \text{GAL4-nos.NGT}1/40/+; P\{w[+mC] = \text{GAL4:VP16-nos.UTR}CG6325[MVD1]/P\{y[+7.7] v[+1.8] = \text{TRiP.HMS00359}attP2$

Sequencing libraries for each sample were synthesized using TruSeq Stranded mRNA kit (Illumina) following supplier recommendations (Sample Preparation Guide—PN 15031047, version Rev.E Oct 2013) and were sequenced on Illumina HiSeq 4000 sequencer as Single-Reads 50 base reads following Illumina's instructions (GenomEast platform, IGBM, Strasbourg, France). Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover. Sequenced reads were mapped to the *Drosophila melanogaster* genome assembly dm6 using TopHat (version 2.1.1) with default option. The aligned reads were assigned to genes by FeatureCounts, run with default options on the dmel-all-r6.15 version of the *Drosophila melanogaster* genome annotation. Differentially expressed genes were identified using the R-package DESeq2 (version 1.14.1). The annotated genes exhibiting an adjusted-P < 0.001 were considered differentially expressed compared to Control.

Chromatin immunoprecipitation, sequencing and analysis

ChIP assays were performed as previously described [41]. Two biological replicates for *control* KD and *lid* KD ovaries (same genotypes as for RNA Seq) were processed and analyzed. For

each biological replicate, eighty ovary pairs were dissected from 2 day-old females and flash frozen. Dissected ovaries were fixed in 1.8% formaldehyde at room temperature for 10 minutes. Chromatin was sonicated using a Diagenod Bioruptor (18 cycles, high intensity, 30s on/30s off) to generate random DNA fragments from 100 to 800 base pairs. Sheared chromatin was incubated overnight at 4°C with H3K4me3 antibody (ab8580 Abcam). Immunoprecipitated samples were treated with RNase A, proteinase K and DNA purified using the ChIP DNA Purification kit (Active Motif #58002) following the manufacturer's instructions. Quantification assessment of purified DNA was done using Qbit dsDNA HS Assay on the Qbit fluorometer (Invitrogen). Immunoprecipitated DNA quality was evaluated on a Bioanalyzer 2100 (Agilent).

Sequencing libraries for each sample were synthesized using Diagenode MicroPlex Library Preparation kit according to supplier recommendations (version 2.02.15) and were sequenced on Illumina HiSeq 4000 sequencer as Paired-End 50 base reads following Illumina's instructions (GenomEast platform, IGBM, Strasbourg, France). Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover. Sequenced reads were mapped to the *Drosophila melanogaster* genome assembly dm6 using Bowtie (version 2.3.3) with default option. Only uniquely aligned reads have been retained for further analyses. Duplicated reads were removed using picard-tools (version 2.17.10). Peak calling was performed for each individual samples and on merged biological replicates using MACS algorithm (version 2.1.1) with default option and a relaxed q-value cut-off of 0.1. Consistent peaks between biological replicates were identified using irreproducible discovery rate (IDR version 2.0.3) with a 0.05 cut-off. Differentially modified H3K4me3 peaks between Control and *lid* Knock-down ovaries were identified using the R-package DiffBind (version 2.2.12) with a 0.05 FDR cut-off.

Data visualization

The Deeptools software was used to convert alignment files to bigwig (bamCoverage) and to generate H3K4me3 heatmap and density profiles (computeMatrix and plotHeatmap). The generated bigwig files were visualized using IGV software.

Supporting information

S1 Fig. Developmental defects of *lid* KD embryos. Embryos were collected for four hours and aged for another four hours at 25°C before DAPI staining and examination in fluorescent microscopy. More than 85% of *lid* KD embryos arrest development before the blastoderm stage. In contrast, 100% of control embryos had reached gastrula or later stages. (TIF)

S2 Fig. Meiosis II is not visibly affected in eggs from *lid* KD females. A—Representative confocal images of eggs in metaphase of meiosis II stained for DNA (blue), alpha-tubulin (red) and ProtA::GFP (green). The tandem of meiotic spindles is shown on the left, the corresponding male nucleus from the same egg is on the right. Bars: 5 µm. B—Quantification of meiosis II phenotypes (normal or abnormal chromosome segregation). (TIF)

S3 Fig. Phenotype of *Sim3A* KD and *rp43* KD eggs/embryos. A—Confocal images of *Sim3A* KD eggs stained for DNA at the indicated stages. The sperm nucleus in the left panel is indicated (arrow). Bar: 10 µm. PB: Polar bodies. B—Confocal images of *rp43* KD early embryos (from ProtA::GFP fathers) stained for DNA and anti-GFP. The sperm nucleus is indicated

(arrows). Bar: 10 μ m. PB: Polar bodies.

(TIF)

S4 Fig. Lid is not directly involved in sperm chromatin remodeling at fertilization. A—Top row: confocal images of stage 10 egg chambers from control (left) and lid KD (right) females stained for DNA (red) and anti-Lid (green). Middle row: detail of a nurse cell nucleus. Bottom row: detail of the oocyte germinal vesicle (oocyte nucleus). Bar: 20 μ m. B—Confocal images of the male pronucleus and the female pronucleus from a control egg in meiosis II stained for DNA and anti-Lid. Bar: 10 μ m. Quantification of Lid positive nuclei is indicated. C—Confocal images of a control (left) and lid KD (right) blastoderm embryo with same staining as in B. Bar: 10 μ m. Quantifications of embryos with a positive/negative nuclear Lid staining are indicated for each genotype.

(JPG)

S5 Fig. Lid and Sin3A control *dhd* expression in female germ cells. A—Principal Component Analysis of Control, *lid* KD and *Sin3A* KD ovarian transcriptomes (two biological replicates for each genotype). B—Volcano plot representations of Differentially-Expressed genes in Control vs *lid* KD (left) and Control vs *Sin3A* KD (right). C—RT-qPCR quantification of *dhd* mRNA levels in ovaries of indicated genotypes. mRNA levels were normalized to *rp49* and shown as relative expression in MTD>+ control. Error bars represent SD (Dunnett's multiple comparisons test to the control MTD>+, *** $P < 0.0001$). D—Western blot analysis of DHD in adult ovaries (left) and 0-30min postfertilization embryos (right). α -tubulin was used as a loading control. E—Western blot analysis of DHD in adult ovaries of indicated genotypes. α -tubulin was used as a loading control.

(TIF)

S6 Fig. Top twelve most downregulated and upregulated genes in *lid* KD and *Sin3A* KD transcriptomes.

(TIF)

S7 Fig. The JmjC domain of Lid is required for normal *dhd* expression. A—Embryo hatching rates from females of indicated genotypes. B—RT-qPCR quantification of *dhd* (left) and *lid* (right) mRNA levels in ovaries of indicated genotypes. mRNA levels were normalized to *rp49* and shown as relative expression in *w¹¹¹⁸* control. Error bars represent SD (Dunnett's multiple comparisons test to the control (** $P < 0.01$; *** $P = 0.0002$). C—Analysis of paternal GFP::Cid expression in late embryos from indicated females (as in Fig 1B).

(TIF)

S8 Fig. *trr* KD does not affect *dhd* expression. A—Confocal images of representative embryos of the indicated genotypes stained for DNA and anti-DHD. The fertilizing sperm nucleus is magnified in insets. Bar: 20 μ m. B—Details of maternal chromosomes (top row) and sperm nucleus (bottom row) from a representative *trr* KD egg stained for ProtA::GFP and histones.

(JPG)

S1 Table. Haploid TRiP genetic screen.

(XLSX)

S2 Table. Differentially expressed genes in *lid* KD and *Sin3A* KD ovaries.

(XLSX)

S3 Table. Quantitative analysis of H3K4me3 differential enrichment in Control vs *lid* KD ovarian ChIP-Seq.

(XLSX)

Acknowledgments

We thank Ferran Azorin, Cayetano Gonzalez, Cristina Molnar and Paulo Navarro-Costa for sharing fly stocks and antibodies. We are grateful to the TRIP projects and Bloomington Drosophila Stock Center for shRNA stocks. The authors would like to thank the members of the IGBMC GenomEast platform for their help. We also thank Raphaëlle Dubruielle for her helpful comments on the manuscript.

Author Contributions

Conceptualization: Daniela Torres-Campana, Shuhei Kimura, Guillermo A. Orsi, Benjamin Loppin.

Data curation: Guillermo A. Orsi, Gérard Benoit.

Formal analysis: Gérard Benoit.

Funding acquisition: Benjamin Loppin.

Investigation: Daniela Torres-Campana, Shuhei Kimura, Béatrice Horard, Benjamin Loppin.

Methodology: Gérard Benoit.

Supervision: Béatrice Horard, Benjamin Loppin.

Validation: Daniela Torres-Campana, Guillermo A. Orsi, Béatrice Horard, Gérard Benoit, Benjamin Loppin.

Writing – original draft: Benjamin Loppin.

Writing – review & editing: Daniela Torres-Campana, Guillermo A. Orsi, Béatrice Horard, Gérard Benoit, Benjamin Loppin.

References

1. Loppin B, Dubruielle R, Horard B (2015) The intimate genetics of *Drosophila* fertilization. *Open Biol* 5: 150076. <https://doi.org/10.1098/rsob.150076> PMID: 26246493
2. Hamm DC, Harrison MM (2018) Regulatory principles governing the maternal-to-zygotic transition: insights from *Drosophila melanogaster*. *Open Biol* 8: 180183. <https://doi.org/10.1098/rsob.180183> PMID: 30977698
3. Spradling AC (1993) Developmental genetics of oogenesis. *The Development of Drosophila melanogaster*—Cold Spring Harbor Laboratory Press. 1–69.
4. Eirin-López JM, Ausió J (2009) Origin and evolution of chromosomal sperm proteins. *Bioessays* 31: 1062–1070. <https://doi.org/10.1002/bies.200900050> PMID: 19708021
5. Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R (2014) Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta* 1839: 155–168. <https://doi.org/10.1016/j.bbaggm.2013.08.004> PMID: 24091090
6. Tirmarche S, Kimura S, Dubruielle R, Horard B, Loppin B (2016) Unlocking sperm chromatin at fertilization requires a dedicated egg thiorodexin in *Drosophila*. *Nature Commun* 7: 13539. <https://doi.org/10.1038/ncomms13539>
7. Petrova B, Liu K, Tian C, Kitaoka M, Freinkman E, Yang J, et al. (2018) Dynamic redox balance directs the oocyte-to-embryo transit via developmentally controlled reactive cysteine changes. *Proc Natl Acad Sci USA* 115: E7978–E7986. <https://doi.org/10.1073/pnas.1807918115> PMID: 30082411
8. Horard B, Sapey-Triomphe L, Bonnefoy E, Loppin B (2018) ASF1 is required to load histones on the HIRA complex in preparation of paternal chromatin assembly at fertilization. *Epigenetics Chromatin* 11: 19. <https://doi.org/10.1186/s13072-018-0189-x> PMID: 29751347
9. Ni J-Q, Zhou R, Czech B, Liu L-P, Holdersbaum L, Yang-Zhou D, et al. (2011) A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat Meth* 8: 405–407. <https://doi.org/10.1038/nmeth.1592>
10. Delabaere L, Orsi GA, Sapey-Triomphe L, Horard B, Couble P, Loppin B (2014) The Spartan Ortholog Maternal Haploid Is Required for Paternal Chromosome Integrity in the *Drosophila* Zygote. *Curr Biol* 24: 2281–2287. <https://doi.org/10.1016/j.cub.2014.08.010> PMID: 25242033

11. Gildea JJ, Lopez R, Shearn A (2000) A screen for new trithorax group genes identified little imaginal discs, the *Drosophila melanogaster* homologue of human retinoblastoma binding protein 2. *Genetics* 156: 645–663. PMID: 11014813
12. Secombe J, Li L, Carlos L, Eisenman RN (2007) The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Genes & Dev* 21: 537–551. <https://doi.org/10.1101/gad.1523007>
13. Lloret-Linares M, Carr CM, Vaquero A, de Olano N, Azorin F (2008) Characterization of *Drosophila melanogaster* JmjC+N histone demethylases. *Nucleic Acids Res*. 36: 2852–2863. <https://doi.org/10.1093/nar/gkn098> PMID: 18375980
14. Lloret-Linares M, Perez-Lluch S, Rossell D, Moran T, Ponsa-Cobas J, Auer H, et al. (2012) dKDM5/LID regulates H3K4me3 dynamics at the transcription-start site (TSS) of actively transcribed developmental genes. *Nucleic Acids Res*. 40: 9493–9505. <https://doi.org/10.1093/nar/gks773> PMID: 22904080
15. Grzenda A, Lombark G, Zhang J-S, Urrutia R (2009) Sin3: Master scaffold and transcriptional corepressor. *Biochim Biophys Acta* 1789: 443–450. <https://doi.org/10.1016/j.bbagr.2009.05.007> PMID: 19505602
16. Spain MM, Caruso JA, Swaminathan A, Pile LA (2010) *Drosophila* SIN3 Isoforms Interact with Distinct Proteins and Have Unique Biological Functions. *J Biol Chem* 285: 27457–27467. <https://doi.org/10.1074/jbc.M110.130245> PMID: 20566628
17. Gajan A, Barnes VL, Liu M, Saha N, Pile LA (2016) The histone demethylase dKDM5/LID interacts with the SIN3 histone deacetylase complex and shares functional similarities with SIN3. *Epigenetics Chromatin* 9:4 <https://doi.org/10.1186/s13072-016-0053-9> PMID: 26848313
18. Liu M, Pile LA (2017) The Transcriptional Corepressor SIN3 Directly Regulates Genes Involved in Methionine Catabolism and Affects Histone Methylation, Linking Epigenetics and Metabolism. *J Biol Chem* 292: 1970–1976. <https://doi.org/10.1074/jbc.M116.749754> PMID: 28028175
19. Barnes VL, Bhat A, Unnikrishnan A, Heydari AR, Arking R, Pile LA (2014) SIN3 is critical for stress resistance and modulates adult lifespan. *Aging* 6: 645–660. <https://doi.org/10.18632/aging.100684> PMID: 25133314
20. Moshkin YM, Kan TW, Goodfellow H, Bezstarosti K, Maeda RK, Pilyugin M, et al. (2009) Histone Chaperones ASF1 and NAP1 Differentially Modulate Removal of Active Histone Marks by LID-RPD3 Complexes during NOTCH Silencing. *Mol Cell* 35: 782–793. <https://doi.org/10.1016/j.molcel.2009.07.020> PMID: 19782028
21. Lee N, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2009) The H3K4 Demethylase Lid Associates with and Inhibits Histone Deacetylase Rpd3. *Mol Cell Biol* 29: 1401–1410. <https://doi.org/10.1128/MCB.01643-08> PMID: 19114561
22. Kronja J, Yuan B, Eichhorn SW, Dzeyk K, Kringsvold J, Bartel DP, et al. (2014) Widespread Changes in the Posttranscriptional Landscape at the *Drosophila* Oocyte-to-Embryo Transition. *Cell Reports* 7: 1495–1508. <https://doi.org/10.1016/j.celrep.2014.05.002> PMID: 24882012
23. Navarro-Costa P, McCarthy A, Prudêncio P, Greer C, Guilgur LG, Becker JORD, et al. (2016) Early programming of the oocyte epigenome temporally controls late prophase I transcription and chromatin remodelling. *Nature Commun* 7: 1–15. <https://doi.org/10.1038/ncomms12331>
24. Manier MK, Belote JM, Berben KS, Novikov D, Stuart WT, Pitnick S (2010) Resolving Mechanisms of Competitive Fertilization Success in *Drosophila melanogaster*. *Science* 328: 354–357. <https://doi.org/10.1126/science.1187096> PMID: 20299550
25. Zhaunova L, Ohkura H, Breuer M (2016) Kdm5/Lid Regulates Chromosome Architecture in Meiotic Prophase I Independently of Its Histone Demethylase Activity. *PLoS Genet*. 12: e1006241. <https://doi.org/10.1371/journal.pgen.1006241> PMID: 27494704
26. Dreton C, Belalcazar HM, Secombe J. The Histone Demethylase KDM5 Is Essential for Larval Growth in *Drosophila* (2018) *Genetics* 209: 773–787. <https://doi.org/10.1534/genetics.118.301004> PMID: 29764801
27. Liu X, Secombe J (2015) The Histone Demethylase KDM5 Activates Gene Expression by Recognizing Chromatin Context through Its PHD Reader Motif. *Cell Reports* 13: 2219–2231. <https://doi.org/10.1016/j.celrep.2015.11.007> PMID: 26673323
28. Emelyanov AV, Fyodorov DV (2016) Thioredoxin-dependent disulfide bond reduction is required for protamine eviction from sperm chromatin. *Genes & Dev* 30: 2651–2656. <https://doi.org/10.1101/gad.290916.116>
29. Sandler JE, Stathopoulos A (2016) Quantitative Single-Embryo Profile of *Drosophila* Genome Activation and the Dorsal-Ventral Patterning Network. *Genetics* 202: 1575–1584. <https://doi.org/10.1534/genetics.116.186783> PMID: 26896327

30. Zamurrad S, Hatch HAM, Drelon C, Belalcazar HM, Secombe J (2018) A *Drosophila* Model of Intellectual Disability Caused by Mutations in the Histone Demethylase KDM5. *Cell Reports* 22: 2359–2369. <https://doi.org/10.1016/j.celrep.2018.02.018> PMID: 29490272
31. Li L, Greer C, Eisenman RN, Secombe J (2010) Essential functions of the histone demethylase lid. *PLoS Genet* 6: e1001221. <https://doi.org/10.1371/journal.pgen.1001221> PMID: 21124823
32. Lussi YC, Mariani L, Friis C, Pelltonen J, Myers TR, Krag C, et al. (2016) Impaired removal of H3K4 methylation affects cell fate determination and gene transcription. *Development* 143: 3751–3762. <https://doi.org/10.1242/dev.139139> PMID: 27578789
33. Kidder B., Hu G, Zhao K (2014) KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biol* 15: R32. <https://doi.org/10.1186/gb-2014-15-2-r32> PMID: 24495580
34. Outchkourov NS, Muñio JM, Kaufmann K, van Ijcken WFJ, Koerkamp MJG, van Leenen D, et al. (2013) Balancing of Histone H3K4 Methylation States by the Kdm5c/SMCX Histone Demethylase Modulates Promoter and Enhancer Function. *Cell Reports* 3: 1071–1079. <https://doi.org/10.1016/j.celrep.2013.02.030> PMID: 23545502
35. Freeman M, Nüsslein-Volhard C, Glover DM (1986) The dissociation of nuclear and centrosomal division in *gnu*, a mutation causing giant nuclei in *Drosophila*. *Cell* 46: 457–468. [https://doi.org/10.1016/0092-8674\(86\)90666-5](https://doi.org/10.1016/0092-8674(86)90666-5) PMID: 3089628
36. Herz HM, Mohan M, Garruss AS, Liang K, Takahashi YH, Mickey K, et al. (2012) Enhancer-associated H3K4 monomethylation by Trithorax-related, the *Drosophila* homolog of mammalian Mll3/Mll4. *Genes & Dev* 26: 2604–2620. <https://doi.org/10.1101/gad.201327.112>
37. Prudêncio P, Guilgur LG, Sobral J, Becker JD, Martinho RG, Navarro-Costa P (2018) The Trithorax group protein dMLL3/4 instructs the assembly of the zygotic genome at fertilization. *EMBO Rep* 19: e45728. <https://doi.org/10.15252/embr.201845728> PMID: 30037897
38. Schuth M, Lehner CF, Heidmann S (2007) Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr Biol* 17: 237–243. <https://doi.org/10.1016/j.cub.2006.11.051> PMID: 17222555
39. Salz HK, Flickinger TW, Mitterdorf E, Pellicena-Palle A, Petschek JP, Brown Abrecht E (1994) The *Drosophila* Maternal Effect Locus deadhead Encodes a Thioredoxin Homolog Required for Female Meiosis and Early Embryonic Development. *Genetics* 136: 1075–1086. PMID: 7516301
40. Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2^{-ΔΔCT} Method. *Methods* 25: 402–408. <https://doi.org/10.1006/meth.2001.1262> PMID: 11846609
41. Grentzinger T, Armenise C, Brun C, Mugat B, Serrano V, Pelisson A, et al. (2012) piRNA-mediated transgenerational inheritance of an acquired trait. *Genome Res* 22: 1877–1888. <https://doi.org/10.1101/gr.136614.111> PMID: 22555593

Supplemental data:

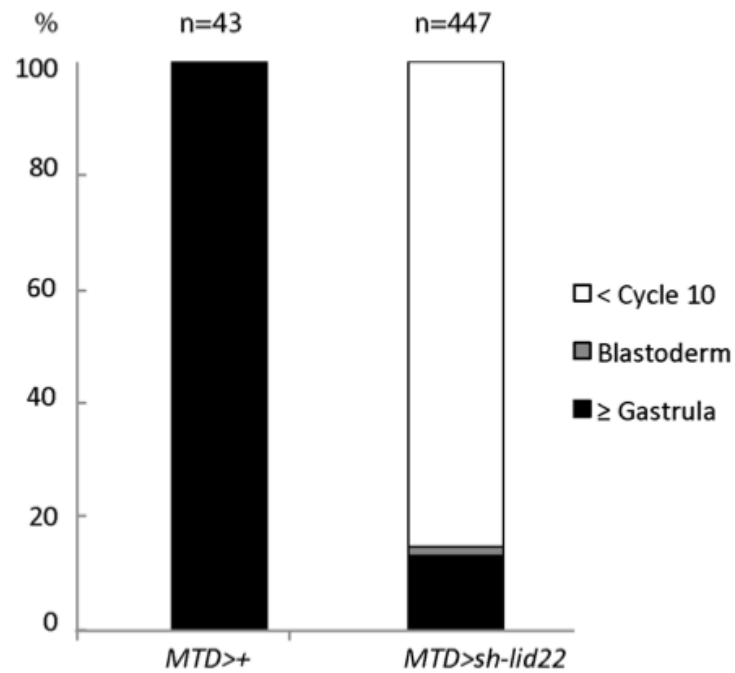


Fig S1. Developmental defects of *lid* KD embryos.

Embryos were collected for four hours and aged for another four hours at 25°C before DAPI staining and examination in fluorescent microscopy. More than 85% of *lid* KD embryos arrest development before the blastoderm stage. In contrast, 100% of control embryos had reached gastrula or later stages.

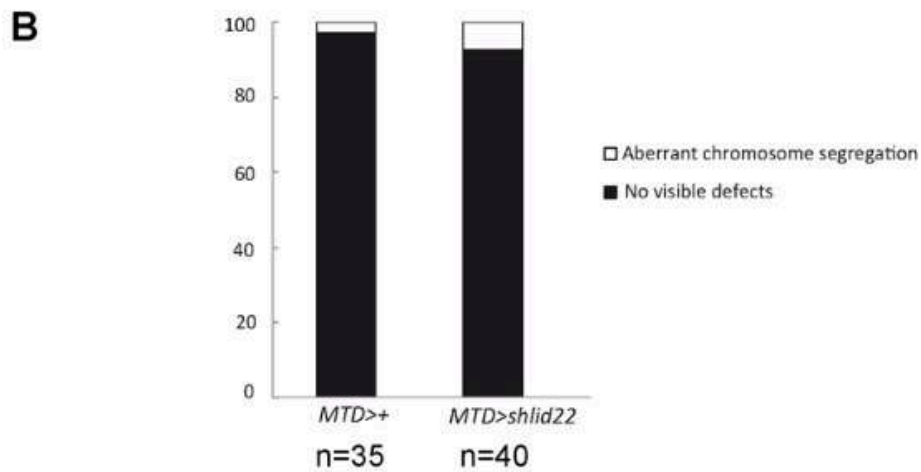
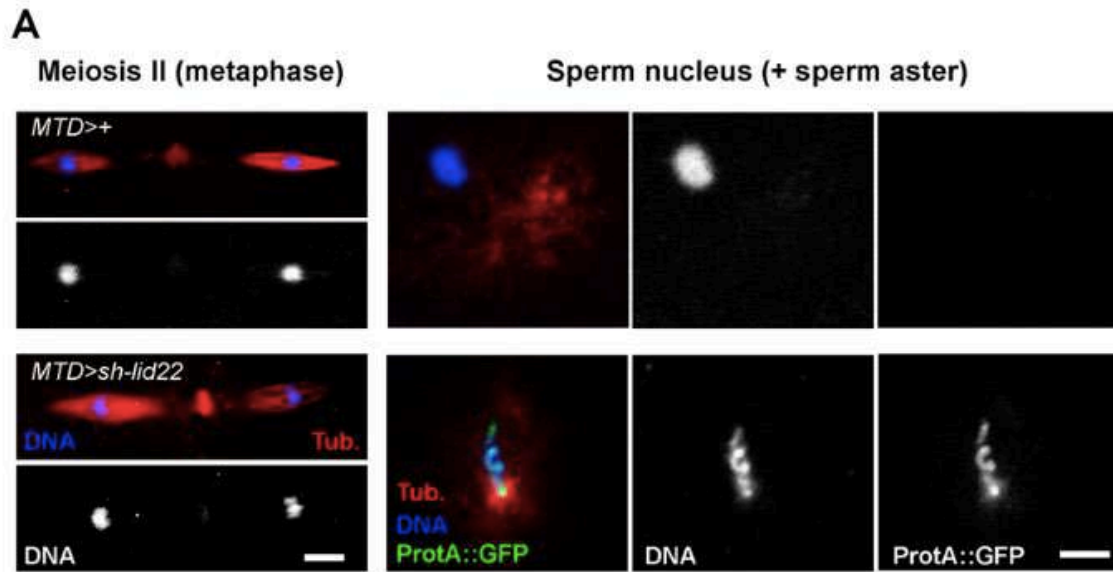


Fig S2. Meiosis II is not visibly affected in eggs from *lid* KD females.

A-Representative confocal images of eggs in metaphase of meiosis II stained for DNA (blue), alpha-tubulin (red) and ProtA::GFP (green). The tandem of meiotic spindles is shown on the left, the corresponding male nucleus from the same egg is on the right. Bars: 5 μ m. B-Quantification of meiosis II phenotypes (normal or abnormal chromosome segregation).

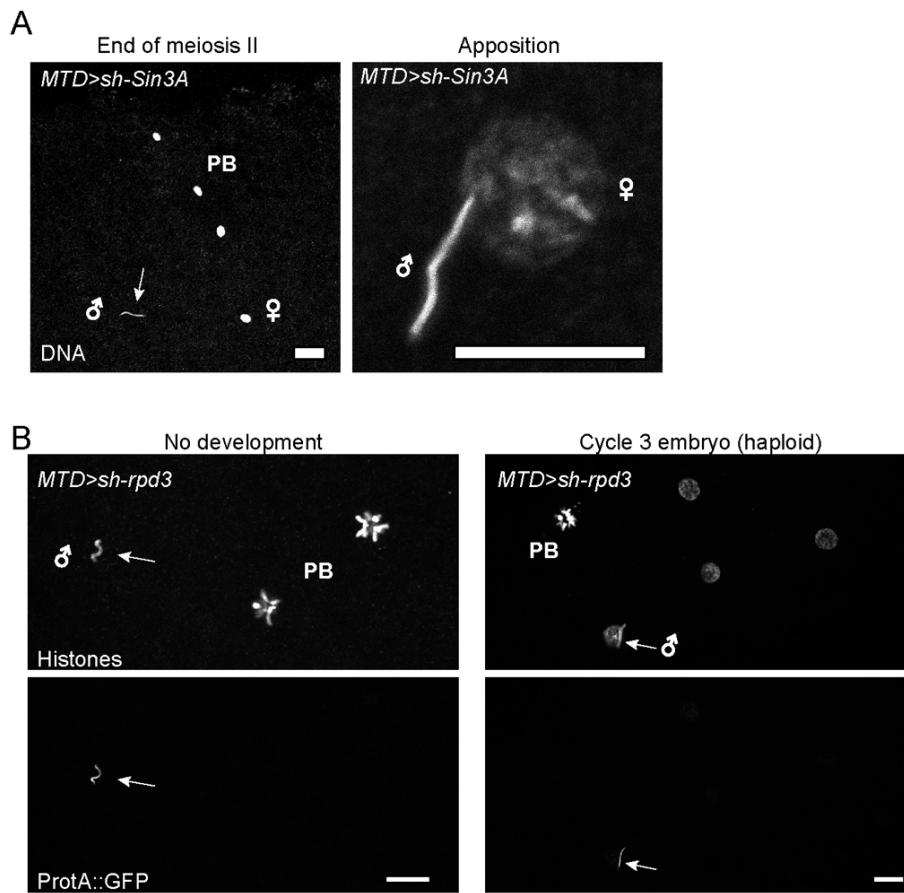


Fig S3. Phenotype of *Sin3A* KD and *rpd3* KD eggs/embryos.

A-Confocal images of *Sin3A* KD eggs stained for DNA at the indicated stages. The sperm nucleus in the left panel is indicated (arrow). Bar: 10 μ m. PB: Polar bodies. B-Confocal images of *rpd3* KD early embryos (from ProtA::GFP fathers) stained for DNA and anti-GFP. The sperm nucleus is indicated (arrows). Bar: 10 μ m. PB: Polar bodies.

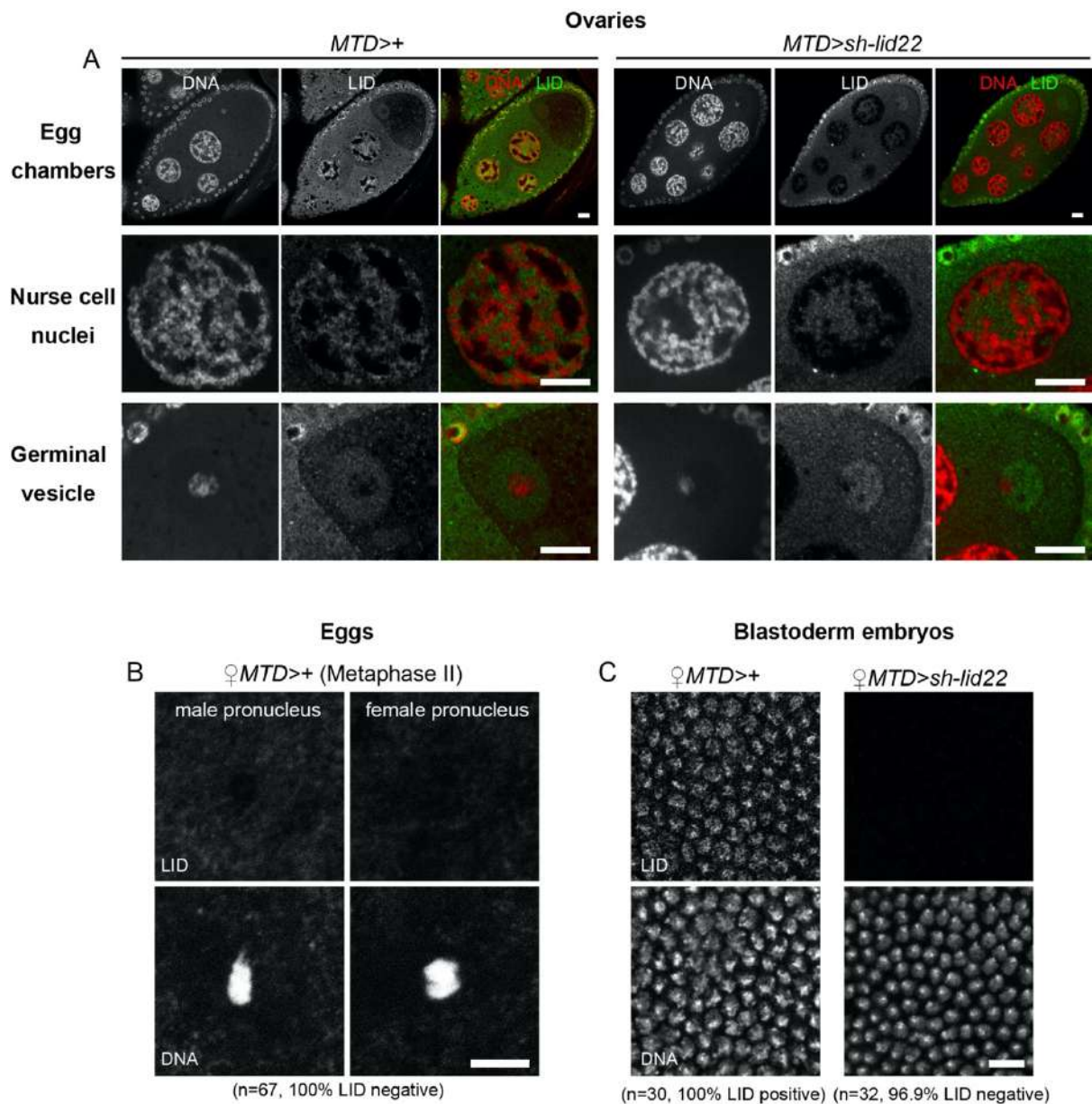


Fig S4. Lid is not directly involved in sperm chromatin remodeling at fertilization.

A-Top row: confocal images of stage 10 egg chambers from control (left) and lid KD (right) females stained for DNA (red) and anti-Lid (green). Middle row: detail of a nurse cell nucleus. Bottom row: detail of the oocyte germinal vesicle (oocyte nucleus). Bar: 20 μ m. B-Confocal images of the male pronucleus and the female pronucleus from a control egg in meiosis II stained for DNA and anti-Lid. Bar: 10 μ m. Quantification of Lid positive nuclei is indicated. C-Confocal images of a control (left) and lid KD (right) blastoderm embryo with same staining as in B. Bar: 10 μ m. Quantifications of embryos with a positive/negative nuclear Lid staining are indicated for each genotype.

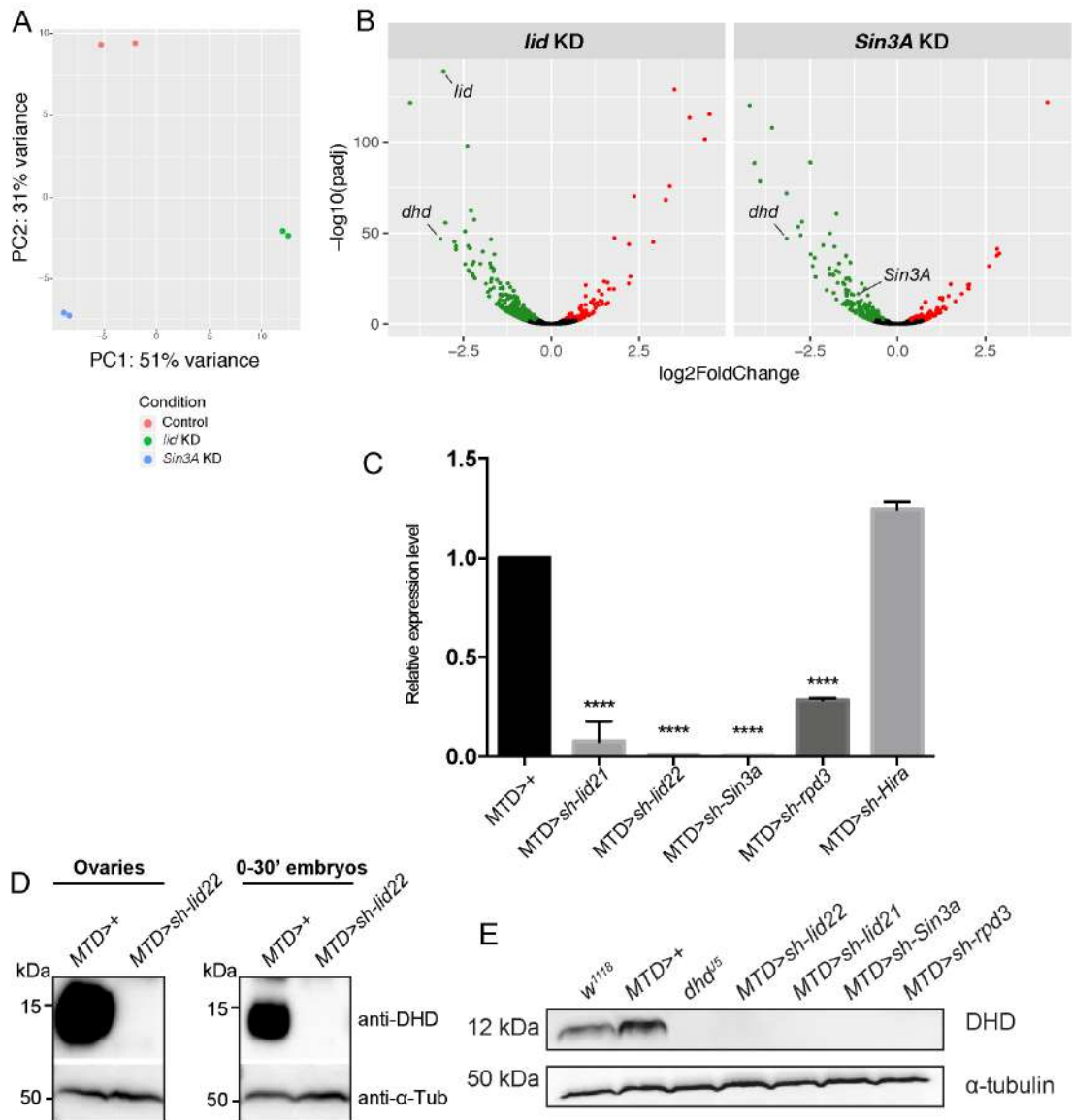


Fig S5. Lid and Sin3A control *dhd* expression in female germ cells.

A-Principal Component Analysis of Control, *lid* KD and *Sin3A* KD ovarian transcriptomes (two biological replicates for each genotype). B-Volcano plot representations of Differentially-Expressed genes in Control vs *lid* KD (left) and Control vs *Sin3A* KD (right). C-RT-qPCR quantification of *dhd* mRNA levels in ovaries of indicated genotypes. mRNA levels were normalized to rp49 and shown as relative expression in MTD>+ control. Error bars represent SD (Dunnett's multiple comparisons test to the control MTD>+, **** P < 0.0001). D—Western blot analysis of DHD in adult ovaries (left) and 0-30min postfertilization embryos (right). α -tubulin was used as a loading control. E—Western blot analysis of DHD in adult ovaries of indicated genotypes. α -tubulin was used as a loading control.

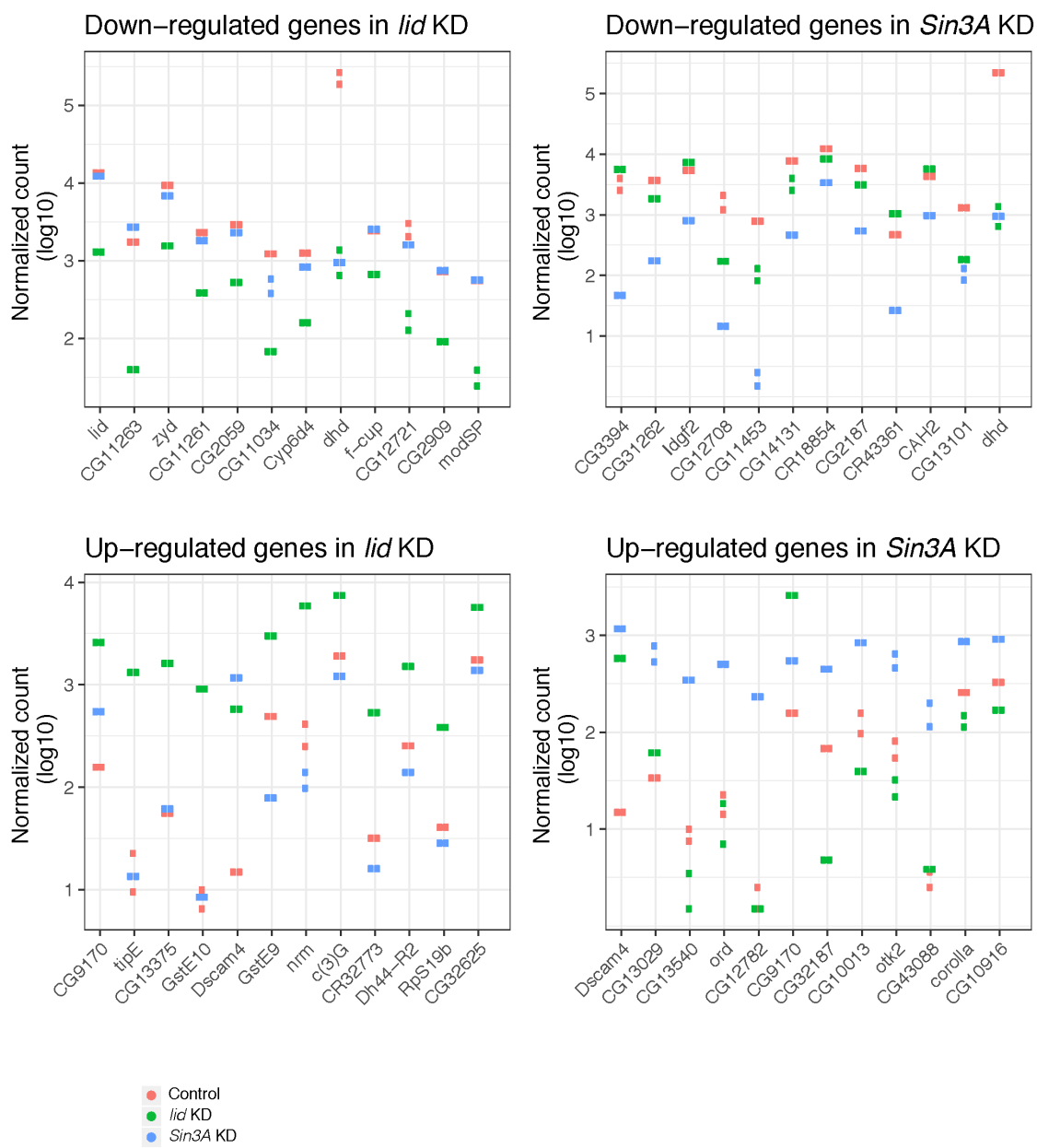
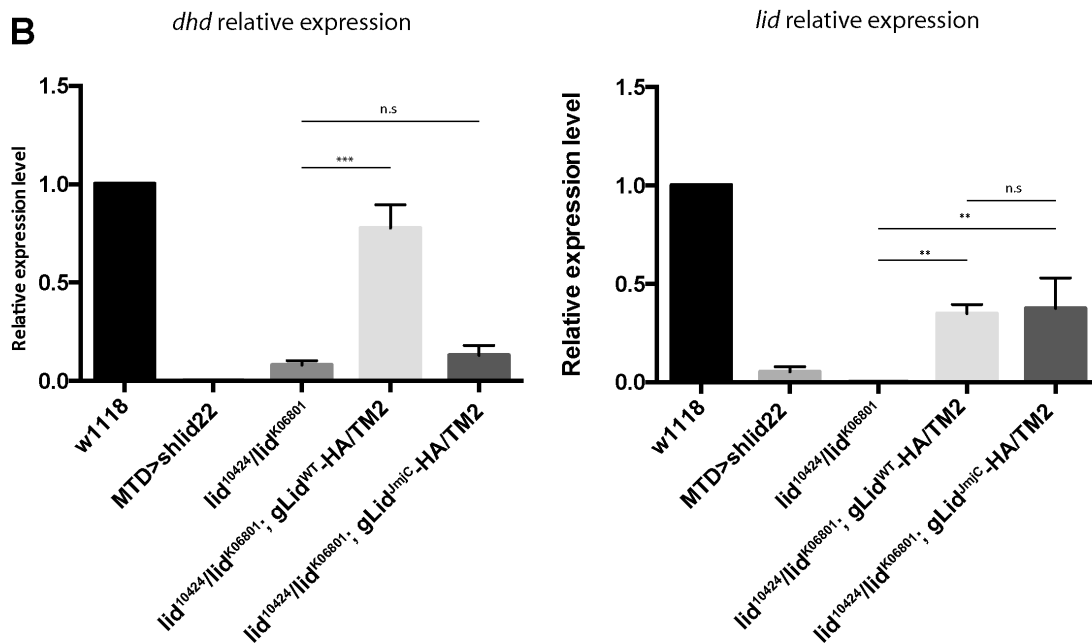


Fig S6. Top twelve most downregulated and upregulated genes in *lid* KD and *Sin3A* KD transcriptomes.

A

Female Genotype	No. eggs	Hatching rate
<i>w¹¹¹⁸</i>	330	97.2
<i>lid¹⁰⁴²⁴/lid^{K06801}; gLid^{WT}-HA/TM2</i>	921	63.4
<i>lid¹⁰⁴²⁴/lid^{K06801}; gLid^{JmjC}-HA/TM2</i>	117	16.2

B**C**

Female Genotype	No. eggs	Gastrula or beyond	GFP neg.	GFP pos.
<i>lid¹⁰⁴²⁴/lid^{K06801}; gLid^{WT}-HA/TM2</i>	53	34	0	34
<i>lid¹⁰⁴²⁴/lid^{K06801}; gLid^{JmjC}-HA/TM2</i>	14	6	3	3

Fig S7. The JmjC domain of Lid is required for normal *dhd* expression.

A-Embryo hatching rates from females of indicated genotypes. B-RT-qPCR quantification of *dhd* (left) and *lid* (right) mRNA levels in ovaries of indicated genotypes. mRNA levels were normalized to rp49 and shown as relative expression in *w¹¹¹⁸* control. Error bars represent SD (Dunnett's multiple comparisons test to the control (** P < 0.01; *** P = 0.0002). C-Analysis of paternal GFP::Cid expression in late embryos from indicated females (as in Fig 1B).

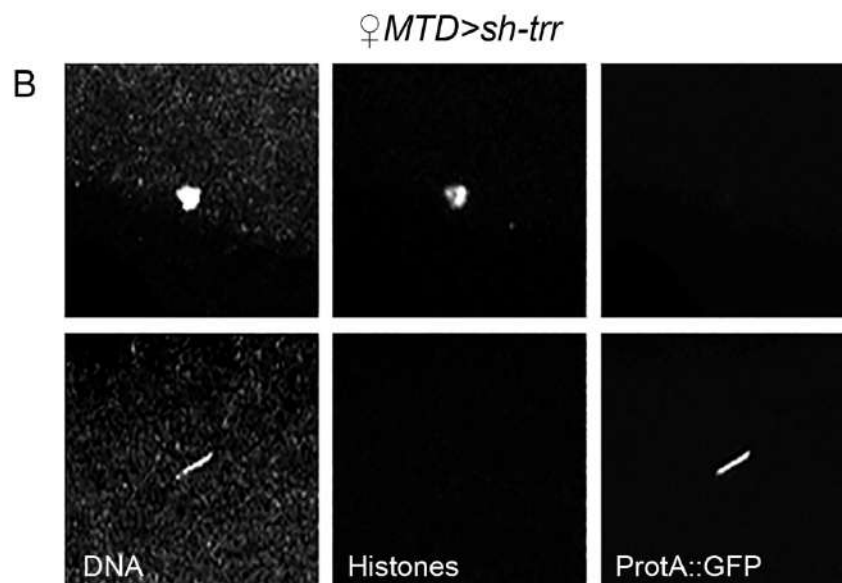
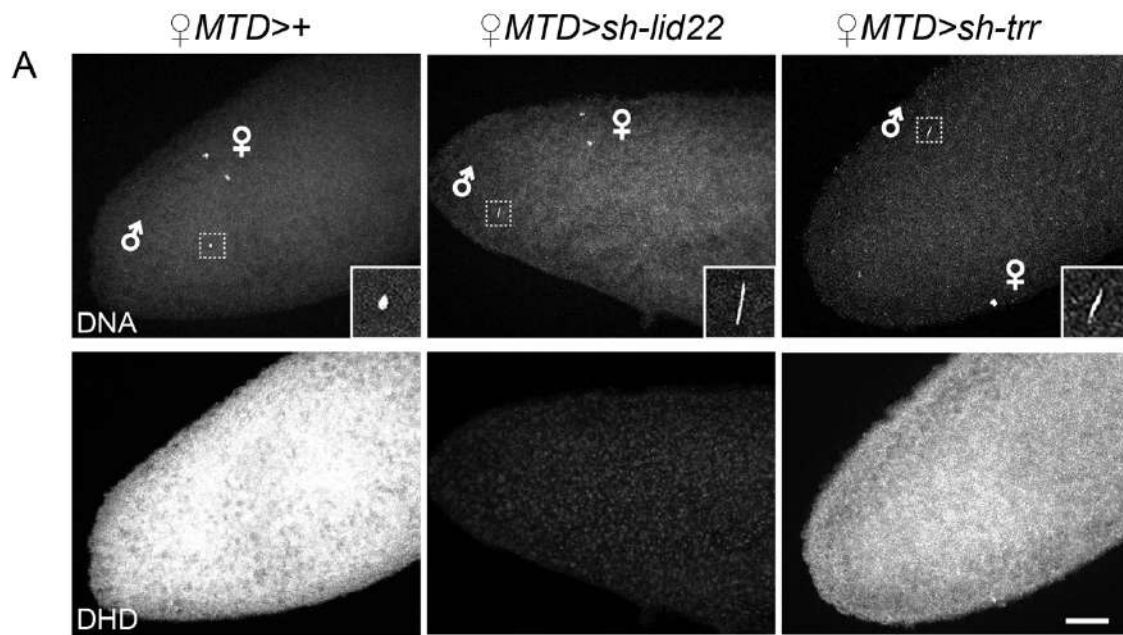


Fig S8. *trr* KD does not affect *dhd* expression.

A-Confocal images of representative embryos of the indicated genotypes stained for DNA and anti-DHD. The fertilizing sperm nucleus is magnified in insets. Bar: 20 μ m. B-Details of maternal chromosomes (top row) and sperm nucleus (bottom row) from a representative *trr* KD egg stained for ProtA::GFP and histones.

S1 Table - Haploid TRIP screen							
	Fertile						"dhd-like": needle-shaped sperm nucleus, as in <i>dhdJ5</i> mutant.
	Sterile, class 1: No egg laying.						"ssm-like": small, round male nucleus, as in <i>Hira(ssm)</i> mutant.
	Sterile, class 2: Few abnormal eggs, no hatching.						
	Sterile, class 3: Eggs do not hatch (HR<5%) and remain white.						
	Sterile, class 4: Eggs do not hatch (HR<5%), some turn brown.						
BDSC#	TRIP#	CG#	Gene Symbol	Vector	Female fertility	GFP::Cid expression	Sperm nuclear phenotype
35334	GL00243	CG9638	Ada2b	pVALIUM22	Fertile	Not tested	Not tested
35232	GL00117	CG40300	AGO3	pVALIUM22	Fertile	Not tested	Not tested
36803	GL01012	CG8887	ash1	pVALIUM22	Fertile	Not tested	Not tested
32503	HMS00507	CG4303	Bap60	pVALIUM20	Fertile	Not tested	Not tested
36781	GL00127	CG6046	Bin1	pVALIUM22	Fertile	Not tested	Not tested
43174	GL01516	CG2009	bip2	pVALIUM22	Fertile	Not tested	Not tested
42502	HMJ02067	CG1845	Br140	pVALIUM20	Fertile	Not tested	Not tested
42658	HMS02494	CG14514	Brd8	pVALIUM20	Fertile	Not tested	Not tested
35623	GL00467	CG31256	Brf	pVALIUM22	Fertile	Not tested	Not tested
35211	GL00090	CG5942	brm	pVALIUM22	Fertile	Not tested	Not tested
35163	GL00031	CG7597	Cdk12	pVALIUM22	Fertile	Not tested	Not tested
33701	HMS00578	CG10133	CG10133	pVALIUM20	Fertile	Not tested	Not tested
43137	GL01475	CG10445	CG10445	pVALIUM22	Fertile	Not tested	Not tested
41847	GL01275	CG11329	CG11329	pVALIUM22	Fertile	Not tested	Not tested
36834	GL01075	CG12877	CG12877	pVALIUM22	Fertile	Not tested	Not tested
33731	HMS00614	CG13900	CG13900	pVALIUM20	Fertile	Not tested	Not tested
36914	GL01123	CG14352	CG14352	pVALIUM22	Fertile	Not tested	Not tested
36812	GL01030	CG1815	CG1815	pVALIUM22	Fertile	Not tested	Not tested
42488	HMJ02050	CG2051	CG2051	pVALIUM20	Fertile	Not tested	Not tested
35803	GL00427	CG32772	CG32772	pVALIUM22	Fertile	Not tested	Not tested
35475	GL00403	CG33694,CG33695	cana,CG33695	pVALIUM22	Fertile	Not tested	Not tested
38211	GL00652	CG3509	CG3509	pVALIUM20	Fertile	Not tested	Not tested
34077	HMS00584	CG4049	CG4049	pVALIUM20	Fertile	Not tested	Not tested
32973	HMS00770	CG4078	CG4078	pVALIUM20	Fertile	Not tested	Not tested
41976	HMS02374	CG4203	CG4203	pVALIUM20	Fertile	Not tested	Not tested
34606	HMS00577	CG43320	CG43320	pVALIUM20	Fertile	Not tested	Not tested
42533	HMJ02100	CG4400	CG4400	pVALIUM20	Fertile	Not tested	Not tested
33696	HMS00568	CG4747	CG4747	pVALIUM20	Fertile	Not tested	Not tested
42833	HMS02525	CG4854	CG4854	pVALIUM20	Fertile	Not tested	Not tested
32441	HMS00439	CG6418	CG6418	pVALIUM20	Fertile	Not tested	Not tested
42552	HMJ02125	CG6506	CG6506	pVALIUM20	Fertile	Not tested	Not tested
41824	GL01252	CG6937	CG6937	pVALIUM22	Fertile	Not tested	Not tested
32485	HMS00488	CG7200	CG7200	pVALIUM20	Fertile	Not tested	Not tested
35229	GL00113	CG7878	CG7878	pVALIUM22	Fertile	Not tested	Not tested
35593	GL00430	CG7911	CG7911	pVALIUM22	Fertile	Not tested	Not tested
41874	GL01305	CG8223	CG8223	pVALIUM22	Fertile	Not tested	Not tested
33361	HMS00233	CG8289	CG8289	pVALIUM20	Fertile	Not tested	Not tested
41649	GL01231	CG9272	CG9272	pVALIUM22	Fertile	Not tested	Not tested
42766	GL00722	CG9418	CG9418	pVALIUM22	Fertile	Not tested	Not tested
35807	GL00455	CG9576	CG9576	pVALIUM22	Fertile	Not tested	Not tested
35652	GLV21017	CG10890,CG13399	mus201,Chrac-14	pVALIUM21	Fertile	Not tested	Not tested
36084	GL00503	CG10712	Chro	pVALIUM22	Fertile	Not tested	Not tested
35387	GL00306	CG2714	cm	pVALIUM22	Fertile	Not tested	Not tested
35354	GL00266	CG8591	CTCF	pVALIUM22	Fertile	Not tested	Not tested
34732	HMS01212	CG7405	CycH	pVALIUM20	Fertile	Not tested	Not tested
35168	GL00037	CG6292	CycT	pVALIUM22	Fertile	Not tested	Not tested
32451	HMS00450	CG7098	dik	pVALIUM20	Fertile	Not tested	Not tested
32397	HMS00391	CG42799	dikar	pVALIUM20	Fertile	Not tested	Not tested
33410	HMS00290	CG7143	DNApol-eta	pVALIUM20	Fertile	Not tested	Not tested
36648	GL00608	CG7602	DNApol-iota	pVALIUM22	Fertile	Not tested	Not tested
35353	GL00265	CG32346	E(bx)	pVALIUM22	Fertile	Not tested	Not tested
32346	HMS00337	CG12238	e(y)3	pVALIUM20	Fertile	Not tested	Not tested
33904	HMS00846	CG12756	Eaf6	pVALIUM20	Fertile	Not tested	Not tested
37017	HMS01255	CG6755	EloA	pVALIUM20	Fertile	Not tested	Not tested
36906	GL01110	CG10215	Ercc1	pVALIUM22	Fertile	Not tested	Not tested
43285	HMS02657	CG1474	Es2	pVALIUM20	Fertile	Not tested	Not tested
33891	HMS00829	CG5899	Etl1	pVALIUM20	Fertile	Not tested	Not tested
35805	GL00442	CG4274	fzy	pVALIUM22	Fertile	Not tested	Not tested
42787	GL01157	CG10670	Gen	pVALIUM22	Fertile	Not tested	Not tested
41893	GL01325	CG42803	gpp	pVALIUM20	Fertile	Not tested	Not tested
33705	HMS00582	CG8887	ash1	pVALIUM20	Fertile	Not tested	Not tested
34817	HMS00127	CG2995	G9a	pVALIUM20	Fertile	Not tested	Not tested
34727	HMS01206	CG6662	GstO1	pVALIUM20	Fertile	Not tested	Not tested
34033	HMS01004	CG4976	NSD	pVALIUM20	Fertile	Not tested	Not tested
33703	HMS00580	CG8651	lrx	pVALIUM20	Fertile	Not tested	Not tested
40931	HMS02179	CG40351	Set1	pVALIUM20	Fertile	Not tested	Not tested
35276	GL00176	CG40080	Haspin	pVALIUM22	Fertile	Not tested	Not tested
42620	HMS02455	CG5825	His3.3A	pVALIUM20	Fertile	Not tested	Not tested
36639	GL00599	CG13363	Hmt4-20	pVALIUM22	Fertile	Not tested	Not tested
32401	HMS00396	CG7041	HP1b	pVALIUM20	Fertile	Not tested	Not tested
33962	HMS00919	CG6990	HP1c	pVALIUM20	Fertile	Not tested	Not tested
37473	GL00616	CG31212	Ino80	pVALIUM22	Fertile	Not tested	Not tested
35819	GL00286	CG5247	Irbp	pVALIUM22	Fertile	Not tested	Not tested
37545	GLV21092	CG13201	ix	pVALIUM21	Fertile	Not tested	Not tested
40855	HMS02022	CG3654	Jarid2	pVALIUM20	Fertile	Not tested	Not tested
35438	GL00363	CG4029	jumu	pVALIUM22	Fertile	Not tested	Not tested
33699	HMS00574	CG11033	Kdm2	pVALIUM20	Fertile	Not tested	Not tested
34629	HMS01304	CG15835	Kdm4A	pVALIUM20	Fertile	Not tested	Not tested
34912	HMS01260	CG10080	mahj	pVALIUM20	Fertile	Not tested	Not tested
33709	HMS00587	CG3753	Marcal1	pVALIUM20	Fertile	Not tested	Not tested
34012	HMS00978	CG5465	MED16	pVALIUM20	Fertile	Not tested	Not tested
42634	HMS02470	CG14802	MED18	pVALIUM20	Fertile	Not tested	Not tested
33710	HMS00588	CG5546	MED19	pVALIUM20	Fertile	Not tested	Not tested
34577	HMS01051	CG18780	MED20	pVALIUM20	Fertile	Not tested	Not tested
34573	HMS01047	CG3034	MED22	pVALIUM20	Fertile	Not tested	Not tested

33755	HMS01097	CG7999	MED24	pVALIUM20	Fertile	Not tested	Not tested
42501	HMJ02066	CG12254	MED25	pVALIUM20	Fertile	Not tested	Not tested
32459	HMS00458	CG5121	MED28	pVALIUM20	Fertile	Not tested	Not tested
34697	HMS01176	CG8609	MED4	pVALIUM20	Fertile	Not tested	Not tested
35371	GL00284	CG4252	mer-41	pVALIUM22	Fertile	Not tested	Not tested
36605	GL00565	CG11301	Mes4	pVALIUM22	Fertile	Not tested	Not tested
36870	GL01011	CG3025	mof	pVALIUM22	Fertile	Not tested	Not tested
35241	GL00128	CG6363	MRG15	pVALIUM22	Fertile	Not tested	Not tested
35737	HMS01479	CG7003	Msh6	pVALIUM20	Fertile	Not tested	Not tested
35345	GL00256	CG10385	msl-1	pVALIUM22	Fertile	Not tested	Not tested
35272	GL00170	CG8631	msl-3	pVALIUM22	Fertile	Not tested	Not tested
42906	HMS02599	CG10692	Mt2	pVALIUM20	Fertile	Not tested	Not tested
34905	HMS01251	CG2244	MTA1-like	pVALIUM20	Fertile	Not tested	Not tested
32941	HMS00735	CG8274	Mtor	pVALIUM20	Fertile	Not tested	Not tested
35430	GL00354	CG18543	mtm	pVALIUM22	Fertile	Not tested	Not tested
35732	GLC01346	CG4644	mtRNApol	pVALIUM22	Fertile	Not tested	Not tested
43244	GLC01431	CG11156	mus101	pVALIUM22	Fertile	Not tested	Not tested
32446	HMS00444	CG12124	mxc	pVALIUM20	Fertile	Not tested	Not tested
34741	HMS01221	CG42279	Nedd4	pVALIUM20	Fertile	Not tested	Not tested
36682	HMS01570	CG15319	nej	pVALIUM20	Fertile	Not tested	Not tested
35328	GL00235	CG17256	Nek2	pVALIUM22	Fertile	Not tested	Not tested
32897	HMS00686	CG5874	Nelf-A	pVALIUM20	Fertile	Not tested	Not tested
42547	HMJ02119	CG32721	NELF-B	pVALIUM20	Fertile	Not tested	Not tested
32835	HMS00525	CG5994	Nelf-E	pVALIUM20	Fertile	Not tested	Not tested
35595	GL00432	CG33554	Nipped-A	pVALIUM22	Fertile	Not tested	Not tested
32837	HMS00527	CG4453	Nup153	pVALIUM20	Fertile	Not tested	Not tested
32391	HMS00385	CG4738	Nup160	pVALIUM20	Fertile	Not tested	Not tested
33707	HMS00585	CG3736	okr	pVALIUM20	Fertile	Not tested	Not tested
35447	GL00372	CG7467	osa	pVALIUM22	Fertile	Not tested	Not tested
43263	GLC01452	CG11237	Oseg6	pVALIUM22	Fertile	Not tested	Not tested
39009	HMS01926	CG5581	Ote	pVALIUM20	Fertile	Not tested	Not tested
35792	GL00229	CG40411	Parp	pVALIUM22	Fertile	Not tested	Not tested
33972	HMS00929	CG1800	pasha	pVALIUM20	Fertile	Not tested	Not tested
34667	HMS01144	CG5208	Patr-1	pVALIUM20	Fertile	Not tested	Not tested
36070	GL00488	CG32443	Pc	pVALIUM22	Fertile	Not tested	Not tested
35601	GL00439	CG4107	Pcaf	pVALIUM22	Fertile	Not tested	Not tested
33946	HMS00897	CG5109	Pcl	pVALIUM20	Fertile	Not tested	Not tested
34690	HMS01169	CG3291	pcm	pVALIUM20	Fertile	Not tested	Not tested
35632	GL00479	CG17509	pds5	pVALIUM22	Fertile	Not tested	Not tested
35206	GL00082	CG17743	pho	pVALIUM22	Fertile	Not tested	Not tested
35724	GLV21089	CG3941	pita	pVALIUM21	Fertile	Not tested	Not tested
35178	GL00049	CG7001	Pk17E	pVALIUM22	Fertile	Not tested	Not tested
32840	HMS00531	CG11375	polybromo	pVALIUM20	Fertile	Not tested	Not tested
35322	GL00228	CG3307	pr-set7	pVALIUM22	Fertile	Not tested	Not tested
34622	HMS01297	CG8877	Prp8	pVALIUM20	Fertile	Not tested	Not tested
35297	GL00199	CG3886	Psc	pVALIUM22	Fertile	Not tested	Not tested
33700	HMS00576	CG5383	PSR	pVALIUM20	Fertile	Not tested	Not tested
35269	GL00167	CG32133	ptip	pVALIUM22	Fertile	Not tested	Not tested
34784	HMS00093	CG7138	r2d2	pVALIUM20	Fertile	Not tested	Not tested
32882	HMS00670	CG9862	Rae1	pVALIUM20	Fertile	Not tested	Not tested
33005	HMS00805	CG5252	Ranbp9	pVALIUM20	Fertile	Not tested	Not tested
32466	HMS00466	CG4879	RecQ5	pVALIUM20	Fertile	Not tested	Not tested
36654	GL00614	CG12189	Rev1	pVALIUM22	Fertile	Not tested	Not tested
34684	HMS01162	CG30149	rig	pVALIUM20	Fertile	Not tested	Not tested
43259	GLC01448	CG18145	Ripalpha	pVALIUM22	Fertile	Not tested	Not tested
32472	HMS00472	CG5422	Rox8	pVALIUM20	Fertile	Not tested	Not tested
34587	HMS01061	CG9273	RPA2	pVALIUM20	Fertile	Not tested	Not tested
36800	GL01005	CG7471	Rpd3	pVALIUM22	Fertile	Not tested	Not tested
43282	HMC02681	CG3180	Rpil140	pVALIUM20	Fertile	Not tested	Not tested
35420	GL00343	CG3178	Rrp1	pVALIUM22	Fertile	Not tested	Not tested
42604	HMS02437	CG7292	Rrp6	pVALIUM20	Fertile	Not tested	Not tested
38914	GL00686	CG4385	S	pVALIUM22	Fertile	Not tested	Not tested
35446	GL00371	CG5595	Sce	pVALIUM22	Fertile	Not tested	Not tested
36637	GL00597	CG30390	Sg29	pVALIUM22	Fertile	Not tested	Not tested
34630	HMS01305	CG9936	skd	pVALIUM20	Fertile	Not tested	Not tested
34949	HMS00313	CG18271	slx1	pVALIUM20	Fertile	Not tested	Not tested
36598	GL00558	CG6057	SMC1	pVALIUM22	Fertile	Not tested	Not tested
35477	GL00406	CG5263	smg	pVALIUM22	Fertile	Not tested	Not tested
34087	HMS00842	CG4013	Smr	pVALIUM20	Fertile	Not tested	Not tested
36125	HMS01540	CG4494	smt3	pVALIUM20	Fertile	Not tested	Not tested
36881	GL01045	CG31755	SoYb	pVALIUM22	Fertile	Not tested	Not tested
38898	GL00669	CG7948	spn-A	pVALIUM22	Fertile	Not tested	Not tested
43158	GL01500	CG3325	spn-B	pVALIUM22	Fertile	Not tested	Not tested
35148	GL00016	CG3169	Spt3	pVALIUM22	Fertile	Not tested	Not tested
35437	GL00362	CG8396	Ssb-c31a	pVALIUM22	Fertile	Not tested	Not tested
38344	HMS01811	CG14216	Ssu72	pVALIUM20	Fertile	Not tested	Not tested
34996	HMS01406	CG11268	ste14	pVALIUM20	Fertile	Not tested	Not tested
36867	GL01006	CG17149	Su(var)3-3	pVALIUM22	Fertile	Not tested	Not tested
33402	HMS00280	CG8013	Su(z)12	pVALIUM20	Fertile	Not tested	Not tested
33403	HMS00281	CG3905	Su(z)2	pVALIUM20	Fertile	Not tested	Not tested
36893	GL01080	CG7869	SuUR	pVALIUM22	Fertile	Not tested	Not tested
42770	GL00726	CG9874	Tbp	pVALIUM22	Fertile	Not tested	Not tested
35429	GL00353	CG3710	Tflls	pVALIUM22	Fertile	Not tested	Not tested
42931	HMS02624	CG9984	TH1	pVALIUM20	Fertile	Not tested	Not tested
41843	GL01271	CG1981	Thd1	pVALIUM22	Fertile	Not tested	Not tested
35603	GL00443	CG10387	tos	pVALIUM22	Fertile	Not tested	Not tested
35790	GL00397	CG10897	tou	pVALIUM22	Fertile	Not tested	Not tested
35794	GL00258	CR34649	tre-1	pVALIUM22	Fertile	Not tested	Not tested
41582	GL00699	CG33261,CG42507	Trl,CG42507	pVALIUM22	Fertile	Not tested	Not tested
36116	HMS01492	CG42595	uex	pVALIUM20	Fertile	Not tested	Not tested
34076	HMS00575	CG5640	Utx	pVALIUM20	Fertile	Not tested	Not tested
35468	GL00394	CG12701	vfl	pVALIUM22	Fertile	Not tested	Not tested
36877	GL01036	CG10728	vis	pVALIUM22	Fertile	Not tested	Not tested
36616	GL00576	CG3707	wapl	pVALIUM22	Fertile	Not tested	Not tested

32469	HMS00469	CG4448	wda	pVALIUM20	Fertile	Not tested	Not tested
33363	GLC01357	CG9226	WDR79	pVALIUM20	Fertile	Not tested	Not tested
35680	GLV21045	CG4148	wek	pVALIUM21	Fertile	Not tested	Not tested
32894	HMS00683	CG4548	XNP	pVALIUM20	Fertile	Not tested	Not tested
35208	GL00084	CG7803	z	pVALIUM22	Fertile	Not tested	Not tested
35228	GL00112	CG12314	zuc	pVALIUM22	Fertile	Not tested	Not tested
34846	HMS00164	CG7471	Rpd3	pVALIUM20	Fertile	Not tested	Not tested
32853	HMS00638	CG17149	Su(var)3-3	pVALIUM20	Fertile	Not tested	Not tested
33726	HMS00608	CG17149	Su(var)3-3	pVALIUM20	Fertile	Not tested	Not tested
38285	HMS01738	CG7467	osa	pVALIUM20	Fertile	Not tested	Not tested
33954	HMS00909	CG4303	Bap60	pVALIUM20	Fertile	Not tested	Not tested
34520	HMS00050	CG5942	brm	pVALIUM20	Fertile	Not tested	Not tested
41857	GL01285	CG4193	dhd	pVALIUM22	Fertile	Not tested	Not tested
36867	GL01006	CG17149	Su(var)3-3	pVALIUM22	Fertile	Not tested	Not tested
32993	HMS00793	CG3143	foxo	pVALIUM20	Fertile	Not tested	Not tested
32427	HMS00422	CG3143	foxo	pVALIUM20	Fertile	Not tested	Not tested
34774	HMS00083	CG1770	HDAC4	pVALIUM20	Fertile	Not tested	Not tested
44103	HMS02823	CG44132	mamo	pVALIUM20	Fertile	Not tested	Not tested
37486	GLV21091	CG44132	mamo	pVALIUM21	Fertile	Not tested	Not tested
40850	HMS02017	CG8591	CTCF	pVALIUM20	Fertile	Not tested	Not tested
42536	HMJ02105	CG6384	Cp190	pVALIUM20	Fertile	Not tested	Not tested
33944	HMS00895	CG6384	Cp190	pVALIUM20	Fertile	Not tested	Not tested
33903	HMS00845	CG6384	Cp190	pVALIUM20	Fertile	Not tested	Not tested
35642	GLV21006	CG10159	BEAF-32	pVALIUM21	Fertile	Not tested	Not tested
34006	HMS00970	CG8573	su(Hw)	pVALIUM20	Fertile	Not tested	Not tested
33906	HMS00848	CG8573	su(Hw)	pVALIUM20	Fertile	Not tested	Not tested
34351	HMS01340	CG6057	SMC1	pVALIUM20	Fertile	Not tested	Not tested
33431	HMS00318	CG9802	Cap	pVALIUM20	Fertile	Not tested	Not tested
33622	HMS00016	CG32443	Pc	pVALIUM20	Fertile	Not tested	Not tested
33964	HMS00921	CG32443	Pc	pVALIUM20	Fertile	Not tested	Not tested
42926	HMS02619	CG17743	pho	pVALIUM20	Fertile	Not tested	Not tested
38261	HMS01706	CG3886	Psc	pVALIUM20	Fertile	Not tested	Not tested
40940	HMS02188	CG33261,CG42507	Tri,CG42507	pVALIUM20	Fertile	Not tested	Not tested
43962	GL01314	CG10798	dm	pVALIUM22	Fertile	Not tested	Not tested
36126	HMS01541	CG6376	E2f	pVALIUM20	Fertile	Not tested	Not tested
33669	HMS00082	CG18412	ph-p	pVALIUM20	Fertile	Not tested	Not tested
35207	GL00083	CG18412	ph-p	pVALIUM22	Fertile	Not tested	Not tested
36720	HMS01610	CG1772	dap	pVALIUM20	Fertile	Not tested	Not tested
32890	HMS00678	CG10414	Atac2	pVALIUM20	Sterile, class 1	N/A	N/A
36624	GL00584	CG10726	barr	pVALIUM22	Sterile, class 1	N/A	N/A
35185	GL00057	CG9748	bel	pVALIUM22	Sterile, class 1	N/A	N/A
36067	GL00485	CG10480	Bj1	pVALIUM22	Sterile, class 1	N/A	N/A
33421	HMS00304	CG31132	BRWD3	pVALIUM20	Sterile, class 1	N/A	N/A
32989	HMS00789	CG7581	Bub3	pVALIUM20	Sterile, class 1	N/A	N/A
35199	GL00073	CG3319	Cdk7	pVALIUM22	Sterile, class 1	N/A	N/A
34857	HMS00175	CG10333	CG10333	pVALIUM20	Sterile, class 1	N/A	N/A
43205	GL01550	CG10565	CG10565	pVALIUM22	Sterile, class 1	N/A	N/A
42810	GL01184	CG3430	CG3430	pVALIUM22	Sterile, class 1	N/A	N/A
35020	HMS01433	CG5033	CG5033	pVALIUM20	Sterile, class 1	N/A	N/A
43254	GLC01443	CG5290	CG5290	pVALIUM22	Sterile, class 1	N/A	N/A
32334	HMS00325	CG5589	CG5589	pVALIUM20	Sterile, class 1	N/A	N/A
43206	GL01551	CG5800	CG5800	pVALIUM22	Sterile, class 1	N/A	N/A
34565	HMS01037	CG6015	CG6015	pVALIUM20	Sterile, class 1	N/A	N/A
34804	HMS00113	CG7185	CG7185	pVALIUM20	Sterile, class 1	N/A	N/A
36609	GL00589	CG8142	CG8142	pVALIUM22	Sterile, class 1	N/A	N/A
36589	GL00549	CG9253	CG9253	pVALIUM22	Sterile, class 1	N/A	N/A
41679	HMS02243	CG9422	CG9422	pVALIUM20	Sterile, class 1	N/A	N/A
34564	HMS01036	CG5602	DNA-lig1	pVALIUM20	Sterile, class 1	N/A	N/A
34827	HMS00142	CG9696	dom	pVALIUM20	Sterile, class 1	N/A	N/A
35149	GL00017	CG1828	dre4	pVALIUM22	Sterile, class 1	N/A	N/A
35271	GL00169	CG7776	E(Pc)	pVALIUM22	Sterile, class 1	N/A	N/A
32345	HMS00336	CG6474	e(y)1	pVALIUM20	Sterile, class 1	N/A	N/A
35764	HMS01512	CG8648	Fen1	pVALIUM20	Sterile, class 1	N/A	N/A
41716	HMS02281	CG5148	cal1	pVALIUM20	Sterile, class 1	N/A	N/A
33666	HMS00076	CG7269	Hel25E	pVALIUM20	Sterile, class 1	N/A	N/A
41988	HMS02387	CG33882,CG33880,CG33876,C	His2B	pVALIUM20	Sterile, class 1	N/A	N/A
35299	GL00202	CG6620	ial	pVALIUM22	Sterile, class 1	N/A	N/A
34552	HMS01024	CG10850	ida	pVALIUM20	Sterile, class 1	N/A	N/A
35366	GL00279	CG12165	lncnp	pVALIUM22	Sterile, class 1	N/A	N/A
41916	HMS02313	CG6189	l(1)1Bi	pVALIUM20	Sterile, class 1	N/A	N/A
43981	GL01591	CG10563	l(2)37Cd	pVALIUM22	Sterile, class 1	N/A	N/A
33002	GLC01635	CG8426	l(2)NC136	pVALIUM22	Sterile, class 1	N/A	N/A
34024	HMS02950	CG5931	l(3)72Ab	pVALIUM20	Sterile, class 1	N/A	N/A
34575	HMS01049	CG12031	MED14	pVALIUM20	Sterile, class 1	N/A	N/A
34664	HMS01141	CG7957	MED17	pVALIUM20	Sterile, class 1	N/A	N/A
34731	HMS01211	CG17397	MED21	pVALIUM20	Sterile, class 1	N/A	N/A
33743	HMS01081	CG9473	MED6	pVALIUM20	Sterile, class 1	N/A	N/A
34926	HMS01275	CG13867	MED8	pVALIUM20	Sterile, class 1	N/A	N/A
35398	GL00318	CG8103	Mi-2	pVALIUM22	Sterile, class 1	N/A	N/A
35150	GL00018	CG13778	Mnn1	pVALIUM22	Sterile, class 1	N/A	N/A
34990	HMS01400	CG12202	Nat1	pVALIUM20	Sterile, class 1	N/A	N/A
37482	GL00625	CG9938	Ndc80	pVALIUM22	Sterile, class 1	N/A	N/A
33694	HMS00564	CG7421	Nopp140	pVALIUM20	Sterile, class 1	N/A	N/A
32836	HMS00526	CG34407	Not1	pVALIUM20	Sterile, class 1	N/A	N/A
34710	HMS01189	CG4579	Nup154	pVALIUM20	Sterile, class 1	N/A	N/A
33897	HMS00837	CG3820	Nup214	pVALIUM20	Sterile, class 1	N/A	N/A
33003	HMS00803	CG11856	Nup358	pVALIUM20	Sterile, class 1	N/A	N/A
35695	GLV21060	CG6251	Nup62	pVALIUM21	Sterile, class 1	N/A	N/A
32409	GLC01628	CG2917	Orc4	pVALIUM22	Sterile, class 1	N/A	N/A
32838	HMS00528	CG8241	pea	pVALIUM20	Sterile, class 1	N/A	N/A
41997	HMS02398	CG7769	pic	pVALIUM20	Sterile, class 1	N/A	N/A
43984	GLC01465	CG6375	pit	pVALIUM22	Sterile, class 1	N/A	N/A
32865	HMS00652	CG5519	Prp19	pVALIUM20	Sterile, class 1	N/A	N/A
41912	HMS02306	CG4320	raptor	pVALIUM20	Sterile, class 1	N/A	N/A

32415	HMS00410	CG9750	rept	pVALIUM20	Sterile, class 1	N/A	N/A
41861	GL01291	CG11979	Rpb5	pVALIUM22	Sterile, class 1	N/A	N/A
36830	GL01068	CG1554	RplI215	pVALIUM22	Sterile, class 1	N/A	N/A
34567	HMS01040	CG7885	RplI33	pVALIUM20	Sterile, class 1	N/A	N/A
32363	HMS00354	CG2173	Rs1	pVALIUM20	Sterile, class 1	N/A	N/A
38531	HMS01744	CG9198	shtd	pVALIUM20	Sterile, class 1	N/A	N/A
35602	GL00440	CG10212	SMC2	pVALIUM22	Sterile, class 1	N/A	N/A
38371	HMS01840	CG13570	spag	pVALIUM20	Sterile, class 1	N/A	N/A
33914	HMS00857	CG9809	Spargel	pVALIUM20	Sterile, class 1	N/A	N/A
35466	GL00392	CG11451	Spc105R	pVALIUM22	Sterile, class 1	N/A	N/A
32896	HMS00685	CG12372	spt4	pVALIUM20	Sterile, class 1	N/A	N/A
34837	HMS00153	CG7626	Spt5	pVALIUM20	Sterile, class 1	N/A	N/A
32373	HMS00364	CG12225	Spt6	pVALIUM20	Sterile, class 1	N/A	N/A
38948	HMS01862	CG6238	ssh	pVALIUM20	Sterile, class 1	N/A	N/A
41648	GL01230	CG8592	stil	pVALIUM22	Sterile, class 1	N/A	N/A
35415	GL00337	CG3836	stwl	pVALIUM22	Sterile, class 1	N/A	N/A
38255	HMS01699	CG12864	Su(var)2-HP2	pVALIUM20	Sterile, class 1	N/A	N/A
35314	GL00219	CG17603	TafI	pVALIUM22	Sterile, class 1	N/A	N/A
35367	GL00280	CG7704	Taf5	pVALIUM22	Sterile, class 1	N/A	N/A
35428	GL00352	CG10281	TflIFalpha	pVALIUM22	Sterile, class 1	N/A	N/A
35462	GL00388	CG3278	Tif-IA	pVALIUM22	Sterile, class 1	N/A	N/A
36307	GL00491	CG1782	Uba1	pVALIUM22	Sterile, class 1	N/A	N/A
35806	GL00452	CG7528	Uba2	pVALIUM22	Sterile, class 1	N/A	N/A
42642	HMS02478	CG7989	wcd	pVALIUM20	Sterile, class 1	N/A	N/A
32952	HMS00746	CG17437	wds	pVALIUM20	Sterile, class 1	N/A	N/A
38297	HMS01758	CG7670	WRNexo	pVALIUM20	Sterile, class 1	N/A	N/A
36123	HMS01538	CG10798	dm	pVALIUM20	Sterile, class 1	N/A	N/A
51454	HMC03189	CG10798	dm	pVALIUM20	Sterile, class 1	N/A	N/A
34069	HMS00051	CG4236	Caf1	pVALIUM20	Sterile, class 1	N/A	N/A
35139	GL00006	CG2252	fs(1)h	pVALIUM22	Sterile, class 1	N/A	N/A
52111	HMC03356	CG4039	Mcm6	pVALIUM20	Sterile, class 1	N/A	N/A
41842	GL01270	CG4039	Mcm6	pVALIUM22	Sterile, class 1	N/A	N/A
36092	GL00511	CG3938	CycE	pVALIUM22	Sterile, class 1	N/A	N/A
38902	GL00673	CG3938	CycE	pVALIUM22	Sterile, class 1	N/A	N/A
41693	HMS02257	CG2252	fs(1)h	pVALIUM20	Sterile, class 1	N/A	N/A
41943	HMS02340	CG2252	fs(1)h	pVALIUM20	Sterile, class 1	N/A	N/A
44009	HMS02723	CG2252	fs(1)h	pVALIUM20	Sterile, class 1	N/A	N/A
35388	GL00307	CG6677	ash2	pVALIUM22	Sterile, class 2	N/A	Not tested
34390	HMS01384	CG13192	CG13192	pVALIUM20	Sterile, class 2	N/A	Not tested
33704	HMS00581	CG40351	Set1	pVALIUM20	Sterile, class 2	N/A	Not tested
35460	GL00386	CG2161	Rga	pVALIUM22	Sterile, class 2	N/A	Not tested
32473	HMS00473	CG16975	Sfmbt	pVALIUM20	Sterile, class 2	N/A	Not tested
35454	GL00379	CG4735	shu	pVALIUM22	Sterile, class 2	N/A	Not tested
35303	GL00206	CG3158	spn-E	pVALIUM22	Sterile, class 2	N/A	Not tested
35592	GL00429	CG42373,CG31917	Tb5,CG31917	pVALIUM22	Sterile, class 2	N/A	Not tested
32926	HMS00718	CG17293	Wdr82	pVALIUM20	Sterile, class 2	N/A	Not tested
35240	GL00126	CG3733	Chd1	pVALIUM22	Sterile, class 2	N/A	Not tested
36068	GL00486	CG6502	E(z)	pVALIUM22	Sterile, class 2	N/A	Not tested
35323	GL00230	CG5179	Cdk9	pVALIUM22	Sterile, class 2	N/A	Not tested
35201	GL00076	CG6137	aub	pVALIUM22	Sterile, class 3	N/A	Not tested
35443	GL00368	CG10542	Bre1	pVALIUM22	Sterile, class 3	N/A	Not tested
36783	GL00518	CG9802	Cap	pVALIUM22	Sterile, class 3	N/A	Not tested
35324	GL00231	CG10572	Cdk8	pVALIUM22	Sterile, class 3	N/A	Not tested
36899	GL01093	CG14222	CG14222	pVALIUM22	Sterile, class 3	N/A	Not tested
36916	HMS01019	CG3848	trr	pVALIUM20	Sterile, class 3	N/A	Not tested
32517	HMS00522	CG4184	MED15	pVALIUM20	Sterile, class 3	N/A	Not tested
32995	HMS00795	CG32491	mod(mdg4)	pVALIUM20	Sterile, class 3	N/A	Not tested
35630	GL00477	CG18740	mor	pVALIUM22	Sterile, class 3	N/A	Not tested
38190	GL00629	CG12047	mud	pVALIUM22	Sterile, class 3	N/A	Not tested
34580	HMS01054	CG2158	Nup50	pVALIUM20	Sterile, class 3	N/A	Not tested
43215	GL01560	CG3041	Orc2	pVALIUM22	Sterile, class 3	N/A	Not tested
42803	GL01176	CG9183	plu	pVALIUM22	Sterile, class 3	N/A	Not tested
42482	GL01341	CG1404	ran	pVALIUM22	Sterile, class 3	N/A	Not tested
35171	GL00041	CG10683	rhi	pVALIUM22	Sterile, class 3	N/A	Not tested
36794	GL00534	CG3423	SA	pVALIUM22	Sterile, class 3	N/A	Not tested
32372	HMS00363	CG1064	Snr1	pVALIUM20	Sterile, class 3	N/A	Not tested
36095	GL00514	CG13109	tai	pVALIUM22	Sterile, class 3	N/A	Not tested
36786	GL00522	CG17436	vtd	pVALIUM22	Sterile, class 3	N/A	Not tested
43141	GL01482	CG15737	wisp	pVALIUM22	Sterile, class 3	N/A	Not tested
32845	HMS00628	CG8625	lswi	pVALIUM20	Sterile, class 3	N/A	Not tested
33907	HMS00849	CG32491	mod(mdg4)	pVALIUM20	Sterile, class 3	N/A	Not tested
33395	HMS00272	CG3423	SA	pVALIUM20	Sterile, class 3	N/A	Not tested
33659	HMS00066	CG6502	E(z)	pVALIUM20	Sterile, class 3	N/A	Not tested
33654	HMS00060	CG3938	CycE	pVALIUM20	Sterile, class 3	N/A	No tested
33739	HMS01075	CG5451	Smu1	pVALIUM20	Sterile, class 4	Yes	WT
32889	HMS00677	CG7583	CtBP	pVALIUM20	Sterile, class 4	Yes	WT
35242	GL00129	CG7055	dalao	pVALIUM22	Sterile, class 4	Yes	WT
43185	GL01529	CG2207	Df31	pVALIUM22	Sterile, class 4	Yes	WT
35488	GL00417	CG15433	Elp3	pVALIUM22	Sterile, class 4	Yes	WT
33706	HMS00583	CG1716	Set2	pVALIUM20	Sterile, class 4	Yes	WT
35399	GL00319	CG42245,CG1244	CG42245.MEP-1	pVALIUM22	Sterile, class 4	Yes	WT
35759	HMS01506	CG8384	gro	pVALIUM20	Sterile, class 4	Yes	WT
34778	HMS00087	CG2128	Hdac3	pVALIUM20	Sterile, class 4	Yes	WT
35238	GL00123	CG11990	hyx	pVALIUM22	Sterile, class 4	Yes	WT
36597	GL00557	CG3696	kis	pVALIUM22	Sterile, class 4	Yes	WT
34980	HMS01389	CG2684	lds	pVALIUM20	Sterile, class 4	Yes	WT
36652	GL00612	CG9088	lid	pVALIUM22	Sterile, class 4	No	dhd-like
35152	GL00020	CG10895	lok	pVALIUM22	Sterile, class 4	Yes	WT
36829	GL01067	CG1793	MED26	pVALIUM22	Sterile, class 4	Yes	WT
39028	HMS01947	CG16928	mre11	pVALIUM20	Sterile, class 4	Yes	WT
36614	GL00574	CG17704	Nipped-B	pVALIUM22	Sterile, class 4	Yes	WT
33985	HMS00945	CG4118	nxft2	pVALIUM20	Sterile, class 4	Yes	WT
37483	GL00626	CG6122	piwi	pVALIUM22	Sterile, class 4	Yes	WT
35419	GL00342	CG2368	psq	pVALIUM22	Sterile, class 4	Yes	WT

36632	GL00592	CG12352	san	pVALIUM22	Sterile, class 4	Yes	WT
35389	GL00308	CG9495	Scm	pVALIUM22	Sterile, class 4	Yes	WT
32368	HMS00359	CG8815	Sin3A	pVALIUM20	Sterile, class 4	No	dhd-like
40881	HMS02049	CG4413	trem	pVALIUM20	Sterile, class 4	Yes	WT
36835	GL01076	CG18009	Trf2	pVALIUM22	Sterile, class 4	Yes	WT
35706	GLV21071	CG9088	lid	pVALIUM21	Sterile, class 4	No	dhd-like
33725	HMS00607	CG7471	Rpd3	pVALIUM20	Sterile, class 4	No	dhd-like
Control	GL00257	CG12153	Hira	pVALIUM22	Sterile, class 4	No	ssm-like
Control	MTD>attP2				Fertile	Yes	WT
Control	w [1118]				Fertile	Yes	WT
					Female fertility	GFP::Cid expression	Sperm nuclear phenotype

F8gn0260027	CG42495	-0.968787	8.35E-07	3.10E-05
F8gn0030598	CG9903	-0.7904792	8.51E-07	3.14E-05
F8gn0031998	SLSA11	-0.7290666	9.82E-07	3.615E-05
F8gn0031589	CG3714	-0.522241	1.01E-06	3.71E-05
F8gn0033205	CG2064	-0.7867165	1.04E-06	3.79E-05
F8gn0037581	CG7352	-0.8753558	1.08E-06	3.92E-05
F8gn0028468	rlet	-0.6723628	1.43E-06	5.16E-05
F8gn0015569	alpha-Est10	-0.8392161	1.44E-06	5.18E-05
F8gn0031117	GstE3	-0.6298817	1.48E-06	5.2882E-05
F8gn0265761	CR44568	-0.9108888	1.88E-06	6.72E-05
F8gn0030668	CG8128	-0.6943842	1.92E-06	6.81E-05
F8gn0050005	GstI2	-0.6438248	2.01E-06	7.11E-05
F8gn0034032	CG8195	-0.5408707	2.05E-06	7.22E-05
F8gn0030925	Hayan	-0.9826592	2.15E-06	7.55E-05
F8gn0037000	Znf77C	-0.6746768	2.26E-06	7.89E-05
F8gn0004108	Nrt	-0.6964366	2.36E-06	8.20E-05
F8gn0032749	TBC1D16	-0.7115413	2.42E-06	8.36E-05
F8gn0265924	CG42324	-0.8959483	2.45E-06	8.44E-05
F8gn0033233	Kdm4A	-0.5605395	2.53E-06	8.70E-05
F8gn0262902	CR43257	-0.957284	2.61E-06	8.94E-05
F8gn0038012	CG10013	-0.9437117	2.77E-06	9.36E-05
F8gn0039094	CG10184	-0.7620824	2.77E-06	9.36E-05
F8gn0033875	CG6357	-0.8296114	2.97E-06	9.98E-05
F8gn0011660	Pms2	-0.6191626	4.01E-06	1.33E-04
F8gn0023540	CG3630	-0.7750882	4.38E-06	1.44E-04
F8gn0040260	Ugt369c	-0.7920914	4.55E-06	1.49E-04
F8gn0032744	Ttc19	-0.7666054	4.63E-06	1.52E-04
F8gn0034399	CG15083	-0.6891411	4.71E-06	1.54E-04
F8gn0031161	CG15446	-0.559749	5.16E-06	1.66E-04
F8gn0034569	dgl3	-0.5100777	5.16E-06	1.66E-04
F8gn0031011	CG8034	-0.5605182	5.44E-06	0.00017485
F8gn0027499	wde	-0.4701503	5.51E-06	1.77E-04
F8gn0053170	CG33170	-0.7056534	5.85E-06	0.00018566
F8gn0000152	Aox	-0.6035714	5.84E-06	1.86E-04
F8gn0017448	CG2187	-0.7509568	5.90E-06	1.86E-04
F8gn0024222	lK9beta	-0.5333447	6.15E-06	1.94E-04
F8gn0011642	Zyx	-0.6685735	6.47E-06	0.00020304
F8gn0034928	CG13562	-0.9576823	6.71E-06	2.10E-04
F8gn0025702	Sprk9D	-0.9076152	6.96E-06	2.17E-04
F8gn0032210	CPLD	-0.6022837	8.33E-06	2.58E-04
F8gn0035160	hng3	-0.7523256	8.55E-06	2.63E-04
F8gn0040384	CG2795	-0.5216919	8.89E-06	2.72E-04
F8gn0030683	CG8239	-0.5479759	1.16E-05	3.53E-04
F8gn0038704	CG5316	-0.5674446	1.19E-05	3.59E-04
F8gn0261794	lcc	-0.5170776	1.31E-05	3.96E-04
F8gn0267967	orella	-0.7468133	1.40E-05	4.20E-04
F8gn0034985	CG3328	-0.9088474	1.42E-05	4.23E-04
F8gn0058160	CG40160	-0.5573892	1.44E-05	0.00042856
F8gn0263001	CR43309	-0.9140864	1.55E-05	4.59E-04
F8gn0035264	Osepd	-0.8324113	1.58E-05	4.66E-04
F8gn0030467	CG1764	-0.6086858	1.68E-05	4.94E-04
F8gn0031224	CG11454	-0.6511044	1.79E-05	5.23E-04
F8gn0261363	PPO3	-0.908642	1.87E-05	5.43E-04
F8gn0036765	CG7408	-0.8043628	1.92E-05	5.57E-04
F8gn0038966	pinta	-0.6002371	1.93E-05	5.57E-04
F8gn0264979	CG4267	-0.5921441	1.95E-05	5.62E-04
F8gn0031974	CG12560	-0.6468851	2.10E-05	5.99E-04
F8gn0030349	CG10353	-0.5110312	2.26E-05	6.42E-04
F8gn0046685	WscK	-0.554038	2.39E-05	6.80E-04
F8gn0036196	CG11658	-0.6157453	2.51E-05	7.08E-04
F8gn0029232	neur	-0.5215054	2.55E-05	7.16E-04
F8gn0038490	CG5285	-0.6202791	2.59E-05	7.25E-04
F8gn0027535	hntv	-0.5633173	2.64E-05	7.39E-04
F8gn0011676	Nus	-0.882868	2.73E-05	7.61E-04
F8gn0052109	CG32109	-0.6631842	2.96E-05	0.00082025
F8gn0266211	CR44906	-0.84417	3.08E-05	8.49E-04
F8gn0032079	CG31886	-0.7657233	3.21E-05	8.78E-04
F8gn0086898	dgo	-0.6488394	3.30E-05	9.01E-04
F8gn0036331	CG14117	-0.7419989	3.32E-05	9.02E-04
F8gn0033715	CG8490	-0.565824	3.39E-05	9.17E-04
F8gn0034578	CG15653	-0.6741581	3.52E-05	9.50E-04
F8gn0033261	uidl	-0.5640484	3.64E-05	0.00097761
F8gn0085212	CG34183	-0.5802033	3.68E-05	0.00098506
F8gn0030716	CG9170	3.51543874	2.73E-133	1.25E-129
F8gn0030710	tipe	4.51536274	2.28E-119	5.23E-116
F8gn0040370	CG13375	3.94270954	2.25E-117	4.12E-114
F8gn0063499	GstE10	4.38522941	1.29E-105	1.96E-102
F8gn0263219	Dscam4	3.38390085	1.29E-79	1.48E-76
F8gn0063491	GstE9	2.37209129	5.43E-74	5.53E-71
F8gn0267509	nrm	3.26989536	5.19E-72	4.76E-69
F8gn0000246	o3JG	1.80256044	7.60E-51	4.64E-48
F8gn0052773	CR32773	2.91153609	2.21E-48	1.06E-45
F8gn0033744	DM4-R2	2.21228708	3.41E-47	1.56E-44
F8gn0004034	y	2.3656636	1.46E-36	4.63E-34
F8gn0039129	Rpf19b	2.25634285	4.45E-29	9.94E-27
F8gn0052625	CG32625	1.50639584	1.75E-26	3.56E-24
F8gn0003965	v	2.2137325	8.01E-26	1.56E-23
F8gn0028400	Syt4	1.60413573	8.45E-26	1.61E-23
F8gn0035086	CG12851	2.20739805	3.06E-25	5.51E-23
F8gn0014417	CG13397	0.9799056	2.34E-24	4.05E-22
F8gn0052373	CG32373	1.65820956	3.79E-22	5.89E-20
F8gn0039683	RhoGAP100F	1.78355414	4.14E-22	6.32E-20
F8gn0052816	CG32816	1.34949326	3.29E-21	4.78E-19
F8gn0036369	CG10089	1.43828569	3.97E-19	5.58E-17
F8gn0035083	Tina-1	0.97905168	3.91E-18	4.84E-16
F8gn0031645	CG3086	1.20681255	3.21E-16	3.59E-14
F8gn0283427	FASN1			
F8gn0050401	dary			
F8gn00263705	Myle10A			
F8gn00275676	Oblalpha.3			
F8gn0039927	CG11355			
F8gn0037623	CG9801			
F8gn0031157	CG1503			
F8gn0004895	Fnd1			
F8gn0266710	hng			
F8gn0000038	nACtubeta1			
F8gn0039118	CG10208			
F8gn0033458	CG18446			
F8gn0033520	Pnc2540-1			
F8gn0031689	Gpr28d1			
F8gn0032729	L2HG0H			
F8gn0037717	CG8301			
F8gn0033205	CG2064			
F8gn0027621	Pfrx			
F8gn0050460	CG30460			
F8gn0035094	CG9380			
F8gn0051793	CG31793			
F8gn0025885	lncs			
F8gn0037326	CG14669			
F8gn0032719	CG17321			
F8gn0037606	CG8030			
F8gn0026417	Hus1-like			
F8gn0001085	fz			
F8gn0267347	squ			
F8gn0033257	sard			
F8gn0003071	PIK			
F8gn0261041	sil			
F8gn0051676	CG31676			
F8gn0038296	CG6752			
F8gn0022709	Auk1			
F8gn0032161	CG4594			
F8gn0023535	arp			
F8gn0033246	ACC			
F8gn0017482	L3dh			
F8gn0014417	CG13397			
F8gn0031703	CG12512			
F8gn0264816	lckln			
F8gn0037742	lgt3R			
F8gn0033815	CG4676			
F8gn0085339	CG34310			
F8gn0263665	CR43656			
		0.86975756	2.30E-08	1.52E-06
		1.11817354	2.28E-08	1.52E-06
		0.98315534	3.45E-08	2.20E-06
		0.67693092	1.26E-07	7.83E-06
		0.90203488	1.71E-07	1.03E-05
		0.6640328	2.35E-07	1.37E-05
		0.93002568	3.29E-07	1.88E-05
		1.0757224	4.27E-07	2.42E-05
		0.83815038	4.39E-07	2.45E-05
		1.04459491	6.48E-07	3.55E-05
		0.80955118	1.20E-06	6.26E-05
		0.58308966	1.28E-06	6.64E-05
		0.74375135	1.29E-06	6.66E-05
		0.80053352	1.58E-06	8.04E-05
		0.52320052	1.72E-06	8.62E-05
		0.54410035	1.94E-06	9.55E-05
		0.73749684	3.06E-06	1.47E-04
		0.72002681	3.25E-06	1.53E-04
		0.62944429	3.81E-06	1.76E-04
		0.93844635	4.81E-06	0.00021789
		0.65755611	4.93E-06	2.23E-04
		0.61521329	5.50E-06	2.47E-04
		0.84098198	5.58E-06	2.49E-04
		0.73961713	6.03E-06	2.66E-04
		0.48037173	6.10E-06	2.68E-04
		0.61835828	6.46E-06	2.82E-04
		0.60208374	6.78E-06	2.95E-04
		0.60195776	8.58E-06	3.66E-04
		0.9250535	8.99E-06	3.81E-04
		0.57585055	9.37E-06	3.95E-04
		0.81342642	1.07E-05	4.47E-04
		0.89263646	1.42E-05	5.81E-04
		0.66542534	1.56E-05	6.28E-04
		0.55981758	1.65E-05	6.58E-04
		0.57432263	1.75E-05	6.89E-04
		0.90279534	1.76E-05	6.89E-04
		0.51387976	1.98E-05	7.62E-04
		0.78501069	2.08E-05	7.91E-04
		0.40989893	2.1026E-05	0.00079514
		0.55306249	2.15E-05	8.08E-04
		0.43906737	2.28E-05	8.50E-04
		0.90083167	2.34E-05	8.66E-04
		0.88304551	2.39E-05	8.80E-04
		0.71342358	2.46E-05	9.03E-04
		0.67340873	2.72E-05	9.93E-04

FBgn0028743	Uhit	1,400557	1,35E-15	1,44E-13
FBgn0264439	CG43857	1,44206442	3,80E-15	3,91E-13
FBgn0029723	Prnc-R	1,16467501	8,98E-15	9,10E-13
FBgn0036368	CG10738	1,12952567	1,06E-14	1,05E-12
FBgn0267733	CR46064	1,30463417	2,70E-14	2,55E-12
FBgn0266258	CG44953	1,61734556	2,70E-14	2,55E-12
FBgn0038204	CG14357	1,45024866	1,17E-13	1,04E-11
FBgn0039900	Syt7	1,57434289	1,36E-13	1,18E-11
FBgn0031275	GABA-B-R3	1,56925587	1,55E-13	1,34E-11
FBgn0039118	CG10208	1,225912	1,70E-13	1,45E-11
FBgn0267347	sqn	0,97779595	4,75E-13	3,85E-11
FBgn0003866	tsh	1,33447888	7,38E-13	5,88E-11
FBgn0033458	CG18446	0,8595366	9,18E-13	7,25E-11
FBgn0050403	CG30403	0,83287764	1,55E-11	1,10E-09
FBgn0032645	CG15142	1,25685861	2,82E-11	1,94E-09
FBgn0264385	CG43837	0,80630682	7,48E-11	4,93E-09
FBgn0038038	CG5167	0,86009134	2,30E-10	1,46E-08
FBgn0263093	CR43361	0,98968277	7,16E-10	4,26E-08
FBgn0284435	lyn	0,90935674	3,67E-09	1,9771E-07
FBgn0259229	CG42329	0,99577599	4,2094E-09	2,2561E-07
FBgn0035317	Osep2	1,2035749	4,82E-09	2,57E-07
FBgn0030723	dpr18	1,20750721	7,53E-09	3,90E-07
FBgn0266005	CR44779	0,91939442	1,40E-08	7,03E-07
FBgn0283427	FASN1	0,86971339	2,30E-08	1,11E-06
FBgn0037902	CG5281	1,01450627	2,83E-08	1,35E-06
FBgn0261266	zuc	0,59100256	4,51E-08	2,07E-06
FBgn0040001	FASN3	1,15849921	5,08E-08	2,32E-06
FBgn0039341	CG5112	0,52776631	9,83E-08	4,21E-06
FBgn0039938	Sox102F	0,96839303	1,34E-07	5,6477E-06
FBgn0051997	CG31997	1,08525441	1,72E-07	7,21E-06
FBgn0042696	Nfl	1,08171827	2,65E-07	1,07E-05
FBgn0263617	CR43626	1,0436084	3,50E-07	1,40E-05
FBgn0023535	arg	1,07140257	3,57E-07	1,42E-05
FBgn0039543	CG12428	0,51181318	3,79E-07	1,50E-05
FBgn0263660	CR43651	0,94475809	5,01E-07	1,94E-05
FBgn0033246	ACC	0,58306679	1,27E-06	4,62E-05
FBgn0265168	CR44237	0,75743644	2,18E-06	7,63E-05
FBgn0039927	CG11155	0,81040266	2,64E-06	9,00E-05
FBgn0260745	mfaf	0,73309962	2,88E-06	9,71E-05
FBgn0260812	mfaf-D	0,94838747	3,51E-06	1,17E-04
FBgn0027611	LManII	0,78767635	3,80E-06	1,27E-04
FBgn0002905	rnas.306	0,63117721	3,93E-06	1,30E-04
FBgn0004047	Yp3	0,86055593	4,75E-06	1,54E-04
FBgn0025629	CG4045	0,61428341	5,64E-06	1,80E-04
FBgn0264265	CG43765	0,89472755	7,33E-06	2,28E-04
FBgn0004895	foad1	0,94745812	8,43E-06	2,60E-04
FBgn0031703	CG12512	0,57709695	9,14E-06	2,79E-04
FBgn0041581	AttB	0,94032658	9,6168E-06	0,00029282
FBgn0025382	Rab27	0,84551951	1,32E-05	0,000398
FBgn0030583	CG14410	0,57134888	1,4142E-05	0,00042218
FBgn0263916	Enr2	0,54077834	1,78E-05	5,21E-04
FBgn0030660	CG8097	0,47918563	1,96E-05	5,64E-04
FBgn0016926	Hrno	0,5721965	2,07E-05	5,92E-04
FBgn0267001	Tern-a	0,80076692	2,42E-05	6,84E-04
FBgn0037577	CG1443	0,82607808	2,8441E-05	0,00078987
FBgn0026361	Sep6	0,58579619	3,19E-05	8,77E-04
FBgn0085370	Pde11	0,4243612	3,21E-05	8,78E-04
FBgn0052702	Gdm	0,56358184	3,33E-05	0,00090215

S3 Table. Quantitative analysis of H3K4me3 differential enrichment in Control vs *lid* KD ovarian ChIP-Seq.

Table with columns: chr, start, end, Conc_WT, Conc_LidKD, Fold, p-value, FDR, Annotation, Detailed Annotation, Distance to TSS, and Nearest PromoterID. The table contains quantitative analysis data for H3K4me3 differential enrichment across various chromosomes and genomic features.

chrX	6086651	6087131	1146	11.91	0.45	3.92E-05	0.000787	exon (Fg)0003302, exon 3 of 3)	exon (NM) 076498, exon 2 of 2)	381	Fg)0003302 Fg)00070889
chrX	20430219	20430719	11.27	11.73	0.45	4.35E-05	0.000858	promoter-TSS (Fg)0031115)	promoter-TSS (NM) 167704)	335	Fg)0003302 Fg)00077306
chr2R	16150071	16150571	11.16	11.61	0.45	5.09E-05	0.000976	promoter-TSS (NM) 001032251)	promoter-TSS (NM) 001032251)	199	Fg)0003303 Fg)00030204
chr3L	8514309	8514809	11.09	11.55	0.45	7.98E-05	0.00139	exon (Fg)0005902, exon 2 of 2)	exon (NM) 399755, exon 2 of 2)	362	Fg)0003302 Fg)00076655
chr3R	5793722	5794222	11.78	12.23	0.45	0.000119	0.00192	promoter-TSS (Fg)0037379)	promoter-TSS (NM) 1413394)	345	Fg)0003303 Fg)00078668
chr3L	24876774	24881774	12.04	12.5	0.45	0.000126	0.00197	promoter-TSS (Fg)0036290)	promoter-TSS (NM) 1409559)	463	Fg)0003303 Fg)00080470
chr3L	3937452	3937952	11.82	12.18	0.45	0.000394	0.00439	TTS (Fg)0035475)	TTS (NM) 1395582)	308	Fg)0003303 Fg)00073323
chr3L	22877459	22877959	11.08	11.54	0.45	0.000344	0.00444	exon (Fg)0004424, exon 1 of 3)	exon (NM) 162976, exon 3 of 5)	475	Fg)0003303 Fg)00078570
chr3R	29123473	29123973	11.35	11.81	0.45	0.000344	0.00444	promoter-TSS (Fg)0039637, exon 1 of 3)	promoter-TSS (NM) 057545)	308	Fg)0003303 Fg)00085372
chr3R	16435923	16436423	11.65	12.1	0.45	0.000473	0.0057	TTS (Fg)0038942)	TTS (NM) 001038969)	424	Fg)0003302 Fg)00083316
chr3R	19863776	19864276	11.91	12.37	0.45	0.000514	0.0061	promoter-TSS (NM) 142580)	promoter-TSS (NM) 142580)	367	Fg)0003303 Fg)00083366
chr3L	19261554	19262054	11.62	12.07	0.45	0.000584	0.00668	0.00668	TTS (NM) 140840)	411	Fg)0003302 Fg)00074289
chrX	17421350	17421850	11.77	12.22	0.45	0.00105	0.0102	TTS (Fg)0040677)	TTS (NM) 144231)	545	Fg)0003302 Fg)00074440
chr3L	13929484	13929984	12.13	12.58	0.45	0.00115	0.0109	promoter-TSS (Fg)0036386)	promoter-TSS (NM) 140412)	348	Fg)0003303 Fg)00075756
chr2R	13941505	13942005	10.89	11.34	0.45	0.00126	0.0116	TTS (Fg)003885)	TTS (NM) 133073)	355	Fg)0003303 Fg)00076136
chr2R	19310057	19310557	11.5	11.95	0.45	0.00141	0.0126	TTS (NM) 001274753)	TTS (NM) 001274753)	358	Fg)0003303 Fg)00080538
chr3L	14001155	14001655	11.01	11.46	0.45	0.0018	0.0153	promoter-TSS (Fg)0036395)	promoter-TSS (NM) 1404294)	406	Fg)0003303 Fg)00083025
chr2L	18709003	18709503	11.57	12.02	0.45	0.00192	0.0161	promoter-TSS (Fg)0025605)	promoter-TSS (NM) 058040)	408	Fg)0003303 Fg)00081099
chr2R	13197504	13198004	11.82	12.27	0.45	0.00208	0.0171	exon (Fg)00053182, exon 2 of 7)	5' UTR (NM) 001169668, exon 2 of 7)	583	Fg)0003303 Fg)00087770
chr3R	13970746	13971246	11.08	11.53	0.45	0.00223	0.0181	promoter-TSS (Fg)0036314)	promoter-TSS (NM) 048388)	-7.86	Fg)0003304 Fg)00080198
chr2L	7026597	7027097	11.82	12.27	0.45	0.00251	0.0198	promoter-TSS (Fg)0031881)	promoter-TSS (NM) 135249)	300	Fg)0003303 Fg)00079429
chr3R	5613441	5613941	11.92	12.37	0.45	0.00261	0.0204	promoter-TSS (Fg)0017550)	promoter-TSS (NM) 1395090)	534	Fg)0003303 Fg)00078704
chr3L	7348360	7348860	11.38	11.83	0.45	0.00271	0.0211	promoter-TSS (Fg)0035298)	promoter-TSS (NM) 057778)	288	Fg)0003303 Fg)00079445
chr3L	19792989	19793489	10.72	11.17	0.45	0.00273	0.0211	promoter-TSS (Fg)00261283)	promoter-TSS (NM) 168816)	430	Fg)0003303 Fg)00080278
chr2R	20622767	20623267	12.14	12.6	0.45	0.0031	0.0234	promoter-TSS (Fg)0034529)	promoter-TSS (NM) 137652)	312	Fg)0003303 Fg)00080824
chr3R	29357678	29358178	12.1	12.55	0.45	0.00322	0.024	promoter-TSS (Fg)0030132)	promoter-TSS (NM) 142941)	463	Fg)0003304 Fg)00084511
chr2R	21333212	21333712	11.74	12.19	0.45	0.00352	0.0256	exon (Fg)00259876, exon 1 of 5)	exon (NM) 001169664, exon 1 of 5)	321	Fg)0003303 Fg)00089507
chr3L	15836660	15837160	10.84	11.29	0.45	0.00387	0.0275	TTS (Fg)0052150)	TTS (NM) 168635)	572	Fg)0003303 Fg)0008654
chrX	5916001	5916501	10.7	11.15	0.45	0.00406	0.0283	promoter-TSS (Fg)0029823)	promoter-TSS (NM) 001144690)	66	Fg)0003303 Fg)000845317
chr2L	443872	443872	11.78	12.23	0.45	0.00634	0.0388	promoter-TSS (Fg)0051974)	promoter-TSS (NM) 001272671)	300	Fg)0003303 Fg)0008136
chrX	10745555	10746055	11.86	12.31	0.45	0.00674	0.0407	promoter-TSS (Fg)0027675)	promoter-TSS (NM) 001272497)	352	Fg)0003303 Fg)00074194
chr2R	15402326	15402826	11.12	11.57	0.45	0.00682	0.0411	exon (Fg)0009421, exon 1 of 5)	exon (NM) 137735, exon 1 of 5)	548	Fg)0003303 Fg)00084932
chr2R	9782678	9783178	11.39	11.84	0.45	0.00722	0.0439	exon (Fg)0003317, exon 1 of 4)	exon (NM) 135469, exon 1 of 4)	545	Fg)0003303 Fg)00079589
chr3R	26965818	26966318	11.24	11.7	0.45	0.00793	0.0462	promoter-TSS (Fg)0041224)	promoter-TSS (NM) 178570)	430	Fg)0003303 Fg)00080824
chrX	17710050	17710550	10.13	10.58	0.45	0.00826	0.0475	promoter-TSS (Fg)0038964)	promoter-TSS (NM) 133011)	157	Fg)0003303 Fg)00080325
chrX	5419768	5420268	11.36	11.81	0.45	0.00877	0.0495	promoter-TSS (Fg)0029877)	promoter-TSS (NM) 167033)	318	Fg)0003303 Fg)00070795
chr2L	2780006	2780506	12.03	12.48	0.45	0.00889	0.0499	promoter-TSS (Fg)0031455)	promoter-TSS (NM) 001032050)	353	Fg)0003303 Fg)00077125
chr3L	22071345	22071845	11.36	11.82	0.45	0.00947	0.0549	promoter-TSS (Fg)00364727)	promoter-TSS (NM) 073073)	233	Fg)0003303 Fg)00078486
chr3R	8032590	8033090	11.22	11.67	0.45	0.010025	0.0549	promoter-TSS (NM) 0010310812)	promoter-TSS (NM) 0010316496)	169	Fg)0003304 Fg)00084760
chr2R	18706189	18706689	11.26	11.71	0.46	8.83E-05	0.00144	TTS (Fg)0026103)	TTS (NM) 001103899)	463	Fg)0003303 Fg)00086607
chrX	6547653	6547653	11.57	12.03	0.46	0.00138	0.00216	exon (Fg)0002987, exon 1 of 1)	exon (NM) 163029, exon 1 of 1)	601	Fg)0003303 Fg)00079085
chrX	15784785	15784785	11.89	12.35	0.46	0.000239	0.00342	exon (Fg)0002580, exon 1 of 2)	exon (NM) 167460, exon 1 of 2)	341	Fg)0003303 Fg)00071131
chr2R	24611874	24612374	11.43	11.89	0.46	0.000282	0.00386	promoter-TSS (Fg)0035040)	promoter-TSS (NM) 001299221)	243	Fg)0003303 Fg)00084487
chr3L	12571810	12572310	11.48	11.93	0.46	0.000425	0.00526	exon (Fg)00036302, exon 2 of 13)	exon (NM) 001144471, exon 2 of 13)	482	Fg)0003303 Fg)00080887
chr2L	21237280	21237780	10.83	11.29	0.46	0.00075	0.00802	promoter-TSS (Fg)0051626)	promoter-TSS (NM) 165362)	281	Fg)0003303 Fg)00081478
chr3L	10637255	10637755	11.22	11.68	0.46	0.000786	0.00829	promoter-TSS (NM) 168415)	promoter-TSS (NM) 168415)	314	Fg)0003303 Fg)00081738
chr2R	14993315	14993815	11.45	11.9	0.46	0.00127	0.0117	promoter-TSS (Fg)00264378)	promoter-TSS (NM) 073877)	250	Fg)0003303 Fg)000832259
chr3R	16955558	16956058	11.83	12.3	0.46	0.00134	0.0122	promoter-TSS (NM) 001103882)	promoter-TSS (NM) 001103882)	371	Fg)0003303 Fg)000839071
chr3R	19904120	19904620	11.8	12.25	0.46	0.00134	0.0136	promoter-TSS (Fg)0038745)	promoter-TSS (NM) 163890)	498	Fg)0003303 Fg)00080594
chr3R	23758424	23758924	11.7	12.16	0.46	0.00157	0.0164	promoter-TSS (Fg)0029311)	promoter-TSS (NM) 170089)	67	Fg)0003303 Fg)00084546
chr2L	12509749	12510249	10.93	11.39	0.46	0.00327	0.0243	promoter-TSS (Fg)0027264)	promoter-TSS (NM) 001293533)	63	Fg)0003303 Fg)000845021
chr2L	10423554	10424054	12.12	12.58	0.46	0.00393	0.0278	exon (Fg)0003246, exon 3 of 7)	exon (NM) 135556, exon 2 of 6)	471	Fg)0003303 Fg)00080020
chr2R	12515492	12515992	11.64	12.1	0.46	0.00395	0.0278	promoter-TSS (Fg)0003349)	promoter-TSS (NM) 165912)	308	Fg)0003303 Fg)00080939
chrX	2609322	2609372	11.61	12.07	0.46	0.00641	0.0392	promoter-TSS (Fg)0004643)	promoter-TSS (NM) 080129)	331	Fg)0003303 Fg)00084003
chr2L	15270835	15271335	10.85	11.31	0.46	0.00705	0.0421	promoter-TSS (NM) 165112)	promoter-TSS (NM) 165112)	543	Fg)0003303 Fg)000820158
chr3L	22263471	22263971	11.82	12.28	0.46	0.00772	0.045	promoter-TSS (Fg)0027800)	promoter-TSS (NM) 001275276)	353	Fg)0003303 Fg)00078492
chr2L	10412870	10413370	11.24	11.7	0.46	0.00797	0.0483	promoter-TSS (Fg)0032943)	promoter-TSS (NM) 001169466)	343	Fg)0003303 Fg)00080047
chr3L	9722587	9723087	11.73	12.19	0.46	0.00812	0.0489	promoter-TSS (Fg)0036303)	promoter-TSS (NM) 140079)	357	Fg)0003303 Fg)00080639
chr3R	18347085	18347585	11.27	11.73	0.46	0.00825	0.0493	exon (Fg)0007275, exon 2 of 4)	promoter-TSS (NM) 00127143)	347	Fg)0003303 Fg)00081571
chr2R	10051544	10052044	11.58	12.04	0.47	2.09E-05	0.00497	promoter-TSS (Fg)0028169)	promoter-TSS (NM) 001293378)	313	Fg)0003303 Fg)00080833
chr2R	17027772	17028272	11.37	11.84	0.47	2.45E-05	0.00539	promoter-TSS (Fg)0026540)	promoter-TSS (NM) 134587)	333	Fg)0003303 Fg)00080690
chrX	7965754	7966254	11.02	11.49	0.47	2.60E-05	0.00563	exon (Fg)0003656, exon 1 of 11)	5' UTR (NM) 167140, exon 1 of 11)	412	Fg)0003303 Fg)000812125
chr3R	16646364	16646864	11.06	11.53	0.47	3.68E-05	0.00749	promoter-TSS (Fg)0015019)	promoter-TSS (NM) 169727)	390	Fg)0003303 Fg)00083356
chr3L	9863660	9864160	11.21	11.68	0.47	3.89E-05	0.00783	promoter-TSS (Fg)0002767)	promoter-TSS (NM) 168386)	439	Fg)0003303 Fg)00076362
chrX	16287975	16288475	11.01	11.49	0.47	4.59E-05	0.0104	promoter-TSS (Fg)0011571)	promoter-TSS (NM) 001272673)	308	Fg)0003303 Fg)00084245
chr3L	7347919	7348419	11.61	12.08	0.47	5.82E-05	0.0109	promoter-TSS (NM) 001169699)	promoter-TSS (NM) 001169699)	446	Fg)0003303 Fg)00076880
chr3R	24134988	24135488	11.62	12.09	0.47	5.91E-05	0.0111	TTS (Fg)00026317)	TTS (NM) 050670)	459	Fg)0003303 Fg)00084547
chr3R	18908035	18908535	11.37	11.83	0.47	6.89E-05	0.0117	promoter-TSS (Fg)0008602)	promoter-TSS (NM) 140815)	368	Fg)0003303 Fg)000757076
chrX	477461	477961	11.17	11.63	0.47	7.52E-05	0.0119	exon (NM) 001103888, exon 2 of 3)	exon (NM) 001103888, exon 2 of 3)	344	Fg)0003303 Fg)00070795
chr3R	10761423	10761923	11.68	12.16	0.47	8.04E-05	0.014	promoter-TSS (Fg)0051390)	promoter-TSS (NM) 204721)	-8	Fg)0003303 Fg)00080273
chr2L	1704977	1705477	11.53	11.99	0.47	0.000176	0.00268	promoter-TSS (Fg)0011570)	promoter-TSS (NM) 001272948)	94	Fg)0003303 Fg)000833130
chr3L	24851805	24852305	11.71	12.18	0.47	0.000182	0.00276	TTS (Fg)00039233)	TTS (NM) 143039)	455	Fg)0003303 Fg)00080470
chr3L	11122718	11123218	11.84	12.31	0.47	0.000283	0.00386	promoter-TSS (Fg)00260795)	promoter-TSS (NM) 140194)	248	Fg)0003303 Fg)00076223
chr3L	1517965	1518465	11.77	12.24	0.47	0.000287	0.00391	promoter-TSS (NM) 138269)	promoter-TSS (NM) 138269)	363	Fg)0003303 Fg)00081072
chr3R	30056983	30057483	11.92	12.39	0.47	0.000317	0.00423	TTS (Fg)00039744)	TTS (NM) 143512)	3	

chr3R	8806560	8807060	11.52	12.01	0.49	4.53E-06	0.000139	TIS (#Rgn037622)	TIS (NM_141578)	489	Fgpn037643 Fba0081919
chr3R	7919439	7919939	11.01	11.51	0.49	7.23E-06	0.000201	paramoter-TSS (#Rgn002542)	paramoter-TSS (NM_080111)	422	Fgpn037531 Fba0300625
chr3L	8340495	8341495	11.41	11.91	0.49	8.22E-06	0.000221	TIS (NM_139993)	TIS (NM_139993)	461	Fgpn037567 Fba0076004
chrX	54093	54593	11.44	11.94	0.49	1.58E-05	0.000036	Intergenic	HE A LINE Jockey	682	Fgpn058469 Fba0307454
chr3L	13411941	13412441	11.67	12.16	0.49	1.74E-05	0.000046	paramoter-TSS (#Rgn026049)	paramoter-TSS (NM_169543)	2619	Fgpn036566 Fba0309341
chr2R	15577076	15577576	11.61	12.11	0.49	1.87E-05	0.000033	TIS (#Rgn026189)	TIS (NM_142248)	344	Fgpn064746 Fba0082538
chr3L	19251015	19260425	11.22	11.71	0.49	2.15E-05	0.000047	exon (#Rgn0010315, exon 2 of 4)	exon (NM_167473, exon 2 of 4)	138	Fgpn064746 Fba0082538
chrX	15508598	15509098	11.54	12.03	0.49	3.18E-05	0.000065	exon (#Rgn0010315, exon 2 of 6)	exon (NM_167473, exon 2 of 5)	828	Fgpn0103815 Fba00300391
chr2R	10880200	10880700	11.18	11.67	0.49	5.68E-05	0.000107	paramoter-TSS (#Rgn010356)	paramoter-TSS (NM_057609)	477	Fgpn033564 Fba00289471
chr3R	8663253	8663753	11.88	12.37	0.49	6.81E-05	0.000124	paramoter-TSS (#Rgn0127606)	paramoter-TSS (NM_141554)	396	Fgpn086669 Fba0081887
chr3L	5751294	5751794	11.64	12.13	0.49	6.89E-05	0.000126	exon (#Rgn0263106, exon 1 of 3)	exon (NM_168128, exon 1 of 1)	673	Fgpn0263106 Fba0071233
chr3R	8348422	8348922	11.13	11.62	0.49	7.53E-05	0.000133	TIS (#Rgn014536)	TIS (NM_141536)	645	Fgpn033564 Fba0081887
chr3R	27112672	27113172	11.85	12.34	0.49	8.94E-05	0.000154	paramoter-TSS (#Rgn039487)	paramoter-TSS (NM_13136513)	352	Fgpn027564 Fba0085106
chr2R	11452320	11452820	11.18	11.67	0.49	9.77E-05	0.000166	exon (#Rgn0000581, exon 2 of 10)	5'UTR (NM_001273968, exon 2 of 11)	207	Fgpn0000581 Fba0085106
chr3R	16412818	16413318	11.63	12.12	0.49	0.00017	0.000261	paramoter-TSS (#Rgn028429)	paramoter-TSS (NM_142956)	325	Fgpn033849 Fba0082899
chr3R	27273763	27274263	11.99	12.47	0.49	0.000233	0.000497	paramoter-TSS (#Rgn046247)	paramoter-TSS (NM_01104477)	309	Fgpn039255 Fba0085158
chr2R	2356490	2356990	11.33	11.81	0.49	0.000441	0.00041	paramoter-TSS (#Rgn018950)	paramoter-TSS (NM_01160791)	651	Fgpn034891 Fba00129860
chr2R	9290949	9291449	11.61	12.11	0.49	0.000453	0.00054	paramoter-TSS (#Rgn028408)	paramoter-TSS (NM_136624)	366	Fgpn0262626 Fba0088599
chr3R	25292521	25293021	10.92	11.41	0.49	0.000453	0.00054	TIS (#Rgn051102)	TIS (NM_120213)	206	Fgpn051089 Fba0084863
chr2R	20644166	20644666	11.82	12.31	0.49	0.000604	0.000687	paramoter-TSS (#Rgn061361)	paramoter-TSS (NM_01032273)	381	Fgpn043587 Fba0086232
chr3R	14662572	14663072	11.77	12.26	0.49	0.000665	0.000744	paramoter-TSS (#Rgn038252)	paramoter-TSS (NM_142126)	375	Fgpn025808 Fba0082974
chr2L	20797873	20798373	10.44	10.92	0.49	0.000731	0.000792	exon (#Rgn002886, exon 1 of 3)	5'UTR (NM_165344, exon 1 of 3)	209	Fgpn032886 Fba0081415
chr3L	16419643	16420143	11.95	12.44	0.49	0.000791	0.000859	exon (#Rgn0027080)	paramoter-TSS (NM_140638)	497	Fgpn033849 Fba0085106
chr2L	4459825	4460325	11.88	12.37	0.49	0.00103	0.00119	paramoter-TSS (#Rgn028929)	paramoter-TSS (NM_011273099)	276	Fgpn034891 Fba0077444
chr2R	23363474	23363974	11.44	11.93	0.49	0.00109	0.0105	exon (#Rgn0266011)	exon (NM_148508, exon 1 of 5)	788	Fgpn0404860 Fba0082233
chr3R	23762634	23763134	12.03	12.52	0.49	0.00109	0.0105	paramoter-TSS (#Rgn031114)	paramoter-TSS (NM_142926)	370	Fgpn033114 Fba0084664
chr2R	8950319	8950819	12.04	12.53	0.49	0.0011	0.0111	exon (#Rgn003356, exon 2 of 3)	exon (NM_165638, exon 2 of 2)	517	Fgpn003356 Fba00307344
chr2L	4692020	4692520	12.17	12.66	0.49	0.00128	0.0118	paramoter-TSS (#Rgn_135040)	paramoter-TSS (NM_135040)	467	Fgpn031657 Fba0087994
chr3R	24187998	24188498	11.93	12.43	0.49	0.00149	0.0132	TIS (#Rgn029503)	TIS (NM_142971)	515	Fgpn039836 Fba0084525
chr2R	17733502	17734002	11.45	11.94	0.49	0.00168	0.0145	TIS (#Rgn034266)	TIS (NM_001169212)	513	Fgpn052554 Fba0112388
chr3L	15512053	15512553	10.52	11.01	0.49	0.00183	0.0155	exon (#Rgn0000565, exon 2 of 4)	exon (NM_206731, exon 2 of 4)	415	Fgpn0000565 Fba0180176
chr3R	29862066	29862566	11.79	12.28	0.49	0.00203	0.0185	paramoter-TSS (#Rgn039714)	paramoter-TSS (NM_170444)	452	Fgpn033849 Fba0081088
chrX	91427768	91432768	11.69	12.18	0.49	0.00232	0.0246	exon (#Rgn0000014, exon 2 of 3)	exon (NM_132931, exon 2 of 3)	526	Fgpn0000014 Fba0081299
chr2L	18931434	18931934	10.97	11.47	0.49	0.00244	0.0302	paramoter-TSS (#Rgn032125)	exon (NM_148508, exon 1 of 5)	456	Fgpn0404860 Fba0082233
chr3L	21633723	21634223	12.23	12.71	0.49	0.00268	0.0417	exon (#Rgn0037109, exon 2 of 4)	exon (NM_143084, exon 2 of 4)	473	Fgpn0037109 Fba0087864
chr3R	17671866	17672366	9.75	10.24	0.49	0.00272	0.0433	exon (#Rgn0000063, exon 2 of 4)	exon (NM_142887, exon 2 of 4)	846	Fgpn0000063 Fba0083478
chr3R	1645267	1645767	11.34	11.84	0.5	2.13E-06	7.80E-05	TIS (#Rgn038413)	TIS (NM_001043252)	466	Fgpn038413 Fba0081291
chr2R	17569013	17569513	11.44	11.94	0.5	3.06E-06	0.000102	paramoter-TSS (#Rgn034243)	paramoter-TSS (NM_137388)	532	Fgpn0402494 Fba0086228
chr3R	10045008	10046008	11.61	12.11	0.5	3.50E-06	0.000113	paramoter-TSS (#Rgn037777)	paramoter-TSS (NM_141713)	385	Fgpn037777 Fba0082173
chr2L	20925542	20926042	11.39	11.88	0.5	4.94E-06	0.000147	TIS (#Rgn029077)	TIS (NM_136234)	105	Fgpn032966 Fba0081438
chr3R	1432892	1432892	11.55	12.04	0.5	5.10E-06	0.000145	TIS (#Rgn038224)	TIS (NM_162659)	292	Fgpn038224 Fba0082233
chr3L	1323660	1323710	11.57	12.06	0.5	5.52E-06	0.000051	TIS (#Rgn038402)	TIS (NM_00134771)	460	Fgpn038402 Fba0087585
chr2R	13593843	13594343	11.71	12.22	0.5	6.66E-06	0.000227	paramoter-TSS (#Rgn027581)	paramoter-TSS (NM_001274029)	450	Fgpn0338519 Fba0087623
chr2R	8168111	8168611	11.21	11.71	0.5	9.76E-06	0.000253	TIS (#Rgn021995)	TIS (NM_141713)	314	Fgpn033273 Fba0086789
chr3L	20426982	20427482	11.35	11.85	0.5	2.82E-05	0.000604	TIS (#Rgn026988)	TIS (NM_140966)	492	Fgpn026988 Fba0087232
chr3R	5606001	5606501	11.69	12.18	0.5	2.90E-05	0.000614	exon (#Rgn0261004, exon 2 of 3)	exon (NM_141300, exon 2 of 3)	282	Fgpn0261004 Fba0087743
chr3R	10043306	10043806	11.57	12.07	0.5	6.07E-05	0.000112	paramoter-TSS (#Rgn035208)	paramoter-TSS (NM_001316573)	297	Fgpn035208 Fba0087174
chr3L	19062658	19063158	11.37	11.87	0.5	7.19E-05	0.000129	exon (#Rgn0036846, exon 1 of 3)	exon (NM_079300, exon 1 of 2)	239	Fgpn0036846 Fba0087506
chr2R	12985752	12986252	11.63	12.13	0.5	0.000103	0.000172	paramoter-TSS (#Rgn037895)	paramoter-TSS (NM_176160)	293	Fgpn033849 Fba0087803
chr2L	2710109	2710509	10.75	11.25	0.5	0.000134	0.000279	TIS (#Rgn068912)	TIS (NM_134667)	784	Fgpn033849 Fba0087840
chr2R	2771274	2771774	12.38	12.87	0.5	0.000183	0.000295	paramoter-TSS (#Rgn028855)	paramoter-TSS (NM_001272882)	211	Fgpn028855 Fba0081904
chr3R	9712774	9713274	11.46	11.96	0.5	0.00048	0.00077	exon (#Rgn010316, exon 1 of 3)	exon (NM_057400, exon 1 of 3)	405	Fgpn010316 Fba0088469
chr3R	16433085	16433585	11.86	12.36	0.5	0.000686	0.00758	TIS (#Rgn051272)	TIS (NM_001120159)	338	Fgpn0384241 Fba0083286
chr3L	1731078	1731578	10.49	10.99	0.5	0.000801	0.00838	paramoter-TSS (#Rgn_167911)	paramoter-TSS (NM_167911)	-221	Fgpn027914 Fba0087278
chr3R	21130471	21130971	11.63	12.13	0.5	0.00106	0.0103	exon (#Rgn038860, exon 1 of 1)	exon (NM_142689, exon 1 of 1)	507	Fgpn038860 Fba0084015
chr3L	11065187	11065687	11.41	11.91	0.5	0.00116	0.0109	paramoter-TSS (#Rgn052075)	paramoter-TSS (NM_140171)	446	Fgpn052075 Fba0087603
chr2R	16287275	16287775	10.93	11.43	0.5	0.00118	0.011	paramoter-TSS (#Rgn040079)	paramoter-TSS (NM_135939)	280	Fgpn040079 Fba0083814
chr2R	14161465	14161965	11.68	12.18	0.5	0.00135	0.0122	paramoter-TSS (#Rgn033900)	paramoter-TSS (NM_137086)	296	Fgpn033901 Fba0087520
chr2L	943050	9431000	10.25	10.75	0.5	0.0014	0.0126	paramoter-TSS (#Rgn029177)	paramoter-TSS (NM_080166)	302	Fgpn029177 Fba0084623
chr2R	20618774	20619274	11.39	11.89	0.5	0.00128	0.0129	exon (#Rgn0031814)	exon (NM_164460)	393	Fgpn031814 Fba0082625
chr3R	9400179	9400279	11.59	12.09	0.5	0.00279	0.0214	paramoter-TSS (#Rgn032686)	paramoter-TSS (NM_169277)	375	Fgpn032686 Fba0082029
chrX	4319756	4320256	11.8	12.3	0.5	0.00323	0.0241	exon (#Rgn0266570, exon 2 of 2)	exon (NM_167007, exon 2 of 2)	560	Fgpn0266570 Fba0087048
chr3R	22042895	22043395	11.1	11.6	0.5	0.003	0.0244	exon (NM_001120213, exon 2 of 9)	exon (NM_164449 Fba0084178)	959	Fgpn026449 Fba0084178
chr2L	1048062	1048562	11.99	12.49	0.5	0.00348	0.0254	paramoter-TSS (#Rgn032259)	paramoter-TSS (NM_001038806)	409	Fgpn032258 Fba0080038
chr2L	10427068	10427568	12.28	12.78	0.5	0.00366	0.0263	paramoter-TSS (#Rgn_00120181)	paramoter-TSS (NM_00120181)	312	Fgpn0267310 Fba0080021
chrX	16654426	16654926	12.01	12.51	0.5	0.00464	0.031	paramoter-TSS (#Rgn003189)	paramoter-TSS (NM_078653)	309	Fgpn003189 Fba0087426
chrX	21027902	21028402	11	11.5	0.5	0.00478	0.0316	TIS (NM_001144751)	TIS (NM_001144751)	239	Fgpn031144751 Fba0083933
chr3L	1546400	1546800	11.61	12.12	0.5	0.00495	0.0329	paramoter-TSS (#Rgn025913)	paramoter-TSS (NM_167894)	429	Fgpn0404860 Fba0087278
chr2R	17682113	17682613	11.17	11.67	0.5	0.00511	0.037	exon (NM_164286)	exon (NM_164286)	-21	Fgpn038402 Fba0086821
chr2L	6921813	6922313	10.7	11.2	0.5	0.00718	0.0427	intron (#Rgn004838, intron 1 of 4)	intron (NM_164286, intron 1 of 4)	475	Fgpn004838 Fba0091404
chr2R	10478058	10478558	10.31	10.81	0.5	0.00823	0.0473	TIS (#Rgn039528)	TIS (NM_136747)	259	Fgpn0261862 Fba0088307
chr2R	17444193	17444693	11.59	12.09	0.5	0.0087	0.0491	TIS (NM_176214)	TIS (NM_176214)	495	Fgpn027564 Fba0086958
chr2R	9599977	9600477	11.45	11.95	0.51	3.40E-06	0.000111	paramoter-TSS (#Rgn033467)	paramoter-TSS (NM_136694)	230	Fgpn0278301 Fba0082922
chr3R	5623641	5624141	11.03	11.54	0.51	4.81E-06	0.000146	paramoter-TSS (#Rgn037351)	paramoter-TSS (NM_141303)	154	Fgpn0448231 Fba0087239
chr3R	22025026	22025526	11.09	11.61	0.51	5.28E-06	0.000155	TIS (#Rgn038923)	TIS (NM_142746)	430	Fgpn038924 Fba0084186
chr3R	16650490	16650990	11.47	11.98	0.51	6.03E-06	0.000172	TIS (NM_169772)	TIS (NM_169772)	217	Fgpn0015831 Fba0083846
chr2L	10387013	10387513	11.49	11.99	0.51	8.29E-06	0.000221	paramoter-TSS (#Rgn0267821)	paramoter-TSS (NM_001273412)	602	Fgpn032726 Fba0080007
chr2R	15216874	15217									

ch2L	568439	568939	12.09	11.6	0.52	0.00659	0.0389	promoter-TSS #Bp0021874)	promoter-TSS (NM_057946)	350	Fb00031261 Fb00078068
chrX	18404980	18408990	10.92	11.44	0.52	0.00802	0.0465	intron (Bp0057524), intron 1 of 9)	intron (NM_206276, intron 1 of 9)	350	Fb00052594 Fb00074570
chr2L	16943220	16949730	10.88	11.41	0.53	0.00336	0.0278	TIS (Bp00336357)	TIS (NM_164343)	351	Fb00075043
chr3R	17373772	17374272	11.96	12.48	0.52	0.00841	0.048	TIS (Bp0023762)	TIS (NM_00112041)	404	Fb00027368 Fb00078220
chr2L	2740373	2740743	11.22	11.75	0.53	9.39E-07	3.93E-05	promoter-TSS #Bp00031450)	promoter-TSS (NM_164497)	371	Fb00010469 Fb00082949
chr3L	13966631	13967131	10.94	11.47	0.53	2.48E-06	8.66E-05	exon (Bp00263232, exon 2 of 5)	exon (NM_162566, exon 2 of 5)	404	Fb00036389 Fb00075599
chrX	19605348	19605848	11.05	11.58	0.53	5.75E-06	0.000165	exon (Bp0001037, exon 1 of 4)	exon (NM_167669, exon 1 of 5)	371	Fb00010733 Fb00074733
chr3L	3466574	3467074	11.49	12.03	0.53	9.25E-06	0.000242	TIS (Bp00041164)	TIS (NM_00014556)	397	Fb00003317 Fb00073127
chr3L	19843655	19843655	11.55	12.08	0.53	1.34E-05	0.000327	promoter-TSS #Bp00037444)	promoter-TSS (NM_079446)	357	Fb00052271 Fb00030108
chrX	16942262	16942762	11.87	12.4	0.53	2.64E-05	0.000569	promoter-TSS #Bp00032808)	promoter-TSS (NM_001042817)	146	Fb00030016 Fb00074417
chr2R	6015060	6015560	11.35	11.89	0.53	2.65E-05	0.000571	exon (Bp0000034, exon 2 of 2)	exon (NM_078931, exon 2 of 2)	450	Fb00000134 Fb00086029
chr2L	17431148	17431648	11.37	11.9	0.53	0.000293	0.000625	5' UTR (NM_136673, exon 1 of 11)	5' UTR (NM_136673, exon 1 of 11)	375	Fb00027233 Fb00042999
chr2R	24767202	24767702	11.06	11.6	0.53	3.08E-05	0.000671	TIS (Bp00025860)	TIS (NM_166867)	259	Fb00027590 Fb00093426
chr3L	12193004	12193504	11.68	12.21	0.53	0.00012	0.00193	TIS (Bp0000594)	TIS (NM_176323)	434	Fb00036313 Fb00031611
chr2R	22421070	22421570	11.5	12.04	0.53	0.00016	0.00248	promoter-TSS #Bp00034727)	promoter-TSS (NM_079087)	308	Fb00039728 Fb00036468
chr2L	13829867	13830367	11.79	12.32	0.53	0.00023	0.00333	promoter-TSS #Bp0001961)	promoter-TSS (NM_165063)	401	Fb00052717 Fb00080543
chr3R	9015608	9015608	11.59	12.12	0.53	0.000237	0.00341	promoter-TSS #Bp0007655)	promoter-TSS (NM_001370091)	413	Fb00027656 Fb00083167
chr3L	18819945	18820445	10.96	11.49	0.53	0.00039	0.00494	promoter-TSS #Bp00036815)	promoter-TSS (NM_168777)	480	Fb00036814 Fb00052084
chr2L	7222445	7222445	12.08	12.61	0.53	0.000527	0.00621	TIS (Bp000259111)	TIS (NM_001272524)	386	Fb00027313 Fb00031913
chrX	2704802	2705302	12.06	12.59	0.53	0.000554	0.00642	exon (Bp00024936, exon 3 of 5)	exon (NM_130664, exon 3 of 5)	417	Fb00004936 Fb00070479
chr2L	9750963	9751463	11.88	12.41	0.53	0.000576	0.00693	promoter-TSS #Bp00042174)	promoter-TSS (NM_137365)	282	Fb00024249 Fb00072870
chr3L	85901	859401	11.76	12.3	0.53	0.00109	0.0104	TIS (Bp00067055)	TIS (NM_138217)	467	Fb00020761 Fb00030123
chr2R	6613924	6614424	11.5	12.08	0.53	0.00122	0.0113	exon (Bp00086655, exon 2 of 2)	exon (NM_001103730, exon 2 of 2)	409	Fb00056728 Fb00030136
chr2R	1072458	1072458	11.19	11.72	0.53	0.00137	0.0124	promoter-TSS #Bp0026345)	promoter-TSS (NM_124506)	161	Fb00052144 Fb00082854
chr2L	13219080	13219580	11.58	12.11	0.53	0.00168	0.0145	promoter-TSS #Bp00032486)	promoter-TSS (NM_135773)	593	Fb00024851 Fb00080254
chr3R	29033410	29033910	12	12.53	0.53	0.00346	0.0254	promoter-TSS #Bp00027873)	promoter-TSS (NM_134301)	348	Fb00036259 Fb00082519
chr2L	1610735	1611235	11.37	11.91	0.53	0.00366	0.0263	promoter-TSS #Bp0001951)	promoter-TSS (NM_134765)	172	Fb00050195 Fb00072916
chr2R	9139167	9139667	10.55	11.08	0.53	0.00367	0.0269	exon (Bp00265313, exon 2 of 3)	5' UTR (NM_001299303, exon 2 of 3)	356	Fb00076533 Fb00044468
chr3R	25126967	25127467	11.24	11.74	0.53	0.00392	0.0277	promoter-TSS #Bp00259152)	promoter-TSS (NM_001170250)	142	Fb00038280 Fb00094822
chr3L	15837321	15837821	11.08	11.6	0.53	0.00507	0.0328	promoter-TSS #Bp00017567)	promoter-TSS (NM_079989)	358	Fb00024483 Fb00031183
chr2L	4882066	4882566	11.59	12.12	0.53	0.00597	0.0372	promoter-TSS #Bp00031643)	promoter-TSS (NM_135027)	791	Fb00020286 Fb00030488
chr3R	6619823	6619823	12.01	12.54	0.53	0.00724	0.0429	exon (Bp00005885, exon 3 of 4)	exon (NM_079569, exon 3 of 4)	728	Fb00050585 Fb00033461
chr2R	24786678	24787178	10.96	11.5	0.54	1.71E-06	6.45E-05	exon (Bp00041205, exon 2 of 6)	exon (NM_079132, exon 2 of 6)	443	Fb00041205 Fb00072428
chrX	496681	497181	11.47	12.01	0.54	2.89E-06	9.79E-05	TIS (Bp00023234)	TIS (NM_130490)	483	Fb00026316 Fb00034793
chr3L	18901669	18902169	11.73	12.27	0.54	3.01E-06	0.000101	promoter-TSS #Bp00036826)	promoter-TSS (NM_140814)	326	Fb00036827 Fb00075067
chr2L	7800638	7801138	11.17	11.71	0.54	3.17E-06	0.000105	promoter-TSS #Bp00031950)	promoter-TSS (NM_135300)	507	Fb00031951 Fb00030129
chr2R	13230559	13231059	11.38	11.92	0.54	3.90E-06	0.000124	promoter-TSS #Bp00038812)	promoter-TSS (NM_133006)	227	Fb00038813 Fb00087744
chr3R	17305148	17305648	11.24	11.77	0.54	4.02E-06	0.000124	promoter-TSS #Bp00038065)	promoter-TSS (NM_142227)	722	Fb00025894 Fb00083374
chr3L	21317322	21317822	11.77	12.31	0.54	4.85E-05	0.000949	TIS (Bp00267234)	TIS (NM_125033)	472	Fb00010498 Fb00078405
chr3R	29933778	29934278	11.64	12.17	0.54	6.25E-05	0.00115	promoter-TSS #Bp00051040)	promoter-TSS (NM_134933)	342	Fb00020299 Fb00030126
chr3L	4826006	4826506	12.06	12.59	0.54	0.000131	0.00208	promoter-TSS #Bp00061797)	promoter-TSS (NM_168103)	404	Fb00027085 Fb00033461
chr2R	13509669	13510169	11.03	11.56	0.54	0.000161	0.00248	TIS (Bp00024556)	TIS (NM_001169673)	283	Fb00038166 Fb00030188
chr3R	8085066	8085566	11.17	11.71	0.54	0.000272	0.00377	exon (Bp00037633, exon 2 of 3)	exon (NM_141579, exon 1 of 2)	302	Fb00037633 Fb00081934
chr2R	8839142	8839642	11	11.55	0.54	0.000284	0.00388	exon (Bp00033354, exon 2 of 13)	exon (NM_136585, exon 2 of 13)	317	Fb00033354 Fb00080875
chr2R	8444632	8445132	10.32	10.86	0.54	0.00044	0.00541	exon (Bp00014463, exon 2 of 7)	exon (NM_048064)	397	Fb00026231 Fb00030480
chrX	11887458	11887958	11.99	12.53	0.54	0.000562	0.0065	promoter-TSS #Bp00030349)	promoter-TSS (NM_001272545)	472	Fb00040350 Fb00082570
chr3R	18301005	18301505	11.61	12.15	0.54	0.000582	0.00792	TIS (Bp00028602)	TIS (NM_142452)	423	Fb00028602 Fb00080254
chrX	19601253	19601753	11.89	12.43	0.54	0.000944	0.0094	exon (Bp00031036, exon 2 of 4)	exon (NM_134481, exon 2 of 4)	349	Fb00010316 Fb00070767
chr3R	28337137	28337637	10.68	11.22	0.54	0.00117	0.0109	promoter-TSS #Bp00261618)	promoter-TSS (NM_170365)	498	Fb00027773 Fb00034733
chr2R	6675204	6675704	11.55	12.08	0.54	0.00154	0.0136	promoter-TSS #Bp00033892)	promoter-TSS (NM_165484)	351	Fb00033891 Fb00030502
chr2R	20258031	20258531	11.29	11.84	0.54	0.00271	0.0209	promoter-TSS #Bp00025720)	promoter-TSS (NM_001274172)	289	Fb00030448 Fb00086272
chr2R	10820311	10820811	12.25	12.79	0.54	0.0043	0.0295	promoter-TSS #Bp00086898)	promoter-TSS (NM_165800)	404	Fb00029134 Fb00080224
chr2R	19496158	19496658	12.24	12.78	0.54	0.00456	0.0307	exon (Bp00026156, exon 1 of 2)	exon (NM_001274163, exon 1 of 2)	448	Fb00026156 Fb00031920
chrX	18095992	18100492	9.33	9.88	0.54	0.00512	0.033	promoter-TSS #Bp00030894)	promoter-TSS (NM_136760)	791	Fb00030894 Fb00030272
chr3R	6085268	6085768	10.24	10.78	0.54	0.00536	0.034	exon (Bp00032828, exon 4 of 4)	exon (NM_143623, exon 4 of 4)	256	Fb00032828 Fb00082882
chr2L	24556	24606	9.23	9.77	0.54	0.00556	0.0356	exon (Bp00017561, intron 1 of 2)	5' UTR (NM_078557, intron 1 of 2)	416	Fb00017561 Fb00073404
chrX	1557347	1557847	11.23	11.78	0.55	9.94E-07	4.10E-05	intron (Bp00066557, intron 1 of 16)	intron (NM_00128891, intron 1 of 16)	-2720	Fb00026252 Fb00030444
chr2R	9093883	9094383	11.44	11.98	0.55	4.33E-05	4.43E-05	promoter-TSS #Bp00030658)	promoter-TSS (NM_132816)	405	Fb00030657 Fb00080270
chr2R	7924303	7924803	11.24	11.79	0.55	1.06E-06	9.17E-05	exon (Bp00033374, exon 2 of 3)	exon (NM_176109, exon 2 of 3)	521	Fb00033374 Fb00080632
chr3R	18685977	18686477	11.8	12.35	0.55	2.67E-06	9.17E-05	intron (Bp00033233, intron 2 of 3)	intron (NM_00129265, intron 2 of 3)	365	Fb00033233 Fb00080886
chr2R	14871627	14872127	10.86	11.41	0.55	2.66E-06	9.17E-05	promoter-TSS #Bp00051294)	promoter-TSS (NM_169826)	430	Fb00027814 Fb00030360
chr3R	2715321	2715821	11.24	11.79	0.55	5.29E-06	0.000155	promoter-TSS #Bp00264962)	promoter-TSS (NM_001169683)	350	Fb00037733 Fb00087450
chr3L	14163736	14164236	11.74	12.29	0.55	8.66E-05	0.000381	TIS (Bp00029945)	TIS (NM_142452)	625	Fb00029945 Fb00030964
chrX	15984406	15984906	11.38	11.93	0.55	4.09E-05	0.000814	exon (Bp00030717, exon 3 of 4)	exon (NM_132858, exon 3 of 4)	515	Fb00030717 Fb00070134
chr3R	9555865	9556365	11.76	12.3	0.55	4.14E-05	0.000814	promoter-TSS #Bp00037708)	promoter-TSS (NM_141647)	313	Fb00027809 Fb00030766
chr2R	10124922	10125422	11.12	11.67	0.55	5.10E-05	0.000977	promoter-TSS #Bp00050011)	promoter-TSS (NM_165750)	313	Fb00025011 Fb00082959
chr3R	24161787	24162287	11.5	12.05	0.55	7.17E-05	0.00129	promoter-TSS #Bp00039157)	promoter-TSS (NM_001032017)	-313	Fb00039157 Fb00084551
chr3L	20226638	20227138	11.32	11.87	0.55	0.000135	0.00214	exon (Bp00060937, exon 2 of 4)	exon (NM_168235, exon 2 of 4)	508	Fb00026039 Fb00074825
chr3R	27929793	27930293	11.82	12.38	0.55	0.000178	0.0027	exon (Bp00027573, exon 3 of 8)	exon (NM_170359, exon 3 of 8)	419	Fb00027573 Fb00080254
chr3R	12691924	12692424	11.94	12.49	0.55	0.00021	0.00309	promoter-TSS #Bp00050955)	promoter-TSS (NM_080502)	313	Fb00035114 Fb00082664
chr3R	25680533	25681033	11.95	12.5	0.55	0.000244	0.00347	exon (Bp00036242, exon 2 of 5)	exon (NM_170208, exon 2 of 5)	484	Fb00036242 Fb00030931
chr2R	23063627	23064127	11.71	12.26	0.55	0.000784	0.00817	promoter-TSS #Bp00061361)	promoter-TSS (NM_170371)	283	Fb00030474 Fb00071984
chr2R	14852832	14853332	11.32	11.87	0.55	0.00101	0.00984	promoter-TSS #Bp00027783)	promoter-TSS (NM_137151)	142	Fb00028404 Fb00057448
chr3R	31805711	31806211	11.23	11.78	0.55	0.00136	0.0124	promoter-TSS #Bp00026755)	promoter-TSS (NM_143642)	283	Fb00039851 Fb00030513
chr2L	17496710	17497210	10.37	10.9	0.55	0.0018	0.0153	promoter-TSS #Bp00267279)	promoter-TSS (NM_124331)	301	Fb00267279 Fb00034649
chr3L	5753764	5754264	11.66	12.2	0.55						

chx	19750996	19751496	11.08	11.69	0.57	0.008099	0.00879			5' UTR (NM_00127295, exon 1 of 8)	477
chx3L	16739878	16380832	10.57	11.13	0.57	0.001158	0.0138			intron (FBgn0036522, intron 2 of 4)	826
chx	2187087	2187587	10.7	11.27	0.57	0.00173	0.0148			promoter-TSS (FBgn0024984)	-482
chx2L	20084416	20084916	11.6	12.17	0.57	0.00184	0.0156			exon (NM_001273688, exon 1 of 2)	404
chx	8689514	8690014	11.58	12.15	0.57	0.00219	0.0161			promoter-TSS (FBgn0027864)	282
chx3R	30783554	30784054	10.24	10.8	0.57	0.00211	0.0173			promoter-TSS (NM_125303)	-159
chx2L	20680929	20681429	11.42	12	0.57	0.00301	0.0228			exon (NM_165338, exon 1 of 1)	447
chx2L	3037623	3038123	12.02	12.59	0.57	0.00313	0.0256			promoter-TSS (NM_001233053)	368
chx2L	8051377	8051877	10.4	10.97	0.57	0.00382	0.0271			5' UTR (NM_164679, exon 2 of 4)	789
chx3L	14762301	14763301	8.8	9.37	0.57	0.00464	0.031			intron (FBgn0036446, intron 1 of 3)	1024
chx3R	10887776	10888276	10	10.57	0.57	0.00478	0.0316			exon (NM_001274918, exon 2 of 4)	1255
chx3L	9053082	9053582	9.36	9.94	0.57	0.00606	0.0374			exon (FBgn0035392, exon 3 of 7)	674
chx2R	12917919	12918419	10.05	10.62	0.57	0.00707	0.0422			5' UTR (NM_001033885A, exon 2 of 4)	312
chx2L	7027637	7028137	11.47	12.04	0.57	0.00774	0.045			promoter-TSS (FBgn0051908)	362
chx2L	16039162	16039662	9.09	9.66	0.57	0.00791	0.0462			exon (FBgn0028644, exon 8 of 8)	-2634
chx2R	18850823	18851323	11.44	12.02	0.58	6.03e-07	2.69E-05			promoter-TSS (FBgn0034401)	230
chx3L	9968474	9968974	10.96	11.54	0.58	6.15e-07	2.72E-05			promoter-TSS (NM_140119)	246
chx2L	4883510	4884010	11.51	12.09	0.58	1.14E-06	4.56E-05			promoter-TSS (FBgn000296)	363
chx3L	12267595	12268095	11.4	11.98	0.58	1.26E-06	4.95E-05			exon (FBgn0036446, intron 1 of 3)	418
chx3L	2373956	2374456	11.53	12.11	0.58	1.46E-06	5.64E-05			promoter-TSS (NM_13005335)	479
chx2L	10229316	10229816	10.67	11.25	0.58	2.46E-06	8.61E-05			TTS (FBgn0050843)	610
chx3L	21515408	21515908	10.76	11.34	0.58	1.16E-05	0.000294			promoter-TSS (FBgn0037093)	229
chx2L	12027500	12028000	10.49	11.07	0.58	2.60E-05	0.000563			exon (FBgn0010497, exon 1 of 5)	313
chx	7109448	7109948	10.99	11.57	0.58	2.82E-05	0.000604			5' UTR (NM_167119, exon 2 of 4)	623
chx3L	19487111	19488111	11.47	12.05	0.58	2.91E-05	0.000614			exon (NM_168799, exon 2 of 2)	278
chx3R	10144697	10145197	10.59	11.17	0.58	9.98E-05	0.00168			5' UTR (NM_143815, exon 1 of 9)	550
chx2R	11897951	11900251	11.59	12.17	0.58	0.000118	0.00192			exon (NM_136068, exon 1 of 1)	405
chx2L	1981225	1981725	11.47	12.05	0.58	0.000301	0.00404			promoter-TSS (FBgn0031377)	388
chx	1874683	1874983	12.02	12.59	0.58	0.000421	0.00522			promoter-TSS (FBgn0022517)	481
chx2L	8527287	8527787	10.72	11.3	0.58	0.000457	0.00546			TTS (FBgn0030953)	205
chx3L	11215639	11216139	9.98	10.56	0.58	0.00102	0.00995			exon (FBgn0026553, exon 1 of 7)	111
chx3R	15351151	15351851	11.66	12.24	0.58	0.00117	0.0109			promoter-TSS (FBgn0026722)	341
chx2R	9414594	9415094	11.33	11.91	0.58	0.00141	0.0126			exon (NM_140827, exon 1 of 6)	503
chx3R	10863889	10863789	10.44	11.02	0.58	0.00286	0.0219			promoter-TSS (FBgn0037846)	-551
chx2R	17460302	17460802	9.52	10.1	0.58	0.00291	0.0221			promoter-TSS (FBgn002953)	-45
chx2R	6742865	6743365	9.43	10.01	0.58	0.00306	0.0231			exon (NM_136388, exon 4 of 8)	1217
chx3R	9019134	9019814	12.06	12.64	0.58	0.00404	0.0282			TTS (NM_109257)	458
chx2R	9478756	94788356	10.37	10.95	0.58	0.00476	0.0315			promoter-TSS (FBgn0030940)	624
chx2L	10894	10894	11.54	12.12	0.58	0.00508	0.0328			promoter-TSS (FBgn0002378)	228
chx	8895119	8895619	8.98	9.55	0.58	0.00517	0.0332			intron (FBgn0011661, intron 1 of 8)	1174
chx	21043014	21043514	9.15	9.73	0.58	0.0053	0.0338			TTS (FBgn0056613)	195
chx3L	1899800	19000300	11.38	11.97	0.58	5.73E-08	3.75E-08			TTS (FBgn0036039)	558
chx3R	669621	6696021	10.86	11.45	0.59	3.25E-07	1.64E-05			TTS (FBgn0262801)	229
chx3R	17040948	17041448	11.67	12.26	0.59	5.94E-07	1.64E-05			promoter-TSS (NM_142332)	517
chx	9272333	9272833	11.66	12.26	0.59	1.24E-06	4.90E-05			promoter-TSS (FBgn0030122)	377
chx2L	10270241	10270741	11.51	12.1	0.59	3.72E-06	0.00012			TTS (NM_138529)	441
chx3R	15794549	15795049	11.06	11.65	0.59	4.63E-06	0.000142			intron (FBgn0026527, intron 1 of 8)	1018
chx2L	13780891	13781391	11.77	12.36	0.59	4.79E-06	0.000146			promoter-TSS (FBgn0029540)	453
chx	3375116	3375616	10.98	11.57	0.59	7.57E-06	0.000208			exon (NM_00127275, exon 2 of 3)	2208
chx	16271305	16271805	10.75	11.33	0.59	7.65E-06	0.000209			exon (NM_132881, exon 2 of 3)	152
chx3R	29253830	29254330	10.41	11	0.59	2.51E-05	0.000549			intron (FBgn001261357, intron 1 of 1)	1245
chx2L	5235752	5236252	11.09	11.68	0.59	2.99E-05	0.000625			promoter-TSS (FBgn0026302)	-501
chx2R	12880531	12881031	11.54	12.13	0.59	3.15E-05	0.000651			promoter-TSS (NM_165959)	350
chx3R	13647828	13648328	11.06	11.65	0.59	4.92E-05	0.000954			promoter-TSS (NM_169531)	-743
chx2R	11198823	11199323	10.5	11.09	0.59	0.000104	0.00172			intron (FBgn0033396, intron 1 of 6)	798
chx3R	5798344	5798844	11.28	11.87	0.59	0.00015	0.00234			promoter-TSS (FBgn002690)	470
chx3R	24582142	24582642	10.62	11.2	0.59	0.000288	0.00391			TTS (FBgn0039210)	482
chx2R	9390269	9390769	11.56	12.15	0.59	0.000344	0.00424			exon (NM_001262656, exon 2 of 3)	2094
chx3R	17545766	17546266	11.36	11.95	0.59	0.00099	0.00973			TTS (FBgn0029571)	515
chx2L	2769227	2769727	12	12.59	0.59	0.00111	0.0106			promoter-TSS (NM_001234595)	589
chx2L	19132507	19133007	11.86	12.45	0.59	0.00147	0.0132			promoter-TSS (NM_001273632)	435
chx2L	16349154	16349654	8.89	9.48	0.59	0.00502	0.0326			exon (NM_001273570, exon 2 of 2)	879
chx	7969644	7970144	10.01	10.6	0.59	0.00595	0.0371			5' UTR (NM_001162226, exon 1 of 6)	132
chx2R	25268998	25269498	8.23	8.82	0.59	0.0065	0.0396			Intergenic	13930
chx3L	18627239	18628239	11.3	11.89	0.59	0.00877	0.0495			promoter-TSS (FBgn00313717)	238
chx3R	9632267	9632767	10.99	11.58	0.6	1.08E-07	6.24E-06			TTS (FBgn0033394)	559
chx2R	14877688	14878188	10.95	11.54	0.6	1.56E-07	8.77E-06			TTS (FBgn0039978)	153
chx3R	11235246	11235746	11.9	12.49	0.6	3.02E-07	1.83E-05			TTS (NM_130154, exon 2 of 3)	194
chx2R	15931355	15931855	11.16	11.75	0.6	4.90E-07	2.26E-05			TTS (NM_176182)	300
chx2R	10111204	10111704	10.83	11.43	0.6	1.08E-06	4.40E-05			exon (FBgn0003071, exon 1 of 8)	585
chx	4381718	4382218	10.73	11.33	0.6	1.29E-05	0.000318			TTS (NM_169434)	663
chx3L	10631372	10631872	11.51	12.11	0.6	1.71E-05	0.000402			promoter-TSS (FBgn00260643)	281
chx3L	16779039	16779539	11.34	11.94	0.6	1.71E-05	0.000402			promoter-TSS (NM_168063)	447
chx2R	14281223	14281723	11.52	12.12	0.6	0.000256	0.00359			exon (FBgn0026269, exon 2 of 8)	347
chx3R	4289979	4289979	10.83	11.43	0.6	0.000389	0.00494			TTS (NM_001169575)	485
chx2R	22940228	22940728	10.59	11.19	0.6	0.000885	0.00905			exon (FBgn0039404, exon 3 of 4)	318
chx3R	24253919	2425919	9.51	10.21	0.6	0.00122	0.0113			exon (FBgn0013225, exon 2 of 13)	469
chx3R	17534278	17534778	8.79	9.39	0.6	0.00138	0.0124			intron (FBgn0029914, intron 1 of 16)	278
chx2R	12570599	12571099	10.51	11.11	0.6	0.00144	0.0128			exon (FBgn0003527, exon 2 of 4)	483
chx2R	8892022	8892522	11.7	12.3	0.6	0.00156	0.0136			promoter-TSS (FBgn0023337)	441
chx3R	11742240	11742740	11.49	12.09	0.6	0.00162	0.014			promoter-TSS (FBgn0029906)	601
chx4	1304145	1304645	9.13	9.73	0.6	0.00221	0.0179			Intergenic	72513
chx2R	13821258	13821758	9.35	9.95	0.6	0.00229	0.0185			exon (FBgn0000662, exon 3 of 5)	3045
chx3R	5798828	5799328	11.88	12.48	0.6	0.00238	0.019			promoter-TSS (FBgn0037369)	331
chx	2122297	2122797	10.19	10.79	0.6	0.00254	0.02			promoter-TSS (FBgn0004860)	739
chx2R	11358051	11358551	9.91	10.5	0.6	0.00255	0.023			promoter-TSS (FBgn0011987)	-651
chx3R	2993380	2993880	11.02	11.62	0.6	0.00283	0.0249			exon (FBgn0002621, exon 2 of 2)	28
chx3R	31581019	31581519	11.34	11.95	0.6	4.59E-09	4.62E-07			promoter-TSS (FBgn0043900)	316
chx2L	778370	778370	11.04	11.66	0.6	4.65E-08	3.13E-06			TTS (NM_138303)	456
chx3L	6174483	6174983	11.26	11.87	0.6	4.86E-07	2.26E-05			promoter-TSS (FBgn0035088)	486
chx2R	18971980	18972480	11.49	12.11	0.6	5.98E-07	2.68E-05			promoter-TSS (FBgn00267351)	307
chx	3967952	3968452	11.7	12.32	0.61	7.84E-07	3.41E-05			promoter-TSS (NM_123785)	453
chx2L	5721090	5721590	11.56	12.17	0.61	1.17E-06	4.68E-05			promoter-TSS (FBgn0031732)	290
chx3L	22075267	22075767	11.25	11.86	0.61	5.61E-06	0.000162			promoter-TSS (NM_141122)	429
chx	8393990	8393990	10.48	11.08	0.61	1.35E-05	0.000328			promoter-TSS (FBgn0011762)	239
chx	634835	635335	11.87	12.49	0.61	1.75E-05	0.000409			promoter-TSS (FBgn0030575)	460
chx2L	18703675	18704175	11.16	11.77	0.61						

chr2L	10857543	10858043	11.41	12.03	0.62	0.00416	0.0288	paramoter-TSS (FgRn0272363)	paramoter-TSS (NM_058100)	301	FgRn024227 FgRn0808180
chr2R	23987938	23988438	10.53	11.15	0.62	0.00777	0.045	paramoter-TSS (FgRn034990)	paramoter-TSS (NM_138070)	267	FgRn028260 FgRn072201
chrX	59135500	11859580	12.1	12.71	0.63	0.00388	0.0182	paramoter-TSS (FgRn02782)	paramoter-TSS (NM_140046)	419	FgRn027251 FgRn073251
chr1L	11892890	11893390	10.57	11.19	0.62	0.00846	0.0482	paramoter-TSS (FgRn02782)	paramoter-TSS (NM_132538)	802	FgRn026780 FgRn072573
chr3L	11117650	11118150	10.17	10.79	0.62	0.00868	0.0491	exon (FgRn0296134, exon 2 of 9)	exon (NM_168446, exon 2 of 9)	566	FgRn036134 FgRn0300785
chr3R	19447247	19447747	11.6	12.23	0.63	4.09E-08	2.91E-06	exon (FgRn022395, exon 1 of 1)	exon (NM_140854, exon 1 of 1)	325	FgRn022395 FgRn047998
chr2R	12587015	12587515	10.31	10.94	0.63	4.69E-05	0.000916	intron (FgRn0222764, intron 1 of 11)	intron (NM_165916, intron 1 of 11)	1887	FgRn0222764 FgRn030565
chrX	8101130	10101630	10.62	11.25	0.63	0.000101	0.0017	paramoter-TSS (NM_00129997)	paramoter-TSS (NM_00129801)	74	FgRn020997 FgRn0340140
chr2L	275157	2752657	10.47	11.1	0.63	0.000151	0.00236	paramoter-TSS (FgRn031453)	paramoter-TSS (NM_001273031)	367	FgRn031453 FgRn0335125
chr3L	21492519	21493019	11.18	11.81	0.63	0.000219	0.00318	paramoter-TSS (NM_1689064)	paramoter-TSS (NM_1689064)	375	FgRn030781 FgRn078939
chrX	9552439	9552939	10.14	10.77	0.63	0.000522	0.00817	exon (FgRn029527, exon 2 of 2)	exon (NM_132339, exon 2 of 2)	998	FgRn029527 FgRn034141
chr2L	9365424	9365924	11	11.63	0.63	0.000527	0.00817	paramoter-TSS (FgRn035981)	paramoter-TSS (NM_140046)	277	FgRn035981 FgRn076499
chr2R	5987521	5988021	9.58	10.21	0.63	0.001131	0.0107	exon (FgRn020194, exon 2 of 2)	exon (NM_052746, exon 2 of 2)	694	FgRn020194 FgRn032599
chrX	15845445	15845945	9.87	10.5	0.63	0.0019	0.016	exon (FgRn030099, exon 3 of 7)	exon (NM_132954, exon 3 of 7)	1029	FgRn030099 FgRn0340116
chrX	6869281	6869781	11.91	12.54	0.63	0.00254	0.02	exon (FgRn0212863, exon 2 of 4)	exon (NM_10222863, exon 2 of 5)	1094	FgRn026993 FgRn0330262
chr2R	9185146	9185646	11.26	11.89	0.63	0.00267	0.0207	exon (FgRn025267, exon 2 of 5)	exon (NM_165666, exon 2 of 5)	676	FgRn025267 FgRn089932
chr2L	17385480	17385980	8.66	9.29	0.63	0.00291	0.0221	paramoter-TSS (NM_00123598)	paramoter-TSS (NM_00123598)	415	FgRn020507 FgRn033451
chrX	20426662	20427062	11.6	12.24	0.63	0.00306	0.0231	TIS (FgRn031115)	TIS (NM_167204)	229	FgRn040650 FgRn0340402
chr3L	19898509	19899009	8.27	8.9	0.63	0.00348	0.0254	intron (FgRn014037, intron 1 of 4)	intron (NM_140898, intron 1 of 4)	2968	FgRn025238 FgRn0340130
chr2R	12194975	12195475	11.25	11.88	0.63	0.00748	0.044	exon (FgRn0272504, exon 2 of 5)	exon (NM_101222984, exon 2 of 5)	681	FgRn0272504 FgRn033063
chrX	9197517	9198017	11.14	11.77	0.64	1.98E-09	2.31E-07	TIS (FgRn030100)	TIS (NM_167878)	447	FgRn030100 FgRn071939
chr3R	29558358	29558858	11.3	11.94	0.64	4.95E-09	4.85E-07	TIS (FgRn030102)	exon (NM_142481, exon 2 of 2)	328	FgRn030971 FgRn089927
chr2L	21164125	21164625	11.79	12.44	0.64	2.44E-06	8.61E-05	paramoter-TSS (FgRn0266215)	paramoter-TSS (NR_124409)	104	FgRn0266215 FgRn034876
chr3L	5812323	5812823	11.02	11.66	0.64	4.13E-06	0.00013	TIS (FgRn030984)	TIS (NM_01014505)	520	FgRn035640 FgRn071113
chr3R	9763910	9764190	11.36	12	0.64	5.08E-06	0.00015	exon (FgRn027746, exon 2 of 4)	exon (NM_011170105, exon 2 of 4)	407	FgRn027746 FgRn0301870
chr3L	10170655	10171155	10.65	11.29	0.64	1.32E-05	0.000323	exon (FgRn0263241, exon 1 of 4)	exon (NM_176317, exon 1 of 4)	342	FgRn0263241 FgRn070204
chr3R	18169579	18170079	10.06	10.69	0.64	0.000162	0.00251	exon (FgRn0242363, exon 2 of 9)	exon (NM_01144599, exon 2 of 9)	1052	FgRn0242363 FgRn0305034
chr3L	25948106	25948606	10.36	11	0.64	0.000278	0.00383	exon (FgRn0263236, exon 2 of 7)	exon (NM_170406, exon 2 of 7)	548	FgRn0263236 FgRn0854000
chrX	653705	654205	11.22	11.86	0.64	0.00391	0.00494	exon (FgRn0256339, exon 2 of 3)	exon (NM_130497, exon 2 of 3)	983	FgRn0256339 FgRn0308252
chr2R	26024589	26025089	10.82	11.46	0.64	0.000441	0.00541	5' UTR (NM_143857, exon 1 of 6)	5' UTR (NM_143857, exon 1 of 6)	260	FgRn025696 FgRn085319
chr3R	9627884	9628384	10.12	10.76	0.64	0.000593	0.00624	TIS (FgRn025558)	TIS (NM_0295628)	814	FgRn025558 FgRn082077
chr2L	18707513	18708013	10	10.63	0.64	0.000616	0.00699	exon (FgRn025608, exon 2 of 3)	exon (NM_088040, exon 2 of 3)	836	FgRn025608 FgRn0341151
chrX	2121256	2121756	11	11.65	0.64	0.000624	0.00706	exon (FgRn004540, exon 2 of 5)	exon (NM_080950, exon 2 of 5)	286	FgRn004540 FgRn070433
chrX	69976709	69981909	9.74	10.37	0.64	0.000838	0.00867	paramoter-TSS (FgRn0272108)	paramoter-TSS (NM_001169213)	218	FgRn0264385 FgRn0332282
chr2R	7548406	7548906	10.48	11.13	0.64	0.000954	0.00945	exon (FgRn025895, exon 1 of 5)	exon (NM_088957, exon 1 of 5)	600	FgRn025895 FgRn089929
chr2L	8401404	8401904	11.58	12.23	0.64	0.00104	0.0101	paramoter-TSS (FgRn0266326)	paramoter-TSS (NR_124146)	430	FgRn0266326 FgRn0344057
chr3R	9548429	9548929	9.07	9.71	0.64	0.00173	0.0148	paramoter-TSS (NM_001273998)	paramoter-TSS (NM_001273998)	444	FgRn033421 FgRn0310331
chr2L	19791079	19791579	8.58	9.22	0.64	0.00259	0.0199	intron (FgRn0320224, intron 1 of 6)	intron (NM_136158, intron 1 of 6)	1553	FgRn0320224 FgRn0332729
chr3R	26024589	26025089	10.82	11.46	0.64	0.000441	0.00541	exon (FgRn0256339, exon 2 of 3)	exon (NM_130497, exon 2 of 3)	983	FgRn025696 FgRn085319
chrX	3296546	3297046	7.79	8.43	0.64	0.00266	0.0207	intron (FgRn0260427, intron 5 of 15)	intron (NM_163669, intron 5 of 15)	1516	FgRn0260427 FgRn070518
chrX	7681014	7681514	8.09	8.73	0.64	0.00315	0.0237	paramoter-TSS (FgRn0267225)	paramoter-TSS (NR_123832)	14	FgRn0267225 FgRn034937
chr2L	11125211	11113021	11.2	11.84	0.64	0.00319	0.0239	TIS (FgRn036133)	TIS (NM_140182)	969	FgRn033462 FgRn076229
chr3L	14636666	14602166	10.92	11.56	0.64	0.00351	0.0256	paramoter-TSS (FgRn036406)	paramoter-TSS (NM_140433)	388	FgRn026736 FgRn034296
chr2R	6502135	65054165	9.51	10.15	0.64	0.0036	0.0259	paramoter-TSS (FgRn0267444)	paramoter-TSS (NR_133106)	380	FgRn026736 FgRn034296
chrX	5736261	5737661	9.1	9.74	0.64	0.00389	0.0275	intron (FgRn0052758, intron 1 of 5)	intron (NM_167045, intron 1 of 5)	1520	FgRn0052758 FgRn070843
chr2L	13387902	13388402	8.47	9.11	0.64	0.0042	0.029	paramoter-TSS (FgRn0366104)	paramoter-TSS (NR_133117)	817	FgRn0265104 FgRn0346709
chr3L	5901799	5902299	9.99	10.57	0.64	0.00432	0.0296	TIS (FgRn029413)	TIS (NM_001104051)	497	FgRn029413 FgRn077880
chr2R	220123	220173	8.31	8.93	0.64	0.00435	0.0296	exon (FgRn0320221)	exon (NM_001272858, exon 2 of 8)	543	FgRn0320221 FgRn0341151
chrX	12597193	12597693	9.66	10.31	0.64	0.00531	0.0338	exon (FgRn0103944, exon 2 of 13)	exon (NM_01103944, exon 2 of 13)	1134	FgRn0103944 FgRn0341151
chr3R	31794844	31795344	9.81	10.45	0.64	0.00756	0.0444	exon (FgRn0260230, exon 2 of 6)	exon (NM_080486, exon 2 of 6)	1114	FgRn0260230 FgRn085845
chr3L	18683354	18683854	11.65	12.29	0.65	2.13E-07	1.13E-05	TIS (FgRn027826)	TIS (NM_168774)	457	FgRn036805 FgRn0345383
chr2L	19114712	19115212	10.56	11.21	0.65	4.58E-07	2.19E-05	TIS (FgRn000055)	TIS (NM_165278)	562	FgRn020236 FgRn081155
chr3R	564858	5647358	10.41	11.06	0.65	3.82E-06	0.000122	exon (FgRn0272951, exon 1 of 11)	exon (NM_01144539, exon 1 of 11)	802	FgRn0272951 FgRn029206
chrX	699109	6990909	11.93	12.57	0.65	3.80E-05	0.000767	paramoter-TSS (FgRn027108, exon 1 of 1)	paramoter-TSS (NM_01169213, exon 1 of 1)	688	FgRn027108 FgRn0308680
chr2L	19779872	19779872	11.26	11.92	0.65	4.52E-05	0.000887	paramoter-TSS (NM_134291)	paramoter-TSS (NM_134291)	422	FgRn027097 FgRn081285
chr3L	1273669	1273469	9.8	10.46	0.65	0.000235	0.0039	exon (FgRn0261933, exon 2 of 2)	exon (NM_088035, exon 2 of 2)	692	FgRn0261933 FgRn079592
chrX	4155014	41551014	11.7	12.35	0.65	0.000245	0.0039	paramoter-TSS (FgRn024302)	paramoter-TSS (NM_166647)	441	FgRn024302 FgRn071201
chr2R	2401264	2401314	11.7	12.35	0.65	0.00049	0.00531	paramoter-TSS (FgRn034978)	paramoter-TSS (NM_166647)	540	FgRn024959 FgRn072162
chr2R	26462113	26462613	9.21	9.86	0.65	0.000917	0.00928	intron (FgRn039340, intron 1 of 6)	intron (NM_143224, intron 1 of 6)	832	FgRn039340 FgRn085015
chrX	8277733	8277733	10	10.65	0.65	0.00124	0.0114	5' UTR (NM_164805, exon 1 of 10)	5' UTR (NM_164805, exon 1 of 10)	219	FgRn030502 FgRn079618
chr2R	14639874	14620374	9.61	10.26	0.65	0.0014	0.0126	TIS (FgRn010638)	TIS (NM_143789)	1396	FgRn033948 FgRn087500
chr2L	13550610	13551160	10.71	11.36	0.65	0.00185	0.0156	paramoter-TSS (NM_176167)	paramoter-TSS (NM_176167)	590	FgRn035156 FgRn080683
chrX	15747780	15748280	8.69	9.34	0.65	0.00349	0.0255	intron (FgRn010382, intron 1 of 5)	intron (NM_165125, intron 1 of 5)	1726	FgRn0266309 FgRn0344038
chr3L	15046410	15046910	7.52	8.17	0.65	0.0037	0.0265	intron (FgRn030683, intron 1 of 4)	intron (NM_206715, intron 1 of 4)	2267	FgRn030683 FgRn03959
chr3R	9376778	9377278	11.71	12.36	0.65	0.00524	0.0395	TIS (FgRn036982)	TIS (NM_140074)	347	FgRn001225 FgRn076496
chr2R	2202258	2201278	10.2	10.85	0.65	0.00531	0.0361	exon (FgRn029442, exon 2 of 18)	exon (NM_01103944, exon 2 of 18)	117	FgRn029442 FgRn0341151
chr2L	16698867	16699367	9.48	10.13	0.65	0.00576	0.0361	TIS (FgRn026374)	TIS (NR_047974)	360	FgRn026374 FgRn034436
chr2L	3353016	3353516	9.71	10.35	0.65	0.00653	0.0397	exon (FgRn020445, exon 2 of 9)	exon (NM_205900, exon 2 of 9)	1563	FgRn020445 FgRn089979
chr3R	19158394	19158894	10.54	11.19	0.65	0.00698	0.0417	exon (FgRn020262, exon 3 of 4)	exon (NM_057310, exon 3 of 4)	1234	FgRn020262 FgRn083749
chr2L	223559	2236059	9.91	10.55	0.65	0.00729	0.0432	paramoter-TSS (FgRn026494)	paramoter-TSS (NR_073696)	537	FgRn026494 FgRn0335158
chr3L	9694021	9694521	10.95	11.61	0.65	3.94E-09	4.10E-07	paramoter-TSS (FgRn0266596)	paramoter-TSS (NR_124838)	186	FgRn036303 FgRn0345032
chr2R	5927784	5928284	10.84	11.51	0.66	1.12E-08	9.93E-07	paramoter-TSS (FgRn0303523)	paramoter-TSS (NM_00104568)	253	FgRn0303523 FgRn091477
chr3R	21388313	21388813	10.66	11.32	0.66	4.15E-08	2.92E-06	paramoter-TSS (FgRn036428)	paramoter-TSS (NM_029078)	758	FgRn036427 FgRn071608
chr2R	19725208	19725708	11.11	11.77	0.66	2.10E-07	1.12E-05	TIS (FgRn034946)	TIS (NM_166648)	328	FgRn002791 FgRn072756
chr3R	16598466	16598966	11.69	12.35	0.66	2.88E-06	9				

chr3R	24924666	24925106	18.39	11.07	0.68	0.00189	0.0146	promoter-TSS (HG00329251)	-359	Hg0030251 Hb00094718
chrX	6959093	6974003	10.22	10.9	0.68	0.00171	0.0147	exon (NM_132143, exon 1 of 2)	236	Hg0009294 Hb0071003
chr2L	1731574	1732074	10.7	11.38	0.68	0.00174	0.0148	promoter-TSS (HG0254494)	532	Hg0031361 Hb0033543
chr2L	20311605	20312105	11.62	12.3	0.68	0.00195	0.0163	exon (NM_00169540, exon 1 of 16)	637	Hg0003475 Hb00307894
chr2R	9439391	9440431	9.96	10.63	0.68	0.00316	0.0237	intron (HG0025952, intron 1 of 6)	705	Hg0025952 Hb0008550
chr2R	6758341	6758381	11.58	12.27	0.68	0.00394	0.0278	exon (NM_136391, exon 1 of 1)	312	Hg0031011 Hb00086134
chr3R	1989328	19894128	9.71	10.39	0.68	0.00442	0.0301	exon (HG0038744, exon 1 of 9)	334	Hg0038744 Hb0010117
chrX	5965319	5965319	9.94	10.02	0.68	0.00825	0.0474	promoter-TSS (NM_132075)	-309	Hg000780 Hb00040665
chr2R	14160653	14161153	11.16	11.85	0.69	2.425e-09	2.78e-07	promoter-TSS (HG0033901)	707	Hg0033901 Hb0033995
chr3R	8902133	8902633	10.5	11.19	0.69	2.446e-07	1.46e-05	exon (NM_136562, exon 2 of 2)	283	Hg0004921 Hb00088644
chr3R	5792117	5792167	11.31	11.97	0.69	7.599e-06	0.00208	TTS (HG0031789)	490	Hg0031787 Hb0073630
chr3L	20139894	20140394	9.94	10.64	0.69	0.002099	0.00403	exon (HG0036942, exon 1 of 1)	463	Hg0036942 Hb0074954
chrX	10803761	10804749	9.24	9.92	0.69	0.000578	0.00663	intron (HG0258256, intron 2 of 6)	201	Hg0258256 Hb0072314
chr2L	6337024	6337524	9.9	10.59	0.69	0.000646	0.00729	exon (HG0031799, intron 1 of 6)	1352	Hg0031799 Hb0070284
chr2R	24043915	24044415	9.69	10.38	0.69	0.000706	0.00772	TTS (HG0011226)	2640	Hg0011226 Hb00721298
chr3R	5658334	5658384	10.26	10.96	0.69	0.00075	0.00802	promoter-TSS (HG0264902)	-962	Hg0264902 Hb00334945
chr2R	4471777	4472277	9.4	10.09	0.69	0.000789	0.0083	promoter-TSS (HG0260798)	-897	Hg0260798 Hb00345363
chr2R	24398522	24399022	9.76	10.45	0.69	0.00121	0.0118	intron (HG0034977, intron 1 of 6)	703	Hg0034977 Hb0027266
chr3R	14696256	14697026	10.33	11.01	0.69	0.0015	0.0133	TTS (HG0262127)	106	NM_00300400
chrX	6268604	6269104	9.61	10.31	0.69	0.00299	0.0222	exon (HG0028282, exon 3 of 6)	538	NM_1340705, exon 3 of 6
chr2L	11999837	12000337	11.3	11.99	0.69	0.00317	0.0238	promoter-TSS (HG0042125)	1396	promoter-TSS (NM_144407)
chrX	3721602	3722102	8.31	9	0.69	0.00335	0.0248	intron (HG0027252, intron 2 of 7)	494	intron (NM_00122252, intron 2 of 7)
chrX	16003567	16004067	8.33	9.01	0.69	0.00358	0.0259	exon (HG0030969, exon 4 of 11)	1295	NM_206759, exon 4 of 11
chr2L	16688417	16688917	8.83	9.52	0.69	0.004	0.0281	intron (HG0261278, intron 1 of 6)	1078	intron (NM_057663, intron 1 of 6)
chr2L	183691	184191	10.68	11.38	0.69	0.00434	0.0297	exon (HG0016777, exon 2 of 11)	1322	intron (NM_00127275, exon 2 of 11)
chr3R	5672555	5673055	9.2	9.89	0.69	0.00472	0.0313	intron (HG0037863, intron 2 of 5)	3053	CpG 32625
chr1L	16730804	16731304	11.13	11.81	0.69	0.00502	0.0326	exon (HG0029156, exon 2 of 3)	428	exon (NM_052485, exon 2 of 3)
chr3L	9062073	9062573	10.9	11.58	0.69	0.00544	0.0345	exon (HG0003116, exon 3 of 4)	625	exon (NM_163811, exon 3 of 4)
chrX	8063754	8064254	10.22	10.91	0.69	0.00562	0.0354	exon (HG0004657, exon 2 of 7)	2360	exon (NM_080504, exon 2 of 7)
chr2L	12975987	12976487	9.64	10.33	0.69	0.00581	0.0363	exon (HG0039598, exon 3 of 5)	436	exon (NM_001273496, exon 3 of 5)
chr3R	9684458	9684958	10.12	10.81	0.69	0.00589	0.0367	promoter-TSS (HG0037730)	844	Hg0037730 Hb00333194
chrX	12046569	12047069	8.4	9.09	0.69	0.00598	0.0372	intron (HG0267001, intron 1 of 19)	1122	intron (NM_023650, intron 1 of 19)
chr2L	13810823	13811323	11.52	12.22	0.7	7.57e-09	7.11e-07	promoter-TSS (HG0028919)	400	promoter-TSS (NM_001273522)
chr2L	8070771	8071271	11.45	12.14	0.7	8.80e-09	8.13e-07	TTS (HG0013531)	603	TTS (NM_001273294)
chr3L	8112980	8112980	11.12	11.83	0.7	8.47e-06	0.00225	TTS (HG0026252)	408	TTS (NM_060491)
chr2L	25320166	25320666	10.92	11.63	0.7	1.55e-05	0.00037	promoter-TSS (HG0034915)	329	promoter-TSS (NM_00129754, exon 2 of 7)
chr3L	12203352	12203852	10.05	10.75	0.7	2.86e-05	0.000609	TTS (HG0036266)	333	TTS (NM_140302)
chrX	6090240	6090290	9.68	10.38	0.7	0.001044	0.00227	promoter-TSS (HG0022840)	469	promoter-TSS (NM_132066)
chrX	7891161	7891161	10.73	11.44	0.7	0.002043	0.00846	exon (HG0029970, exon 1 of 1)	236	exon (NM_132157, exon 1 of 1)
chr2R	24187504	24188004	9.77	10.47	0.7	0.004003	0.00505	promoter-TSS (HG0266083)	1.16	promoter-TSS (NM_001299902)
chr3L	8441149	8441499	9.13	9.83	0.7	0.000514	0.0061	exon (HG026352, exon 1 of 9)	308	5' UTR (NM_001144440, exon 1 of 9)
chr2R	16729556	16730056	8.33	9.03	0.7	0.000853	0.0088	intron (HG0050463, intron 1 of 11)	1578	intron (NM_166188, intron 1 of 11)
chr3R	10858943	10859443	11.41	12.11	0.7	0.00179	0.0148	TTS (HG0037843)	307	TTS (NM_00117034)
chr3R	10309955	10310455	9.99	9.78	0.7	0.00249	0.0094	exon (HG0037807, exon 2 of 4)	1408	exon (NM_0012071, exon 2 of 4)
chr2L	10430365	10430865	11.25	11.95	0.7	0.00708	0.0422	promoter-TSS (HG00264702)	-805	promoter-TSS (NM_0239420)
chr2L	58730	58730	8.27	8.97	0.7	0.0075	0.0441	exon (HG0267827)	467	promoter-TSS (NM_134466)
chr3R	18403476	18403976	11.29	12	0.71	1.04e-10	1.83e-08	exon (HG0048161, exon 3 of 4)	405	exon (NM_169314, exon 3 of 4)
chr3R	31225184	31225684	10.6	11.31	0.71	1.55e-07	8.75e-06	exon (HG0039313, exon 1 of 6)	535	5' UTR (NM_170545, exon 1 of 6)
chr2L	9570375	9570875	10.52	11.23	0.71	1.66e-06	6.30e-05	exon (HG0030130, exon 1 of 4)	378	5' UTR (NM_135454, exon 1 of 4)
chr2R	12170955	12171455	10.39	11.1	0.71	1.86e-06	6.96e-05	promoter-TSS (HG0014184)	80	promoter-TSS (NM_165864)
chr4	1259421	1259421	6.89	7.6	0.71	0.000332	0.00437	Intergenic	63780	HE T A LINE Jockey
chr3L	14124495	14124995	9.71	10.41	0.71	0.000349	0.00453	TTS (HG0038206)	803	TTS (NM_142081)
chr3R	2203115	2203215	11.63	12.33	0.71	0.000391	0.0049	promoter-TSS (HG0038972)	4318	promoter-TSS (NM_142748)
chr3L	15108490	15108990	9.32	10.03	0.71	0.000557	0.00737	promoter-TSS (HG00264702)	-805	promoter-TSS (NM_0239420)
chr2R	17028861	17029361	9.5	10.2	0.71	0.000677	0.0075	exon (HG0032112, exon 2 of 24)	1342	5' UTR (NM_206147, exon 2 of 24)
chr3R	19488122	19488622	9	9.71	0.71	0.000893	0.00909	exon (HG0034452, exon 7 of 9)	195	Hg0034451 Hb0089502
chr2L	8367028	8367528	10.25	10.95	0.71	0.00123	0.0114	intron (HG0260439, intron 2 of 6)	1291	CpG
chrX	2292232	2292732	10.17	11.62	0.71	0.00156	0.0137	exon (HG0023526, exon 2 of 3)	938	exon (NM_001272255, exon 2 of 3)
chr3R	25128025	25128525	9.92	10.63	0.71	0.00181	0.0153	promoter-TSS (HG0039280)	646	promoter-TSS (NM_001170249)
chr3R	6655042	6655542	11.86	12.57	0.71	0.00221	0.018	exon (HG0029360, exon 2 of 4)	914	exon (NM_167090, exon 2 of 4)
chr3L	20779717	20780217	9.52	10.23	0.71	0.00225	0.0182	TTS (HG0037023)	100	TTS (NM_140997)
chr2R	2476544	2476644	10.39	11.1	0.71	0.0024	0.025	exon (HG0036749, exon 2 of 4)	2024	intron (NM_00125453, intron 2 of 3)
chr3R	8342289	8342789	9.82	10.53	0.71	0.00457	0.0307	TTS (HG0025385)	927	TTS (NM_057339)
chr2L	3143508	3143808	8.59	9.29	0.71	0.00484	0.0318	intron (HG0015600, intron 1 of 10)	1228	intron (NM_057805, intron 1 of 10)
chr2R	17004391	17004891	10.85	11.55	0.71	0.00491	0.0321	promoter-TSS (HG00394179)	508	promoter-TSS (NM_137334)
chrX	657006	657006	10.67	11.37	0.71	0.00747	0.044	exon (HG0025637, exon 2 of 2)	183	exon (NM_001038729, exon 1 of 1)
chr3R	12711596	12712096	11.8	12.52	0.72	4.16e-09	4.27e-07	exon (HG0029005, exon 1 of 7)	425	exon (NM_029560, exon 1 of 7)
chr2R	18845666	18846166	11.2	11.92	0.72	7.34e-09	7.05e-07	TTS (HG0034401)	703	TTS (NM_137353)
chr2L	6724665	6725165	10.7	11.42	0.72	1.84e-08	1.49e-06	exon (HG0040174, exon 2 of 15)	1013	exon (NM_001273225, exon 2 of 15)
chrX	15851914	15852414	10.94	11.66	0.72	5.91e-08	3.82e-06	exon (HG0026707, exon 2 of 5)	1001	exon (NM_001169292, exon 2 of 5)
chr2R	6630139	6630639	10.36	11.07	0.72	2.65e-05	9.17e-03	exon (HG0026749, exon 2 of 10)	2008	intron (NM_001258674, intron 2 of 6)
chrX	17126147	17126647	10.04	10.76	0.72	3.47e-05	0.00071	exon (HG0004350, exon 2 of 10)	276	exon (NM_132967, exon 4 of 11)
chrX	17626894	17627394	11.31	12.02	0.72	5.02e-05	0.000967	exon (HG0030850, exon 2 of 4)	751	exon (NM_00144745, exon 2 of 3)
chr3R	26882113	26882613	9.77	10.49	0.72	7.77e-05	0.00136	intron (HG0004359, intron 1 of 2)	1804	intron (NM_001276070, intron 1 of 2)
chr3R	8671563	8672063	10.42	11.15	0.72	9.70e-05	0.00165	exon (HG0037607, exon 2 of 3)	-1329	exon (NM_169232, exon 2 of 3)
chr2R	17633005	17633505	10.26	10.99	0.72	0.000101	0.0017	TTS (HG0034255)	261	Hg0004400 Hb00082897
chr3L	17655072	17655572	9.36	10.07	0.72	0.000169	0.00259	exon (HG0036745, exon 3 of 3)	706	exon (NM_140743, exon 3 of 3)
chr3L	9130214	9130714	8.4	9.13	0.72	0.000347	0.00451	promoter-TSS (HG0011206)	-109	promoter-TSS (NM_029265)
chr3L	13331041	13331541	10.77	11.49	0.72	0.000422	0.00522	exon (HG0010309, exon 2 of 2)	273	exon (NM_167395, exon 2 of 2)
chr2R	7825467	7825967	9.15	9.87	0.72	0.000469	0.00554	promoter-TSS (HG0016654)	294	promoter-TSS (NM_165364)
chr3R	27254347	27254847	9.85	10.56	0.72	0.00075	0.00802	exon (HG0011289, exon 2 of 7)	1368	exon (NM_057648, exon 2 of 7)
chr3L	20544845	20545345	9.54	10.27	0.72	0.000759	0.00809	TTS (HG0038120)	1003	TTS (NM_033163)
chr3R	5661447	5661947	9.55	10.27	0.72	0.000795	0.00834	exon (HG0046222, exon 3 of 9)	1663	exon (NM_169092, exon 3 of 9)
chrX	19618840	19619340	10.08	10.78	0.72	0.000865	0.0089	TTS (HG001040)	705	TTS (NM_134484)
chr2R	11843282	11843782	10.17	10.88	0.72	0.000881	0.00903	promoter-TSS (HG0262819)	878	promoter-TSS (NM_001253941)
chrX	1045515	1046015	10.38	11.11	0.72	9.00e-04	0.00916	promoter-TSS (HG0026569)	-458	promoter-TSS (NM_001042791)
chr3R	18799338	18799838	9.25	9.96	0.72	0.000907	0.00922	promoter-TSS (HG0038651)	-715	promoter-TSS (NM_142495)
chr2R	14768991	14769491	8.73	9.45	0.72					

chr1R	15232053	15129553	11.05	11.8	0.75	8.66E-06	0.00227	exon #Rgn000533, exon 2 of 3	exon (NM_079638, exon 2 of 3)	5/5	IRgnc000533 (Fb01008308)
chr2L	12038828	12039328	9.99	10.75	0.75	9.75E-06	0.00253	promoter-TSS (NM_164896)	promoter-TSS (NM_164896)	-567	IRgnc002397 (Fb01008309)
chr2L	12098266	12098766	10.25	11	0.75	2.24E-05	5.00E-04	TTS (IRgnc002870)	TTS (NM_001042890)	-709	IRgnc003409 (Fb01008310)
chr3R	153191762	15319262	10.16	10.91	0.75	3.02E-05	0.00361	promoter-TSS (NM_1436318)	promoter-TSS (NM_1436318)	-790	IRgnc003856 (Fb01008378)
chr1R	159164293	159164793	9.38	10.13	0.75	0.00014	0.00222	exon (IRgnc026146), exon 2 of 6	exon (NM_001144749, exon 2 of 6)	904	IRgnc026146 (Fb01029622)
chr3R	17400240	17400740	10.14	10.89	0.75	0.000217	0.00317	intron (IRgnc026504, intron 1 of 2)	ICCTC[CCNA]Sample_repeat[Sample_repeat]	806	IRgnc026504 (Fb01008344)
chr2L	17440395	17440895	9.37	9.31	0.75	0.000112	0.00417	promoter-TSS (NM_001269362)	promoter-TSS (NM_001269362)	-425	IRgnc026432 (Fb01008348)
chr1R	11983923	11984423	10.38	11.13	0.75	0.000717	0.00761	exon (IRgnc026476, exon 3 of 3)	exon (NM_001147179, exon 3 of 3)	1715	IRgnc026476 (Fb01008354)
chr1R	9281010	9281510	9.28	10.03	0.75	0.00447	0.0393	exon (IRgnc026236, exon 3 of 9)	exon (NM_144849, exon 3 of 9)	1857	IRgnc026236 (Fb01008368)
chr1R	2065330	2065830	11.31	12.07	0.76	6.65E-08	4.12E-06	promoter-TSS (IRgnc025626)	promoter-TSS (NM_138017)	126	IRgnc025631 (Fb01034010)
chr1R	621761	622261	10.97	11.73	0.76	1.07E-07	6.24E-06	promoter-TSS (NM_130493)	promoter-TSS (NM_130493)	206	IRgnc016038 (Fb01008972)
chr1R	8062184	8062684	9.95	10.72	0.76	1.05E-05	0.00271	intron (IRgnc004657, intron 1 of 6)	promoter-TSS (NM_206674)	750	IRgnc004657 (Fb01021105)
chr3R	20587192	20587692	10.49	11.25	0.76	1.44E-05	0.00035	promoter-TSS (IRgnc027711)	promoter-TSS (NM_133425)	260	IRgnc027711 (Fb01034751)
chr3R	22133745	22134245	10.03	10.79	0.76	1.45E-05	0.00035	TTS (IRgnc025936)	TTS (NM_206540)	161	IRgnc025162 (Fb01008419)
chr3R	15289997	15290497	9.62	10.38	0.76	0.000114	0.00186	promoter-TSS (NM_079636)	promoter-TSS (NM_079636)	502	IRgnc026675 (Fb01034528)
chr1R	3447601	3448101	8.07	8.83	0.76	0.000211	0.00339	intron (IRgnc029657, intron 2 of 6)	intron (NM_001272776, intron 2 of 6)	1032	IRgnc029657 (Fb01034078)
chr2L	2584076	2584576	9.64	10.4	0.76	0.000258	0.00361	exon (IRgnc026479, exon 2 of 8)	exon (NM_001273017, exon 2 of 8)	1104	IRgnc026479 (Fb01033513)
chr2R	25262610	25263110	8.09	8.85	0.76	0.000267	0.00371	Intergenic	ICLTA[LINE]Jockey	-7542	IRgnc026504 (Fb01033172)
chr2R	12514615	12515115	10.07	10.83	0.76	0.000373	0.00479	promoter-TSS (IRgnc033450)	promoter-TSS (NM_136949)	192	IRgnc033470 (Fb01008294)
chr2R	17591098	17591598	11.43	12.19	0.76	0.000407	0.00509	promoter-TSS (IRgnc034749)	promoter-TSS (NM_001274131)	357	IRgnc029070 (Fb01008832)
chr1R	2146596	2147096	10.27	11.04	0.76	0.000891	0.00509	intron (NM_0004654, intron 2 of 2)	intron (NM_025712, intron 2 of 2)	1483	IRgnc004654 (Fb01033410)
chr2R	12124652	12124752	9.49	10.25	0.76	0.000917	0.00628	exon (IRgnc027504, exon 5 of 5)	exon (NM_001273984, exon 5 of 5)	1894	IRgnc033686 (Fb01008298)
chr1R	16799231	16799731	9.95	10.7	0.76	0.00122	0.0113	exon (IRgnc026791, exon 1 of 3)	exon (NM_206768, exon 1 of 3)	258	IRgnc026791 (Fb01007455)
chr1R	9354806	9355306	9.09	9.86	0.76	0.00125	0.0115	5' UTR (NM_00103452, exon 2 of 10)	5' UTR (NM_00103452, exon 2 of 10)	1121	IRgnc026791 (Fb01007455)
chr2R	17395514	17396014	8.8	9.56	0.76	0.00162	0.0141	intron (IRgnc010620, intron 1 of 3)	intron (NM_079973, intron 1 of 3)	1172	IRgnc010620 (Fb01008909)
chr3L	11159625	11159725	9.86	10.61	0.76	0.000868	0.00668	exon (IRgnc034174, exon 1 of 7)	exon (NM_001259125, intron 1 of 7)	620	IRgnc034174 (Fb01007780)
chr3L	8357618	8358118	9.22	9.99	0.76	0.00086	0.00487	exon (IRgnc001248, exon 2 of 2)	exon (NM_001144478, exon 2 of 2)	257	IRgnc001248 (Fb01007647)
chr3L	21173387	21173887	11.51	12.28	0.77	2.09E-11	4.68E-09	exon (IRgnc027026, exon 1 of 1)	exon (NM_206410, exon 1 of 1)	506	IRgnc027026 (Fb01008946)
chr2R	23994429	23994929	11.06	11.84	0.77	1.11E-09	1.47E-07	TTS (IRgnc02519)	TTS (NM_001014545)	506	IRgnc034951 (Fb01007220)
chr3R	16051946	16052446	10.21	10.99	0.77	6.28E-07	2.76E-05	exon (IRgnc027848, exon 2 of 15)	exon (NM_001170156, exon 2 of 15)	393	IRgnc027848 (Fb01034083)
chr2R	13952647	13953147	9.99	10.76	0.77	1.03E-05	0.00267	intron (NM_001259325, intron 3 of 7)	intron (NM_001259325, intron 3 of 7)	1128	IRgnc026730 (Fb010087614)
chr3R	10265814	10266314	9.35	10.13	0.77	0.000113	0.00185	intron (IRgnc026217, intron 1 of 11)	intron (NM_001260090, intron 1 of 11)	1129	IRgnc026217 (Fb010305354)
chr3R	18747098	18747598	9.6	10.37	0.77	0.000176	0.00268	exon (NM_0260999, exon 2 of 6)	exon (NM_080121, exon 2 of 6)	929	IRgnc026099 (Fb010083674)
chr1R	8419513	8420013	9.13	9.9	0.77	0.000192	0.00288	intron (IRgnc026495, intron 3 of 9)	intron (NM_206654, intron 3 of 9)	1018	IRgnc026495 (Fb010083105)
chr1L	8161804	8162304	10.23	11	0.77	0.000414	0.00518	promoter-TSS (IRgnc016143)	promoter-TSS (NM_132542)	364	IRgnc016143 (Fb010076404)
chr3L	18626550	18627050	9.9	10.68	0.77	0.00077	0.00816	intron (IRgnc036801, intron 1 of 8)	intron (NM_001275088, intron 1 of 8)	442	IRgnc036801 (Fb010075049)
chr3R	15644341	15644841	9.61	10.38	0.77	0.000943	0.0094	intron (NM_0003721, intron 1 of 7)	intron (NM_001260191, intron 1 of 7)	395	IRgnc003721 (Fb010079546)
chr1R	8142743	8143243	11.9	12.67	0.77	0.00159	0.0139	promoter-TSS (IRgnc025800)	promoter-TSS (NM_078524)	614	IRgnc030006 (Fb010071191)
chr3L	1964852	1965352	9.54	10.31	0.77	0.00362	0.026	promoter-TSS (NM_139424)	promoter-TSS (NM_139424)	632	IRgnc035298 (Fb010072844)
chr2R	2145626	2146126	8.71	9.48	0.77	0.00661	0.04	promoter-TSS (IRgnc0110470)	promoter-TSS (NM_001299751)	-352	IRgnc011047 (Fb010342897)
chr2L	12207419	12207919	8.2	8.97	0.77	0.00714	0.0424	promoter-TSS (NM_00100114)	promoter-TSS (NM_00169492)	-532	IRgnc000114 (Fb010080667)
chr2L	6069884	6070384	11.17	11.95	0.78	9.44E-13	2.45E-10	promoter-TSS (NM_078766)	promoter-TSS (NM_078766)	532	IRgnc017771 (Fb010082921)
chr3R	11821495	11821995	11.25	12.03	0.78	1.16E-09	1.48E-07	TTS (IRgnc019756)	TTS (NM_00142657)	320	IRgnc019756 (Fb010085713)
chr3L	2596395	2596895	10.05	10.83	0.78	8.47E-05	0.00335	TTS (IRgnc025355)	exon (NM_139473, exon 2 of 8)	407	IRgnc035285 (Fb010072528)
chr3R	29200641	29201141	11.54	12.32	0.78	2.45E-06	8.61E-05	TTS (IRgnc039620)	TTS (NM_143425)	115	IRgnc011422 (Fb010085390)
chr1R	6358086	6358586	9.94	10.72	0.78	5.43E-05	0.00103	intron (IRgnc026155, intron 1 of 2)	intron (NM_132051, intron 1 of 2)	485	IRgnc026155 (Fb010079907)
chr2L	6371603	6372103	9.32	10.1	0.78	7.66E-05	0.00135	promoter-TSS (IRgnc043854)	promoter-TSS (NM_135176)	546	IRgnc043854 (Fb01029281)
chr3R	25070142	25070642	10.71	11.38	0.78	8.55E-05	0.00147	promoter-TSS (NM_001170247)	promoter-TSS (NM_001170247)	434	IRgnc030721 (Fb010304718)
chr3L	22724569	22725069	9.33	10.12	0.78	0.000185	0.00279	exon (IRgnc025451, exon 2 of 5)	5' UTR (NM_001100205, exon 2 of 5)	783	IRgnc025451 (Fb010302314)
chr1R	9683125	9683625	11.58	12.37	0.78	0.000292	0.00396	exon (IRgnc0161617, exon 2 of 14)	5' UTR (NM_001201647, exon 2 of 14)	292	IRgnc0161617 (Fb010071402)
chr3L	11293716	11294216	10.82	11.6	0.78	0.000304	0.00408	exon (IRgnc036152, exon 1 of 3)	exon (NM_143020, exon 1 of 3)	820	IRgnc036152 (Fb010076183)
chr3R	15804917	15805417	10.19	10.97	0.78	0.000796	0.00834	TTS (IRgnc038665)	TTS (NM_142736)	528	IRgnc026616 (Fb010083119)
chr3R	13013732	13014232	9.57	10.35	0.78	0.000918	0.00928	intron (IRgnc025496, intron 1 of 5)	intron (NM_136491, intron 1 of 5)	234	IRgnc025496 (Fb010082699)
chr3L	7156497	7156997	8.24	9.03	0.78	0.00367	0.0263	exon (IRgnc025917, exon 1 of 8)	5' UTR (NM_163193, exon 1 of 8)	376	IRgnc025917 (Fb01029641)
chr1R	15856855	15857355	11.14	11.9	0.79	4.51E-11	8.66E-09	exon (IRgnc020734, exon 2 of 8)	exon (NM_126734, exon 2 of 8)	328	IRgnc020734 (Fb010074171)
chr3R	2363070	2363570	10.73	11.52	0.79	2.46E-09	2.79E-07	exon (IRgnc038673, exon 1 of 1)	exon (NM_142884, exon 1 of 1)	248	IRgnc038673 (Fb010339619)
chr1R	16357850	16358350	11.08	11.88	0.79	1.10E-08	9.80E-07	exon (IRgnc005411, exon 2 of 5)	exon (NM_057543, exon 2 of 5)	355	IRgnc005411 (Fb010307176)
chr2R	23895284	23895784	10.77	11.56	0.79	4.34E-08	3.00E-06	promoter-TSS (IRgnc034923)	promoter-TSS (NM_166629)	422	IRgnc034924 (Fb01032133)
chr1R	18083713	18084213	10.38	11.17	0.79	4.41E-08	3.00E-06	exon (IRgnc038890, intron 1 of 6)	CpG	1444	IRgnc038890 (Fb010343117)
chr3R	25244418	25244918	10.19	10.99	0.79	2.34E-06	8.41E-05	exon (NM_00000158, exon 1 of 3)	exon (NM_057452, exon 1 of 3)	389	IRgnc000158 (Fb010084869)
chr1R	20166773	20167273	11.46	12.25	0.79	6.22E-06	0.00078	exon (IRgnc031952, exon 2 of 4)	exon (NM_139653, exon 2 of 4)	351	IRgnc031952 (Fb010070007)
chr2R	10416236	10416736	9.75	10.54	0.79	0.00168	0.0128	exon (IRgnc016968, exon 1 of 7)	exon (NM_001273935, exon 1 of 7)	630	IRgnc016968 (Fb010083083)
chr1R	12655108	12655608	9.61	10.4	0.79	1.54E-05	0.00368	intron (IRgnc015024, intron 2 of 3)	promoter-TSS (NM_165145)	-373	IRgnc015024 (Fb010072682)
chr2L	16048928	16049428	9.59	10.38	0.79	2.22E-05	0.00495	promoter-TSS (NM_165145)	exon (NM_137428, exon 1 of 1)	896	IRgnc000182 (Fb010082085)
chr2R	17847294	17847794	11.28	12.07	0.79	3.75E-05	0.00076	exon (IRgnc015522, exon 1 of 1)	exon (NM_137428, exon 1 of 1)	896	IRgnc015522 (Fb010086799)
chr1R	20568222	20568722	11.01	11.79	0.79	5.01E-05	0.00097	promoter-TSS (NM_123969)	promoter-TSS (NM_123969)	754	IRgnc026702 (Fb010345968)
chr3R	5255582	5256082	9.18	9.97	0.79	6.19E-05	0.00114	TTS (IRgnc037804)	TTS (NM_141265)	876	IRgnc037805 (Fb010078821)
chr2L	1156937	1157437	10.8	11.6	0.79	7.49E-05	0.00133	promoter-TSS (IRgnc0261458)	promoter-TSS (NM_001038781)	849	IRgnc0261458 (Fb010100023)
chr1R	5880787	5881287	10.02	10.81	0.79	7.94E-05	0.00139	exon (IRgnc029820, exon 2 of 6)	exon (NM_137074, exon 2 of 6)	325	IRgnc029820 (Fb010345313)
chr2L	15957480	15957980	9.3	10.09	0.79	8.17E-05	0.00142	intron (IRgnc029672, intron 1 of 2)	intron (NM_134291, intron 1 of 2)	1297	IRgnc029672 (Fb010083697)
chr3R	10416236	10416736	9.75	10.54	0.79	0.00168	0.0128	exon (IRgnc027674, exon 2 of 10)	exon (NM_168519, exon 2 of 10)	630	IRgnc016968 (Fb010083083)
chr3L	222232	222282	9.39	10.18	0.79	0.000691	0.00413	intron (IRgnc029737, intron 1 of 3)	CpG	1589	IRgnc029737 (Fb010334706)
chr2R	14711673	14712173	9.44	10.23	0.79	0.00074	0.00797	exon (NM_138171, exon 2 of 9)	exon (NM_138171, exon 2 of 9)	1592	IRgnc035113 (Fb010072574)
chr2R	19670663	19671163	11.76	12.56	0.79	0.00115					

chr3L	15162476	15162976	10.06	10.91	0.84	2.42E-07	1.25E-05	T1S (NM_140550)	902	IRp0036489	IRb0075646
chr2R	25551130	25551630	10.16	11.44	0.84	1.24E-05	0.003099	exon (NM_001104447, exon 2 of 3)	1138	IRp0039766	IRb0084768
chrX	18566494	18566994	10.08	10.92	0.84	2.25E-05	0.005051	exon (NM_001227278, exon 3 of 5)	439	IRp0265958	IRb0340365
chr2R	18407821	18408321	9.83	10.67	0.84	3.41E-05	0.007011	promoter-TSS (NM_366279)	238	IRp0036492	IRb0086676
chr2R	11614930	11615430	9.76	10.59	0.84	6.25E-05	0.00115	intron (IRp0023366, intron 1 of 16)	865	IRp0033836	IRb0088991
chr3L	16039381	16039881	8.91	9.75	0.84	0.000102	0.00171	intron (NM_001274928, intron 1 of 2)	1661	IRp0266265	IRb0075501
chrX	8151441	8151491	10.25	11.1	0.84	0.00304	0.023	exon (IRp0030011, exon 1 of 3)	384	IRp0030011	IRb0071189
chr2L	6481870	6482370	10.79	11.64	0.85	4.81E-11	9.04E-09	exon (IRp0031810, exon 2 of 4)	1390	IRp0267109	IRb0345975
chr2R	16692226	16692726	10.66	11.51	0.85	1.48E-07	1.24E-05	exon (IRp0031961, exon 2 of 3)	776	IRp013631	IRb0087111
chrX	5906535	5906585	10.38	11.23	0.85	9.18E-07	3.88E-05	promoter-TSS (NM_0258221)	4	IRp0025821	IRb0070284
chr3L	24535054	24535554	8.63	9.48	0.85	0.000028	0.00208	exon (IRp0040045, exon 2 of 2)	1215	IRp0040045	IRb0307117
chr3L	11212452	11212952	10.34	10.99	0.85	0.000227	0.0033	promoter-TSS (NM_368443)	856	IRp0036439	IRb0076211
chr2R	18397116	18397616	8.55	9.4	0.85	0.000297	0.00401	exon (IRp0063499, exon 1 of 4)	160	IRp0063499	IRb0086776
chr3R	29766889	29767389	9.02	9.87	0.85	0.000428	0.00528	exon (IRp0001297, exon 1 of 3)	1145	IRp0001297	IRb0099988
chrX	19781974	19782474	11.4	12.25	0.85	0.000458	0.00557	exon (IRp0259789, exon 2 of 2)	919	IRp0259789	IRb0345614
chr2L	3630814	3631314	9.41	10.26	0.85	0.005002	0.06601	intron (IRp0000771, intron 1 of 7)	1090	IRp0000771	IRb0089306
chr2R	14770770	14771270	11.98	12.79	0.85	0.000328	0.00421	exon (IRp0009739, exon 4 of 11)	1262	IRp0009739	IRb0304366
chr2R	18128533	18129033	9.86	10.7	0.85	0.000981	0.00967	T1S (NM_180224)	790	IRp0040665	IRb0086739
chr2R	9293515	9294015	9.98	10.85	0.86	1.63E-07	9.06E-06	intron (IRp0002840, intron 1 of 2)	157	IRp0002840	IRb0333998
chrX	49231	49231	11.46	12.32	0.86	1.31E-05	0.00322	exon (NM_166844, exon 2 of 2)	542	IRp0022531	IRb0070991
chr3R	15198453	15198953	9.39	10.25	0.86	1.80E-05	0.000417	intron (NM_00127045, intron 1 of 5)	1762	IRp0260659	IRb0301102
chrX	16067455	16067955	7.06	7.92	0.86	2.71E-05	0.00582	exon (IRp0030269, exon 5 of 9)	3968	IRp0030269	IRb0074206
chrUn_D5484009.1	1823	2323	8.85	9.71	0.86	0.000531	0.00623	NA	NA	NA	NA
chr2R	13556794	13557294	10.22	11.08	0.86	0.000444	0.0128	T1S (IRp0263459)	642	IRp0038306	IRb0313500
chrX	14770770	14771270	11.98	12.79	0.86	0.000328	0.00421	promoter-TSS (NM_1668008)	631	IRp0026789	IRb0070735
chrX	27273631	27274131	10.35	11.02	0.87	1.60E-07	8.95E-06	promoter-TSS (NM_343295)	892	IRp0024411	IRb0085175
chrX	8713710	8714210	10.88	11.75	0.87	4.93E-06	0.000447	promoter-TSS (NM_300061)	213	IRp0030643	IRb0339260
chr3R	11195144	11195644	7.16	8.03	0.87	1.78E-05	0.000414	promoter-TSS (IRp0033787)	-697	IRp0033787	IRb0082341
chr2L	8410850	8411350	10.4	11.26	0.87	3.73E-05	0.000757	exon (IRp0014417, exon 2 of 2)	508	IRp0014417	IRb0331991
chr3R	12373064	12373564	9.32	10.2	0.87	5.90E-05	0.0011	promoter-TSS (NM_0010040)	247	IRp0010039	IRb0082570
chrX	7138036	7138536	9.08	9.94	0.87	0.000119	0.00192	intron (IRp0002939, intron 1 of 4)	564	IRp0029399	IRb0338215
chrX	21046635	21047135	9.4	10.28	0.87	0.00314	0.0237	exon (IRp0025814, exon 2 of 2)	139	IRp0265631	IRb0304739
chrX	2052927	2052977	9.52	10.4	0.87	0.00357	0.0257	exon (IRp0015789, exon 2 of 4)	199	IRp0015789	IRb0073353
chr3L	1477801	1478301	10.48	11.31	0.88	3.55E-12	8.50E-10	5' UTR (NM_138264, exon 1 of 1)	1799	IRp0028764	IRb0332625
chr3L	3162507	3163007	11.27	12.15	0.88	4.31E-07	2.08E-05	T1S (NM_167991)	578	IRp0204194	IRb0073029
chr2R	11853385	11853885	9.86	10.74	0.88	4.99E-07	2.28E-05	exon (IRp0031504, exon 2 of 5)	1019	IRp0013256	IRb0088033
chrX	13717259	13717759	10.26	11.14	0.88	7.24E-06	0.000201	exon (IRp0267576, exon 15 of 15)	776	IRp0030512	IRb0072844
chrX	468935	468935	8.75	9.63	0.88	1.09E-05	0.000279	exon (IRp0010019, exon 1 of 1)	150	IRp0010019	IRb0070099
chr3R	29040851	29041351	9.88	10.76	0.88	7.11E-05	0.00128	exon (IRp0003093, exon 2 of 11)	1272	IRp0003093	IRb0319150
chrX	15708559	15709059	10.55	11.43	0.88	0.000215	0.00314	exon (IRp0030672, exon 2 of 5)	295	IRp0036758	IRb0074057
chr3R	3180620	3180670	9.72	10.6	0.88	0.000899	0.0110	exon (IRp0033985, exon 2 of 2)	1916	IRp0033985	IRb0085252
chrX	22533834	22534334	8.94	9.82	0.88	0.00463	0.031	exon (IRp0018194, exon 2 of 3)	1389	IRp0018194	IRb0317385
chr2R	12138682	12139182	10.55	11.42	0.89	7.78E-13	2.05E-10	T1S (NM_136285)	146	IRp0033883	IRb0082882
chrX	16662322	16662322	10.33	11.22	0.89	1.07E-10	1.85E-08	exon (IRp0003189, exon 2 of 3)	392	IRp0003189	IRb0340155
chr3R	24231688	24232188	11.51	12.4	0.89	7.46E-10	1.03E-07	promoter-TSS (IRp00266744)	-574	IRp0266744	IRb0345246
chr3L	1473790	1474290	8.36	9.25	0.89	2.58E-05	0.000551	exon (IRp0267487, exon 3 of 3)	1217	IRp0267487	IRb0331611
chr2L	2749623	2750123	11.01	11.89	0.89	0.000209	0.00308	promoter-TSS (NM_001273030)	274	IRp0025066	IRb0330736
chr2R	25281169	25281669	10.03	10.93	0.89	0.000382	0.00487	Inteinspex	26101	IRp0035094	IRb0333172
chr3L	11884472	11885972	9.07	9.95	0.89	0.000203	0.0169	intron (IRp0036191, intron 1 of 5)	1743	IRp0036191	IRb0272362
chr3R	24627404	24627904	9.97	10.86	0.89	0.0024	0.0191	exon (IRp0046685, exon 3 of 3)	1804	IRp0046685	IRb0084636
chr2L	4183850	4184350	8.73	9.62	0.89	0.0034	0.025	Inteinspex	1300	IRp0264866	IRb0334786
chr3L	9715688	9716188	10.24	11.12	0.89	0.00343	0.0251	promoter-TSS (IR_002482)	-864	IRp0060001	IRb0091297
chr3L	7591537	7592037	8.84	9.73	0.89	0.00517	0.0332	promoter-TSS (IR_0018500)	-789	IRp0035800	IRb0076524
chr3R	29675872	29676372	10.67	11.57	0.89	1.14E-10	1.93E-08	promoter-TSS (IRp0025457)	472	IRp0035800	IRb0085452
chr2R	16582939	16583439	10.34	11.24	0.9	6.62E-09	6.42E-07	non coding (IR_1245A, exon 1 of 1)	180	IRp0265295	IRb0340074
chr3L	15140142	15140642	9.41	10.31	0.9	3.34E-06	0.000109	intron (NM_001300142, intron 1 of 1)	-1738	IRp0265295	IRb0345844
chr3L	862998	862998	9.6	10.5	0.9	0.00221	0.000221	promoter-TSS (IRp0035152)	-728	IRp0035152	IRb0333738
chr3R	9764773	9765273	9.26	10.16	0.9	1.28E-05	0.00018	exon (IRp0036043, intron 1 of 12)	1480	IRp0036043	IRb0307618
chr4	1202084	1202584	7.68	8.57	0.9	0.00356	0.027	HEA1LINE Jockey	70452	IRp0052853	IRb0010630
chr2R	18591830	18592330	9.85	10.76	0.91	1.18E-06	4.69E-05	T1S (NM_0038929)	1054	IRp0038929	IRb0292922
chr3L	6049920	6050420	10.27	11.18	0.91	1.51E-06	5.78E-05	promoter-TSS (NM_364679)	-629	IRp0031269	IRb0346115
chrX	18844328	18844828	9.93	10.84	0.91	2.94E-06	9.91E-05	exon (IRp0084541, exon 2 of 5)	916	IRp0084541	IRb0310440
chrX	7284612	7285112	9.51	10.42	0.91	2.98E-06	9.99E-05	promoter-TSS (NM_001299047)	15	IRp0029911	IRb0339819
chr2L	21308406	21308906	11.06	11.97	0.91	4.48E-06	0.000139	exon (IRp0016958, exon 2 of 12)	833	IRp0016958	IRb0301461
chrX	14869974	14874974	9.17	10.09	0.91	0.000176	0.00268	T1S (NM_206739)	500	IRp0030991	IRb0073948
chr2R	18613835	18614335	10.39	11.09	0.91	0.00356	0.0458	promoter-TSS (IRp0034362)	933	IRp0034362	IRb0086734
chr3R	1133481	1133981	11.14	12.06	0.91	1.04E-10	1.42E-08	intron (IRp0019831, intron 1 of 3)	132	IRp0019831	IRb0073142
chr3L	3811858	3812358	10.11	11.03	0.92	3.54E-09	3.72E-07	promoter-TSS (IRp0052263)	846	IRp0052263	IRb0073140
chr2L	5005351	5005851	10.89	11.85	0.92	1.15E-08	1.00E-06	promoter-TSS (NM_001169407)	-673	IRp0053113	IRb0079040
chr2L	15050502	15055502	9.27	10.19	0.92	2.50E-06	8.70E-05	exon (IRp0007418, exon 3 of 4)	2726	IRp0204183	IRb0080721
chrX	12908340	12908840	10.2	11.12	0.92	1.48E-05	0.000356	T1S (IRp0030433)	1032	IRp0030434	IRb0073639
chr3L	11821058	11821558	9.89	10.82	0.92	7.51E-05	0.00133	exon (IRp0030122, exon 2 of 2)	897	IRp0030122	IRb0302136
chr3R	23056511	23057111	9.19	10.12	0.92	0.000111	0.00184	promoter-TSS (NM_133432)	-14	IRp0267258	IRb0347322
chr3R	3866226	3866726	8.75	9.67	0.92	0.000155	0.00241	intron (IRp0265575, intron 3 of 5)	241	IRp0265575	IRb0340205
chr3R	13823389	13823889	7.93	8.81	0.92	0.000205	0.00265	intron (IRp0035726, intron 1 of 9)	2227	IRp0035726	IRb0309745
chrX	15885422	15885922	10.89	11.86	0.93	1.06E-09	1.42E-07	exon (IRp0269398, exon 2 of 14)	1639	IRp00127652	IRb0310065
chr2R	7987992	7988492	10.37	11.3	0.93	9.24E-08	5.63E-06	promoter-TSS (NM_365383)	-719	IRp0033747	IRb0088980
chr3L	15100260	15100760	8.81	9.74	0.93	4.87E-06	0.000147	exon (IRp0264703, exon 2 of 2)	-1008	IRp0264704	IRb0333938
chr2L	19436677	19441677	9.9	10.83	0.93	8.08E-06	0.000219	promoter-TSS (IRp0032783)	-951	IRp0032783	IRb0081209
chr3R	18411364	18411864	8.66	9.6	0.93	2.31E-05	0.00051	promoter-TSS (NM_368621)	17	IRp0038821	IRb0083654
chr3L	15556674	15557174	7.41	8.34	0.93	3.44E-05	0.000705	exon (IRp0087035, exon 1 of 6)	102	IRp0087035	IRb0313333
chr2R	10357576	10358076	11.23	12.16	0.93	3.85E-05	0.000775	exon (IRp0265512, exon 1 of 3)	758	IRp0265512	IRb0333946
chrX	14507	15007	7.54	8.47	0.93	0.000207	0.0027	NA	NA	NA	NA
chrX	7248734	7249234	9.59	10.56	0.93	0.000062	0.00056	intron (IRp0012957, intron 1 of 11)	685	IRp0029597	IRb0085924

chrX	3074850	3074850	8.13	9.13	1	1.77E-06	6.64E-05	intron (Rg0028369, intron 7 of 12)	intron (NM_080320, intron 7 of 12)	-7038	Rg0029649 Rg001070561
chr3R	9548380	9346330	8.85	9.85	1	1.87E-06	6.96E-05	intron (Rg0026410, intron 2 of 5)	intron (NM_001300317, intron 2 of 5)	6214	Rg0026410 Rg0010734368
chrX	15813803	15814043	9.53	10.54	1	1.04E-05	0.000806	5' UTR (NM_001103952, exon 3 of 12)	5' UTR (NM_001103952, exon 3 of 12)	4201	Rg00103345 Rg0010308330
chr2L	16824161	16824616	9.69	10.69	1.01	3.44E-06	0.000112	intron (Rg0022213, exon 2 of 2)	intron (NM_078864, exon 2 of 2)	2841	Rg0022213 Rg001080930
chr3L	21881613	21882113	7.51	8.52	1.01	0.000406	0.0283	intron (Rg0026777, intron 1 of 9)	intron (NM_001299949, intron 1 of 9)	2287	Rg0026777 Rg001031364
chr3R	30528876	30529376	10	11.02	1.02	2.65E-08	2.01E-05	intron (Rg0035109, exon 4 of 4)	intron (NM_370512, exon 4 of 4)	1197	Rg0035109 Rg001025645
chr3L	11490065	11490065	7.84	8.86	1.02	1.23E-05	8.86E-06	intron (Rg0036365, intron 1 of 3)	intron (NM_140213, intron 1 of 3)	2186	Rg0036365 Rg001076093
chr3L	3908183	3908683	9.42	10.44	1.02	0.000641	0.00725	TTS (Rg0041147)	TTS (NM_139585)	1159	Rg0025791 Rg001073176
chr2L	4977718	4978218	10.6	11.62	1.03	4.48E-19	3.60E-16	promoter-TSS (Rg00331661)	promoter-TSS (NM_135044)	657	Rg00331661 Rg001078997
chr3L	1330766	1331266	9.67	10.7	1.03	3.67E-07	1.82E-05	promoter-TSS (NM_1382509)	promoter-TSS (NM_1382509)	777	Rg0032909 Rg001072678
chrX	12762762	12763262	10.62	11.65	1.03	8.65E-06	0.000227	exon (Rg0030420, exon 1 of 1)	exon (NM_132599, exon 1 of 1)	549	Rg0030420 Rg0010346120
chr2L	16690845	16691345	9.09	10.12	1.03	1.72E-05	0.000404	intron (Rg00161278, intron 7 of 7)	intron (NM_165172, intron 7 of 7)	4928	Rg00161278 Rg00100459
chr3R	8347697	8348197	10.07	11.1	1.03	5.49E-05	0.00104	TTS (Rg0032580)	TTS (NM_1692318)	229	Rg0032581 Rg001083184
chr3L	2636877	2637377	9.79	10.83	1.04	1.80E-08	1.47E-06	intron (Rg00303575, exon 2 of 2)	5' UTR (NM_001144403, exon 3 of 8)	223	Rg00303575 Rg001029346
chrX	3968884	3969384	10.07	11.11	1.04	4.83E-08	3.22E-06	promoter-TSS (Rg00296829)	promoter-TSS (NM_130792)	202	Rg00296829 Rg0010346011
chr2R	18852138	18852638	10.6	11.64	1.04	1.51E-06	5.78E-05	TTS (Rg0024402)	TTS (NM_132529)	255	Rg0024403 Rg001072328
chr3L	1695695	1695745	8.2	9.24	1.04	4.42E-06	0.000138	promoter-TSS (NM_001104159)	promoter-TSS (NM_001104159)	-653	Rg0010182 Rg001075269
chr2L	10977804	10978304	9.94	10.98	1.04	3.35E-05	0.000689	exon (Rg00104781, exon 2 of 5)	exon (NM_001273440, exon 2 of 5)	1283	Rg00104781 Rg001080348
chr3L	19634087	19634587	9.39	10.44	1.05	2.86E-08	2.20E-06	exon (Rg00303696, exon 2 of 8)	exon (NM_176363, exon 2 of 8)	1241	Rg00303696 Rg0010302513
chr2L	2298372	2298422	7.43	8.48	1.05	4.73E-07	2.22E-05	intron (Rg0000384, intron 1 of 5)	intron (NM_001273767, intron 1 of 5)	2160	Rg0000384 Rg001113262
chr2R	15479517	15480017	8.97	10.02	1.05	7.13E-06	0.000199	intron (Rg0034021, intron 1 of 4)	intron (NM_001299525, intron 1 of 4)	3067	Rg0034021 Rg0010345399
chr2L	21830190	21830690	7.2	8.25	1.05	3.51E-05	0.000715	intron (Rg0003866, intron 1 of 2)	intron (NM_078891, intron 1 of 2)	658	Rg0003866 Rg001085906
chr2R	21314727	21314777	8.73	9.78	1.05	0.000255	0.00399	intron (Rg0034613, exon 2 of 8)	exon (NM_001169764, exon 2 of 8)	1846	Rg0034613 Rg001071601
chr2L	18853049	18853549	10.04	11.1	1.06	8.85E-08	6.13E-06	promoter-TSS (Rg0032723)	promoter-TSS (NM_1452409)	680	Rg0032723 Rg001085314
chr2L	2976277	2976777	11.22	12.28	1.06	1.86E-07	1.02E-05	promoter-TSS (Rg0023494)	promoter-TSS (NM_134858)	325	Rg0023495 Rg001078179
chr2L	18967104	18967604	9.46	10.52	1.06	1.99E-07	1.08E-05	intron (Rg0015772, intron 1 of 6)	intron (NM_001291146, intron 1 of 6)	1367	Rg0015772 Rg001081185
chr3L	7361658	7362158	10.28	11.35	1.06	4.08E-06	0.000129	exon (Rg00267485, exon 1 of 1)	exon (NM_139854, exon 2 of 4)	1348	Rg00267485 Rg001076847
chr3L	14415323	14415823	9.14	10.2	1.06	1.66E-05	0.000939	intron (Rg00308700, intron 3 of 7)	exon (NM_001144481, exon 3 of 7)	4875	Rg00308700 Rg0010345399
chr2L	12028733	12029233	8.32	9.38	1.06	0.000555	0.0398	intron (Rg0010497, intron 1 of 4)	intron (NM_138760, intron 1 of 4)	1546	Rg0010497 Rg001080292
chr2R	7624073	7624573	8.2	9.26	1.06	0.00879	0.0495	intron (Rg0026527, intron 1 of 14)	intron (NM_001259256, intron 1 of 14)	1082	Rg0026527 Rg0010306933
chrX	19123998	19124498	10.05	11.11	1.07	1.87E-07	1.86E-05	promoter-TSS (NM_001169541)	promoter-TSS (NM_001169541)	-379	Rg0020031 Rg0010346428
chr2L	5278292	5278792	8.17	9.24	1.07	2.42E-06	8.57E-05	Intergenic	Intergenic	-17523	Rg0026383 Rg00303278
chr2R	7958802	7959302	9.63	10.71	1.08	5.67E-05	0.00107	exon (Rg0033240, exon 3 of 5)	exon (NM_001030755, exon 3 of 5)	903	Rg0033240 Rg001113052
chrX	2274004	2274504	9.68	10.77	1.09	1.11E-06	4.49E-05	exon (Rg0031187, exon 3 of 11)	exon (NM_001169174, exon 2 of 3)	524	Rg0031187 Rg001080092
chrX	2245756	2245806	9.68	10.78	1.09	2.28E-06	8.24E-05	intron (Rg00267431, intron 1 of 2)	intron (NM_167776, exon 3 of 11)	923	Rg00267431 Rg00107248
chr3R	727707	728207	8.28	9.38	1.1	6.83E-08	4.19E-06	intron (Rg00267431, intron 1 of 2)	CpG	5336	Rg00267431 Rg001141258
chr2L	21170039	21170539	9.86	10.96	1.1	9.77E-08	5.84E-06	exon (Rg00307103, exon 2 of 4)	exon (NM_001030716, exon 2 of 4)	1124	Rg00307103 Rg001081473
chr2R	1840605	1840655	10.25	11.35	1.1	0.000546	0.00634	promoter-TSS (Rg00063493)	promoter-TSS (NM_137485)	397	Rg00063494 Rg001086674
chr3R	27905196	27905696	7.85	8.95	1.11	5.92E-08	3.82E-06	intron (Rg0058282, intron 1 of 12)	5' UTR (NM_001104475, exon 1 of 12)	295	Rg0058282 Rg0010305292
chr3R	1580585	1580635	8.3	9.41	1.11	5.04E-07	2.29E-05	intron (Rg0026616, intron 1 of 4)	intron (NM_142237, intron 1 of 4)	1858	Rg0026616 Rg001083119
chr2L	589855	589905	7.68	8.8	1.12	3.35E-08	2.63E-06	TTS (Rg0026727)	TTS (NR_1346613)	958	Rg0026727 Rg001045834
chr2L	1862773	1862823	9.07	10.18	1.12	2.76E-07	1.43E-05	intron (Rg0030801, intron 2 of 8)	CpG	2017	Rg0030801 Rg001080574
chr2R	780978	780978	9.48	10.6	1.12	8.14E-10	4.57E-08	promoter-TSS (Rg0051472)	promoter-TSS (NM_169195)	465	Rg0051473 Rg001080174
chr3R	998184	998684	9.22	10.35	1.13	5.35E-08	3.55E-06	Intergenic	Intergenic	47182	Rg0029691 Rg00308892
chr3R	21272495	21272995	9.39	10.53	1.13	1.95E-05	0.000444	exon (Rg0038877, exon 2 of 4)	exon (NM_142706, exon 2 of 4)	1072	Rg0038877 Rg001084102
chrX	3290979	3291479	9.02	10.15	1.13	5.37E-05	0.00102	promoter-TSS (NM_166967)	promoter-TSS (NM_166967)	-14	Rg0000479 Rg001070517
chr2R	15481016	15481516	9.19	10.32	1.13	0.00478	0.0316	intron (Rg0034021, intron 1 of 4)	intron (NM_001299525, intron 1 of 4)	1568	Rg0034021 Rg0010345399
chr3R	15546532	15547032	9.6	10.74	1.14	1.17E-09	1.48E-07	exon (Rg0038349, exon 4 of 9)	exon (NM_001316510, exon 4 of 9)	954	Rg0038349 Rg001083151
chr2L	9915914	9916414	9.88	11.02	1.14	2.34E-06	8.41E-05	exon (Rg0040064, exon 3 of 4)	exon (NM_078804, exon 3 of 4)	912	Rg0040064 Rg001079876
chr3R	29765807	29766307	8.96	10.1	1.14	4.44E-06	0.000138	promoter-TSS (NM_001032197)	promoter-TSS (NM_001032197)	63	Rg001032197 Rg001099988
chr2L	5341671	5342171	9.33	10.47	1.14	0.00102	0.00991	intron (Rg00307103, intron 3 of 2)	intron (NM_135880, intron 3 of 2)	313	Rg00307103 Rg001079077
chr2L	6167005	6167505	9.71	10.85	1.15	2.46E-09	2.97E-07	exon (Rg0021826, exon 2 of 5)	exon (NM_166294, exon 2 of 5)	329	Rg0021826 Rg001080174
chrX	626405	626905	10.6	11.75	1.15	2.91E-09	3.13E-07	TTS (Rg00267031)	TTS (NR_1223663)	1114	Rg00267031 Rg001070116
chr3R	21066756	21067256	8.76	9.91	1.15	2.29E-08	1.80E-06	intron (Rg0000083, intron 1 of 4)	CpG	1124	Rg0000083 Rg00136797
chr3L	11436057	11436557	10.31	11.47	1.15	3.67E-08	2.70E-06	promoter-TSS (NM_141818)	promoter-TSS (NM_141818)	1458	Rg0037902 Rg001082391
chr3R	29148859	29149359	10.66	11.81	1.15	3.41E-07	1.70E-05	promoter-TSS (Rg0039642)	promoter-TSS (NM_143418)	748	Rg0039642 Rg001004882
chr3R	863786	864286	7.88	9.02	1.15	4.70E-07	2.22E-05	intron (Rg00267431, intron 2 of 2)	CpG	141415	Rg00267431 Rg001081258
chr4	1305651	1306151	7.36	8.51	1.15	2.25E-06	8.19E-05	Intergenic	Intergenic	74019	Rg0026363 Rg001091630
chr3R	18853439	18853939	10.11	11.28	1.16	2.17E-10	3.85E-08	promoter-TSS (Rg0034403)	promoter-TSS (NM_137540)	-646	Rg0034403 Rg0010727338
chr3R	31670624	31671124	8.9	10.06	1.16	3.10E-06	0.000189	exon (Rg0039863, exon 3 of 9)	exon (NM_170563, exon 3 of 10)	1410	Rg0039863 Rg001082823
chr2L	19794191	19794691	9.46	10.63	1.16	0.000262	0.00222	exon (Rg00303696, exon 2 of 8)	exon (NM_176363, exon 2 of 8)	1242	Rg00303696 Rg0010302513
chr2L	6945702	6946202	9.98	11.14	1.16	1.65E-05	0.000911	promoter-TSS (Rg0031817)	promoter-TSS (NM_00127242)	-663	Rg0031817 Rg001083116
chr3R	18295667	18296167	8.59	9.76	1.17	0.00042	0.00522	intron (Rg0010877, intron 1 of 3)	intron (NM_143294, intron 1 of 3)	-164	Rg0010877 Rg001033434
chr3L	1479055	1479555	9.57	10.75	1.18	8.83E-10	8.17E-08	promoter-TSS (NM_176273)	promoter-TSS (NM_176273)	-393	Rg0020248 Rg001072713
chrX	3448211	3448711	7.97	9.15	1.18	1.31E-08	1.13E-06	intron (Rg0029657, intron 1 of 6)	intron (NM_001272276, intron 1 of 6)	482	Rg0029657 Rg0010344078
chr3L	9695175	9695675	9.76	10.95	1.19	2.46E-11	5.16E-09	promoter-TSS (NM_168360)	promoter-TSS (NM_168360)	16	Rg0036301 Rg001076416
chr3L	6977484	6977984	8.5	9.69	1.19	0.000253	0.00357	exon (Rg0047205, exon 3 of 3)	non-coding (NR_001944, exon 3 of 3)	1413	Rg0047205 Rg001076899
chr2R	16338012	16338512	9.79	10.99	1.2	2.37E-10	3.60E-08	promoter-TSS (NM_1661681)	promoter-TSS (NM_1661681)	-39	Rg0010611 Rg00108

chr3L	22999735	23000235	626	8.5	2.25	2.68E-16	1.39E-13	intron # Bgn0262509, intron 1 of Z3	DNAREP1_DM RC H3K9me3	1614	FBgn0262509 FBur0304530
chr3L	1474669	1475169	881	11.14	2.33	2.74E-18	1.73E-15	TTS (FBgn0020248)	TTS (NM_176273)	338	FBgn0262467 FBur0310161
chr3L	22997396	22997896	551	7.92	2.41	2.24E-11	4.83E-09	promoter-TSS # Bgn0262509	promoter-TSS (NM_001043156)	725	FBgn0262509 FBur0304530
chr3L	22988072	22988572	755	10.21	2.66	3.15E-35	6.95E-32	promoter-TSS # Bgn0262509	promoter-TSS (NM_001043156)	49	FBgn0262509 FBur0304530
chr3R	27961166	27961666	-918	10.05	10.23	2.83E-116	2.50E-112	promoter-TSS # Bgn0051253	promoter-TSS (NM_170364)	405	FBgn0262408 FBur0346202

II. The concerted activity of epigenomic effectors is essential to establish transcriptional programs during oogenesis

1. Article Presentation

In our first study, a genetic screen allowed us to identify Lid and Sin3A as essential regulators of *dhd* expression. These factors have been found in a repressive complex (Moshkin et al., 2009), which made us wonder if other chromatin complexes were involved in *dhd* regulation. Broadening our analysis to other knockdowns led us to identify two additional factors, essential for *dhd* expression: Snr1, a subunit of the chromatin remodeler Swi/Snf and the chromatin factor Mod(mdg4) mainly known for its insulator role. We thus wondered how are all of these factors promoting the hyperactivation of *dhd*?

To tackle this question, in this second article, I investigated the impact of these different knockdowns on the epigenome, using the chromatin profiling method Cut&Run. I also used a dedicated data analysis strategy to identify potential regulatory elements associated to histone marks.

Of note, this revised version of the article has been accepted by PLOS Genetics.

Three classes of epigenomic regulators converge to hyperactivate the essential maternal gene *deadhead* within a heterochromatin mini-domain.

Daniela Torres-Campana¹ , Béatrice Horard¹ , Sandrine Denaud² , Gérard Benoit¹ , Benjamin Loppin^{1*} & Guillermo A. Orsi^{1,3*}.

¹: Laboratoire de Biologie et Modélisation de la Cellule, CNRS UMR5239, École Normale Supérieure de Lyon, University of Lyon, France.

²: Institute of Human Genetics, UMR 9002, CNRS–University of Montpellier, 34396 Montpellier, France.

³: Current address: Institute for Advanced Biosciences, INSERM U1209, CNRS, Université Grenoble Alpes, France.

*: correspondence: benjamin.loppin@ens-lyon.fr; guillermo.orsi@inserm.fr

Abstract

The formation of a diploid zygote is a highly complex cellular process that is entirely controlled by maternal gene products stored in the egg cytoplasm. This highly specialized transcriptional program is tightly controlled at the chromatin level in the female germline. As an extreme case in point, the massive and specific ovarian expression of the essential thioredoxin Deadhead (DHD) is critically regulated in *Drosophila* by the histone demethylase Lid and its partner, the histone deacetylase complex Sin3A/Rpd3, via yet unknown mechanisms. Here, we identified Snr1 and Mod(mdg4) as essential for *dhd* expression and investigated how these epigenomic effectors act with Lid and Sin3A to hyperactivate *dhd*. Using Cut&Run chromatin profiling with a dedicated data analysis procedure, we found that *dhd* is intriguingly embedded in an H3K27me3/H3K9me3-enriched mini-domain flanked by DNA regulatory elements, including a *dhd* promoter-proximal element essential for its expression. Surprisingly, Lid, Sin3a, Snr1 and Mod(mdg4) impact H3K27me3 and this regulatory element in distinct manners. However, we show that these effectors activate *dhd* independently of H3K27me3/H3K9me3, and that *dhd* remains silent in the absence of these marks. Together, our study demonstrates an atypical and critical role for chromatin regulators Lid, Sin3A, Snr1 and Mod(mdg4) to trigger tissue-specific hyperactivation within a unique heterochromatin mini-domain.

Author Summary

Multicellular development depends on a tight control of gene expression in each cell type. This relies on the coordinated activities of nuclear proteins that interact with DNA or its histone scaffold to promote or restrict gene transcription. For example, we previously showed that the histone modifying enzymes Lid and Sin3A/Rpd3 are required in *Drosophila* ovaries for the massive expression of *deadhead* (*dhd*), a gene encoding for a thioredoxin that is essential for fertility. In this paper, we have further identified two additional *dhd* regulators, Snr1 and Mod(mdg4) and dissected the mechanism behind hyperactivation of this gene. Using the epigenomic profiling method Cut&Run with a dedicated data analysis approach, we unexpectedly found that *dhd* is embedded in an unusual chromatin mini-domain featuring repressive histone modifications H3K27me3 and H3K9me3 and flanked by two regulatory elements. However, we further showed that Lid, Sin3A, Snr1 and Mod(mdg4) behave like obligatory activators of *dhd* independently of this mini-domain. Our study unveils how multiple broad-acting epigenomic effectors operate in non-canonical manners to ensure a critical and specialized gene activation event. These findings challenge our knowledge on these regulatory mechanisms and their roles in development and pathology.

Introduction

Gene expression is tightly controlled in eukaryotic cells by the composition, organization and dynamics of nucleosomes, consisting of an octamer of histone proteins wrapped in ~146bp of DNA. The concerted activity of protein complexes including histone chaperones, readers and writers as well as nucleosome remodelers, defines the positioning, composition and post-translational modifications of nucleosomes [1–3]. The resulting chromatin landscape is further organized by insulator proteins that delimit tridimensional contacts along the genome, forming sub-nuclear domains and guiding contacts between promoters and their cognate regulatory elements [4]. This tightly regulated epigenomic environment profoundly influences RNA Polymerase access to DNA and transcriptional activity.

Tremendous efforts in the past decades aimed at dissecting the roles of these epigenomic effectors *in vivo*. A privileged method is ablation or dosage manipulation of each component to measure its impact on gene expression. While these approaches can yield precious functional insight, the ubiquitous expression and wide range of activities of these factors, as well as redundancies in their interactions, make it difficult to infer their precise function. Understanding their function therefore requires identifying biologically relevant situations where disrupting these effectors impacts transcription in a critical and specific manner. We previously described one of such cases, where perturbation of the histone H3K4 demethylase Lid/KDM5 or the histone deacetylase complex Sin3A/Rpd3 in *Drosophila* ovaries dramatically abrogated the expression of the maternal gene *deadhead* (*dhd*), which is essential for female fertility [5].

The *Drosophila* egg is loaded with maternal gene products synthesized by germline nurse cells that enable early embryonic development in the absence of zygotic

transcription [6]. An extreme example of this specialized transcriptome, *dhd* is among the most highly expressed genes in adult ovaries, while it is almost completely silent in any other tissue and developmental stage [5,7–9]. The DHD protein is a thioredoxin involved in regulating the general redox state in oocytes [10,11]. In addition, DHD plays a critical role at fertilization to reduce cysteine-cysteine disulfide bonds on the Protamine-like proteins that replace histones on chromatin during spermiogenesis [9,12]. In the absence of DHD, paternal chromosomes fail to decondense and are excluded from the first zygotic nucleus, leading to haploid gynogenetic development and embryonic lethality. The *dhd* locus, which produces a single, short (952bp), intronless transcript is packed within a 1369bp region that separates its flanking genes *Trx-T* and *CG4198*. Remarkably, these two genes are expressed exclusively in the male germline, thereby constituting an apparently unfavorable environment for *dhd* transcription in ovaries. In addition, we showed that a 4305bp transgene spanning only *Trx-T*, *dhd* and part of *CG4198* largely recapitulates the expression of *dhd* [5,9], indicating that regulatory elements sufficient for *dhd* activation are contained within this restricted region. Our previous study further found that Lid and Sin3A are essential activators of *dhd* in *Drosophila* ovaries, in striking contrast to their otherwise relatively modest impact on the rest of the transcriptome. Considering these unusual features, we postulated that the exquisite sensitivity of *dhd* to these broad-acting chromatin effectors revealed a singular mode of epigenomic regulation that enables its massive and specific ovarian expression [5].

Here, we exploited this singular model locus to understand how multiple classes of epigenomic effectors converge to achieve programmed transcriptional hyperactivation. We identified the Brahma chromatin remodeler component Snr1 [13] and the BTB/POZ-domain protein Mod(mdg4) [14] as factors that share with Lid and

Sin3A a critical and highly specific role in activating *dhd*. By exploiting the chromatin profiling method Cut&Run [15] and an adapted data analysis strategy, we found that *dhd* is unexpectedly embedded within a heterochromatin mini-domain flanked by two border regulatory elements. One of these is a *dhd*-proximal element, which encompasses a DNA Replication-related Element (DRE-box) motif [16] that is essential for *dhd* expression. Yet, exploiting knockdown and transgenic tools, we found that Lid, Sin3A, Snr1 and Mod(mdg4) activate *dhd* independently of the associated heterochromatin mini-domain. Furthermore, this mini-domain is not required to restrict *dhd* expression to ovaries. Together, our results put into perspective our understanding on these epigenomic regulators by revealing how they exert a biologically essential control of *dhd* via non-canonical mechanisms.

Results

Mod(mdg4) and Snr1 are essential for *dhd* expression

We previously performed a female germline RNA interference screen to identify chromatin factors required for paternal chromosome incorporation into the zygote at fertilization. As part of that screen, the histone H3K4 demethylase Lid, Sin3A and Rpd3, which participate in deacetylase complexes targeting various lysine residues in H3 and H4 [17,18], were identified as essential regulators of *dhd* expression. Because Lid and Sin3A can interact within a co-repressor complex [19,20], we asked whether other chromatin regulatory complexes might also be involved in *dhd* regulation. We therefore broadened our analysis to other knockdowns that caused maternal effect sterility associated with a *dhd*-like mutant phenotype, i.e. defective sperm nuclear decompaction

at fertilization. Among these, we focused on two additional UAS-controlled small hairpin RNA (shRNA) constructs from the TRiP collection [21], respectively targeting *mod(mdg4)* and *Snr1*. *Snr1* is an essential subunit of the Brahma chromatin remodeler that mediates protein-protein interactions within this complex as well as with external interacting partners [13,22]. The *mod(mdg4)* gene codes for up to 31 isoforms [23], all of which are targeted by the shRNA construct. Among these, the most well characterized, *Mod(mdg4)^{67.2}* is a common component of boundary insulators in the *Drosophila* genome [24], but other non-insulating isoforms exhibiting activator functions have also been identified [4,25,26]. These two candidates belonged to two classes of epigenomic effectors distinct from *Lid* and *Sin3A*, and we thus decided to investigate their function during the oocyte to zygote transition.

When activated by the Maternal Triple Driver (MTD) Gal4 source, these shRNAs efficiently reduced the levels of *mod(mdg4)* and *Snr1* transcripts (FigS1-A). Previous studies reported defective oogenesis and diminished egg production in *mod(mdg4)* as well as *Snr1* mutant females [22,25]. Consistently, females with ovarian knockdown of *mod(mdg4)* or *Snr1* (hereby referred to as *mod(mdg4)* KD or *Snr1* KD females) were almost completely sterile (Table 1). Indeed, while KD females were able to lay more eggs than mutants, these almost systematically failed to hatch. Focusing on paternal chromatin organization at fertilization in these embryos, we found that both *mod(mdg4)* and *Snr1* ovarian KDs systematically led to failure of male pronucleus decondensation, which remained elongated (Fig1-A). Concomitantly, these eggs exhibited retention of the protamine fluorescent marker *Mst35Ba::GFP* (*ProtA::GFP*) [27] in paternal chromatin, as observed in *dhd* loss of function mutants [9,12].

Table 1. Embryo hatching rates.

The *w¹¹¹⁸* strain is used as reference.

<i>Knockdowns</i>			
Female Genotype	Male Genotype	Number of eggs	Hatch. rate (%)
<i>Control</i>	<i>w¹¹¹⁸</i>	1561	98.27%
<i>Snr1</i> KD	<i>w¹¹¹⁸</i>	791	0.00%
<i>mod(mdg4)</i> KD	<i>w¹¹¹⁸</i>	1589	1.01%
<i>lid</i> KD (val22)	<i>w¹¹¹⁸</i>	1403	2.14%
<i>lid</i> KD (val21)	<i>w¹¹¹⁸</i>	1144	1.05%
<i>Sin3a</i> KD	<i>w¹¹¹⁸</i>	1221	0.25%
<i>E(z)</i> KD	<i>w¹¹¹⁸</i>	843	0.00%
<i>Rescue with WT or ΔDRE mutant transgene</i>			
<i>w¹¹¹⁸</i>	<i>w¹¹¹⁸</i>	344	97.67%
<i>dhd^{Δ5}</i>	<i>w¹¹¹⁸</i>	375	0.00%
<i>dhd^{Δ5}::;pW8-dhd^{WT}</i>	<i>w¹¹¹⁸</i>	663	85.67%
<i>dhd^{Δ5}::;pW8-dhd^{ΔDRE}</i>	<i>w¹¹¹⁸</i>	475	2.15%
<i>dhd^{Δ5}::;pW8-dhd^{FD}</i>	<i>w¹¹¹⁸</i>	410	87.80%
<i>Knockdown rescue with the WT transgene</i>			
<i>dhd^{Δ5}::; lid KD(val21), pW8-dhd^{WT}</i>	<i>w¹¹¹⁸</i>	175	1.92%
<i>dhd^{Δ5}::; Sin3a KD, pW8-dhd^{WT}</i>	<i>w¹¹¹⁸</i>	271	0.37%
<i>dhd^{Δ5}::; Snr1 KD, pW8-dhd^{WT}</i>	<i>w¹¹¹⁸</i>	247	0.00%
<i>dhd^{Δ5}::; mod(mdg4) KD, pW8-dhd^{WT}</i>	<i>w¹¹¹⁸</i>	462	0.43%

The above results suggest that Mod(mdg4) and Snr1 could regulate *dhd* expression. RNA-sequencing on *mod(mdg4)* and *Snr1* KD ovaries indeed revealed that *dhd* is dramatically downregulated in both KDs, with a fold reduction of almost two orders of magnitude (Fig1-B,C, and FigS1-B,E). *dhd* was the first most strongly affected gene in *mod(mdg4)* KD ovaries in terms of fold-change in expression, and the 14th most affected gene in *Snr1* KD ovaries. Consistently, DHD protein levels assessed by Western Blot in KD ovaries were also dramatically reduced (Fig1-D). This was in contrast to a more modest impact of both KDs on the rest of the transcriptome and the limited overlap in their effects (Fig S1-B,C,D). In particular, genes in the vicinity of *dhd* were not significantly affected by the KDs (Fig1-C and FigS1-B). Therefore, despite the packed

genomic organization of the *dhd* locus, its expression strictly and singularly depends on multiple epigenomic effectors.

Cut&Run with dedicated analysis reveals both the distribution of histone modifications and their associated regulatory elements

To more precisely characterize the chromatin landscape at the *dhd* locus, we next implemented the Cut&Run epigenomic profiling method [15]. In Cut&Run, histone modifications of interest are targeted *in situ* by a specific antibody following tissue permeabilization. Target-bound antibodies are subsequently coupled to a fusion between the bacterial Protein A and Micrococcal Nuclease (ProteinA-MNase) that cleaves exposed DNA in the vicinity of the antibody, releasing target nucleosomal particles into solution. Importantly, MNase is expected to also cleave exposed DNA in the immediate spatial vicinity of the nucleosome-bound antibody, causing the release of DNA particles bound by other proteins such as polymerases or DNA sequence-specific transcription factors (Fig2-A). In particular, DNA regulatory elements occupied by sequence-specific transcription factors are typically associated with MNase footprints distinctly shorter than nucleosomes [28–30]. Partially unwrapped dynamic nucleosomes typically associated with regulatory elements can also produce such distinctly short footprints [31]. Following paired-end sequencing, such released DNA fragments can be distinguished and separated by their size, yielding a map of nucleosomes (>146bp) and sub-nucleosomal particles, putatively corresponding to

regulatory elements (<120bp). A single Cut&Run experiment should thus identify DNA regulatory elements that are in physical proximity of target histone modifications.

With this in mind, we conducted H3K27me3 Cut&Run in *Drosophila* ovaries. Using only 12 pairs of ovaries per sample, we robustly revealed H3K27me3 domains. Remarkably, visualization of Cut&Run fragments shorter than 120bp (which excludes fully wrapped octameric nucleosomes) revealed that these were enriched at discrete peaks within H3K27me3 domains. Genome-wide analysis identified 679 peaks of fragments <120bp (hereon referred to as “short fragment peaks”) that were ~250bp-wide in average (Fig2-B,D). We hypothesized that short fragment peaks represented H3K27me3-associated regulatory elements occupied by transcription factors. Within H3K27me3 domains, we expected these to include Polycomb Response Elements (PREs) as well as insulators. For example, short fragment peaks corresponded to several well-described PREs and insulators in the Bithorax complex H3K27me3 domain [26,32,33] (Fig2-C), consistent with observations in larval tissue [34]. To ask whether this reflects a broader genome-wide trend, we compared short fragment peaks with PRE and insulator markers genome-wide. Although there is scarce genome-wide data available for *Drosophila* ovaries, H3K27me3 domains are generally present in most cell types. We thus exploited datasets from embryonic-derived S2 and Kc cell lines. Consistent with their occupancy by transcription factors, ATAC-seq peaks [35] -revealing hyper-accessible DNA- coincide with Polycomb regulatory elements in flies and mice [35,36]. Genome-wide, our small fragment peaks identified in ovaries were enriched for ATAC-seq signal, arguing that these indeed correspond to DNA regulatory elements (Fig2-D). Enrichment at these peaks of the Polycomb protein [37] and the insulator protein CP190 [38] further argues that these elements often correspond to functional PREs or insulators. Accordingly, at the borders of H3K27me3 domains, short fragment peaks

were more frequently associated with CP190, confirming previous reports that this factor is associated with H3K27me3 domain boundaries [24,39] (Fig2-D). Instead, Polycomb was rather enriched at peaks localized internally within these domains. Our Cut&Run analysis strategy can therefore be used to reveal not only the breadth of histone modification domains in ovaries but also their associated DNA regulatory elements.

***dhd* lies within an H3K27me3/H3K9me3 mini-domain flanked by DNA regulatory elements.**

To gain insight on *dhd* regulation, we next sought to analyze its associated chromatin configuration. We previously showed that the active transcription modification H3K4me3 is enriched at the *dhd* promoter and that this mark is lost in *lid* KD ovaries [5](FigS2-A). Using available ChIP-seq datasets from embryonic derived S2 cells, we further observed that that *dhd* lies within a ~5kbp mini-domain featuring two types of repressive histone modifications: H3K27me3 and H3K9me3 (FigS2-A) [40]. H3K27me3 is the hallmark of Polycomb-based repression [41,42], whereas H3K9me3 dictates Heterochromatin Protein 1 (HP1)-based repression [43,44]. Interestingly, this mini-domain was also found in ChIP-seq data from fly ovaries (FigS2-A) [45]. Potentially regulating this chromatin environment, Lid and Sin3A were described as participating in a co-repressor complex [20], but their global impact on repressive histone modifications is unclear. In turn, previous reports showed that depletion of insulator proteins Mod(mdg4), as well as CTCF, Su(Hw), CP190 or BEAF-32, did not affect the spread of Polycomb-associated domains but instead caused a general decrease in H3K27me3 levels [24]. In contrast, the Brahma/BAF complex is typically considered as

counteracting Polycomb repression based on work in mammals [46], but this interplay has not been analyzed in *Drosophila*.

Based on these observations, we sought to better characterize the distribution of H3K27me3 and H3K9me3 in ovaries at the *dhd* locus. Using our Cut&Run approach, we confirmed that *dhd* is included in a ~5450bp heterochromatic H3K27me3/H3K9me3 mini-domain that extends from the promoter region of *dhd* to the promoter of the next gene active in ovaries, *Sas10* (Fig3-A). While our whole-tissue Cut&Run approach cannot distinguish which cells harbor this domain, published H3K27me3 ChIP-seq profiles from either somatic or germline cells in ovaries indicate that this mark is present in both (Figure S2-A). We segmented the Cut&Run H3K27me3 signal in control ovaries and identified 278 discrete H3K27me3 domains, ranging from 3 to 240kb in width. The *dhd* H3K27me3 domain stood out when considering its enrichment in this mark relative to its length, compared to other domains (Fig3-B). Our analyses indicate that despite its reduced size, the *dhd* locus is capable of accumulating proportionately high amounts of this repressive modification. While we also identified other previously described H3K9me3 domains lodged in euchromatic regions (FigS3), the signal is, as expected, largely dominated by pericentric heterochromatin. We could thus not robustly call such euchromatic H3K9me3 domains, preventing us from conducting an analogous analysis for *dhd* on this mark.

Surprisingly, short fragment peak analysis revealed two putative DNA regulatory elements associated with both histone marks, precisely at the mini-domain borders, with no internal peaks present (Fig3-A). ATAC-seq data from S2 cell lines confirmed presence of only these two border elements (FigS2-B). In addition, in Kc cells, *dhd* border elements are occupied by CP190 and Mod(mdg4), both of which can be found at the boundaries of *Drosophila* H3K27me3 domains [24,47]. However, the insulating

isoform Mod(mdg4)^{67.2} is not found at the *dhd* locus (FigS2-B), suggesting that a different isoform playing an activator function is responsible for *dhd* regulation. Finally, the *dhd*-proximal 5' border element featured a significant, although very modest enrichment for PRE markers Polycomb and Polyhomeotic (Fig-S2-B). This regulatory architecture was quite unusual, as we could not find any other H3K27me3 domain in the genome sharing this particular organization with two border elements and no internal elements. Together, these results revealed that *dhd* lies within a unique H3K27me3/H3K9me3 mini-domain featuring only border elements.

Sin3A, Snr1 and Mod(mdg4) control the regulatory architecture of the *dhd* mini-domain.

We next aimed at evaluating the potential role for Lid, Sin3A, Mod(mdg4) and Snr1 in regulating this heterochromatin mini-domain. We used KD ovaries for these factors and included as a control a KD for the H3K27 methyltransferase Enhancer of zeste (*E(z)*), induced in germ cells by the MTD-Gal4 driver. While *E(z)* KD females were sterile as previously described [48,49] (Table1), they were able to lay eggs and displayed only a moderate effect on *dhd* expression (a 25% reduction compared to controls) (FigS4-A). Immunofluorescence staining on dissected control ovaries showed that H3K27me3 marks follicle cell nuclei, the karyosome (i.e the oocyte nucleus) and nurse cell nuclei, although nurse cell staining was relatively weaker (FigS3-B), consistent with previous reports [48]. As expected, H3K27me3 was undetectable in the karyosome and in nurse cells of *E(z)* KD ovaries, whereas follicle cells (which do not express MTD-driven shRNAs) still carried this mark at normal levels. While *lid*, *Sin3a* and *mod(mdg4)* KD ovaries displayed normal H3K27me3 staining, we observed a

moderate reduction in H3K27me3 levels in nurse cells in *Snr1* KD ovaries, even while H3K27me3 levels were not affected on the karyosome (FigS4-B).

We next carried out H3K27me3 Cut&Run on ovaries from all KDs. Within our 278 identified H3K27me3 domains (see above), we compared the average enrichment in H3K27me3 signal in control and KD ovaries (Fig3-C). In *E(z)* KD ovaries, Cut&Run experiments revealed only a moderate loss of H3K27me3 signal (35% average reduction at these domains compared to controls) (Fig3-C), contrasting with the strong global reduction in H3K27me3 immunofluorescence signal. This difference is likely to reflect the fact that the H3K27me3 signal from Cut&Run experiments originates from both germline and somatic cells. Accordingly, *E(z)* KD completely abrogated H3K27me3 signal at the *spen*, *Corto* or *ptc* loci, all of which are decorated with H3K27me3 in nurse cells but not in follicle cells (FigS4-C) [50]. In contrast, the *gl*, *dpp*, or *repo* loci, which show stronger H3K27me3 in follicle cells compared to nurse cells, were only slightly affected in *E(z)* KD ovaries (FigS4-C). Together, these results show that our Cut&Run strategy detects H3K27me3 signal from both germline and somatic cells and is able to detect quantitative differences in the averaged signal when nurse cells are strongly affected. Consistent with immunofluorescence experiments, *lid*, *Sin3a* and *mod(mdg4)* KDs had only a modest global impact on average H3K27me3 levels (5% reduction compared to controls), and no effect on the spread of H3K27me3 domains (Fig3-C). Also consistent with our immunofluorescence experiments, *Snr1* KD led to a more severe average reduction of H3K27me3 Cut&Run signal compared to controls (20%), although not as dramatic as *E(z)* KD.

In agreement with genome-wide observations, the levels of H3K27me3 in the *dhd* mini-domain were reduced in *E(z)* KD ovaries and unaffected in *Sin3a* or *mod(mdg4)* KD ovaries. More surprisingly, the domain was not measurably affected in

Snr1 KD ovaries, despite the fact that H3K27me3 is globally impacted by this knockdown (Fig3-D and FigS5-A). Within the sensitivity limits of our approach, these results indicate that *Sin3a*, *Snr1* and *mod(mdg4)* KDs have little if any impact on H3K27me3 at the *dhd* locus. Conversely, in *lid* KD ovaries, in which global H3K27me3 levels were unaffected, we detected an increase in H3K27me3 levels at the *dhd* mini-domain (Fig3-C,D and FigS5-A). This raised the possibility that Lid could facilitate *dhd* expression by counteracting Polycomb-mediated repression.

Since the *dhd* mini-domain also featured H3K9me3, we next turned to Cut&Run followed by qPCR to evaluate its status in KD ovaries. H3K27me3 Cut&Run-qPCR measures the expected enrichments at H3K27me3 domains and detects variations in the signal coherent with Cut&Run-seq results (Fig3-E and FigS5-B). To validate the H3K9me3 Cut&Run-qPCR approach in ovaries, we exploited the *CG12239* gene as a positive control [45], and detected an expected enrichment in H3K9me3 signal at this locus (Fig3-E and FigS5-B). At the *dhd* locus, H3K9me3 was enriched as expected from ChIP-seq results. Importantly, knockdown of *lid*, *Sin3a*, *mod(mdg4)* or *Snr1* had no effect on this enrichment. (Fig3-E). The *dhd* heterochromatin mini-domain including H3K27me3 and H3K9me3 is thus independent of Sin3A, Snr1, whereas Lid counteracts H3K27me3.

We next evaluated the impact of different KDs on the *dhd* mini-domain short fragment peaks at border regions. We first analyzed the effect of our different knockdowns on the full set of 679 peaks previously defined (Fig2-B). Both *E(z)* and *Snr1* KD led to a strong (~63%) decrease in short fragment peak average counts genome-wide (Fig4-A). Since these KDs also affect global H3K27me3 levels, this reduction could result from a general absence of histone modification-targeted MNase on chromatin. Remarkably, *Sin3a* KD led to a similarly strong effect on short fragment peak counts that

could not be attributed to its global impact on H3K27me3. Instead, this data suggests that Sin3A is required to ensure proper occupancy and organization of transcription factors and/or nucleosomes at DNA regulatory elements associated with H3K27me3. In contrast, *lid* or *mod(mdg4)* KD did not globally affect short fragment peak counts, indicating that these factors do not play such a role (Fig4-A).

Consistent with their effects genome-wide, short fragment counts at the *dhd* mini-domain border elements were strongly diminished upon *E(z)* and *Sin3a* KDs (Fig4-B and FigS5-C). Intriguingly, *mod(mdg4)* KD led to a similar impact on these border elements (particularly the *dhd*-proximal one), even though it did not globally affect H3K27me3-associated elements genome-wide (Fig4-B and FigS5-C). This observation could indicate that the *dhd* border elements become less frequently occupied by transcription factors, that these factors become less frequently associated with H3K27me3, and/or that their nucleosomal organization is compromised. In all cases, this suggests that Mod(mdg4) is required to ensure chromatin organization of the border DNA regulatory elements at the *dhd* mini-domain. Remarkably, *Snr1* KD led to a similar effect on border elements without affecting H3K27me3 levels at the *dhd* mini-domain, suggesting that Snr1 is also required for the proper organization of the *dhd* border elements. In striking contrast, *lid* KD had no detectable effect on these regulatory elements (Fig4-B and FigS5-C). We concluded that Lid, although essential for *dhd* expression, was not required to ensure the proper organization of *dhd* border elements.

Altogether our results, summarized in Fig4-C, indicate that Lid, Sin3A, Snr1 and Mod(mdg4), impact H3K27me3 or its associated regulatory elements genome-wide and/or at the *dhd* mini-domain in four distinct manners.

The *dhd* promoter-proximal DNA regulatory element is required for *dhd* expression independently of its heterochromatin mini-domain.

We next performed sequence analysis of the *dhd* mini-domain border elements, screening against the flyreg.v2 transcription factor DNA binding motif database [51,52]. At the 5' border element, which mapped to the *dhd* promoter region, we identified four perfect matches for the DNA replication-related element (DRE) motif, TATCGATA (Fig5-A). This motif is recognized by the insulator-associated factor BEAF-32 [53] and the core-promoter factor DREF [16]. These four DRE motifs overlap in the palindromic sequence TATCGATATCGATA, 37bp upstream of the *dhd* transcription start site. Consistently, BEAF-32 and DREF both occupy this element in Kc cells (FigS6) [38]. Previous studies showed that BEAF-32 null females are partially fertile (~40% hatching rate) [54], indicating that this factor is not essential for *dhd* expression. In turn, DREF is essential in a cell-autonomous manner and indeed *dref* mutations cause oogenesis defects [55]. Accordingly, we observed severe atrophy and failure to produce oocytes in *dref* KD ovaries. Because this precluded studying the role of DREF in *dhd* regulation, we instead sought to probe the importance of the DRE motifs themselves.

The *dhd*^{l5} null allele is a 1.4 kb deletion affecting the entire promoter region including the promoter-proximal regulatory element, and part of the coding region of *dhd* [7,9] (Fig5-A). A *pW8-dhd*^{WT} transgenic construct, bearing the entire *dhd* gene - including its promoter region-, restores *dhd* expression as well as fertility in *dhd*^{l5} mutants [9] (Fig5-A,B, Table1). We now constructed a second rescue transgene based on the *pW8-dhd*^{WT}, where the 14bp carrying the DRE motifs were deleted (*pW8-dhd*^{ADRE})

(Fig5-A). These constructs were inserted into the same genomic location as *pW8-dhd^{WT}* (62E1) and combined with the *dhd^{l5}* deficiency. In striking contrast to *pW8-dhd^{WT}*, the *pW8-dhd^{ADRE}* construct was unable to rescue *dhd* expression, or substantially improve fertility in *dhd^{l5}* deficient flies (Fig5-B, Table1). The DRE motifs are thus essential to ensure *dhd* expression.

To test a role for this regulatory element and its DRE motifs in regulating the H3K27me3/H3K9me3 mini-domain, we performed Cut&Run-seq and Cut&Run-qPCR on homozygous *dhd^{l5}* ovaries, as well as rescue *dhd^{l5}::pW8-dhd^{WT}* and non-rescued *dhd^{l5}::pW8-dhd^{ADRE}* ovaries. Strikingly, the 5.4kbp *dhd* H3K27me3 mini-domain was completely lost in *dhd^{l5}* ovaries (Fig5-C,D and FigS7), despite the fact that 90% of this mini-domain were intact in the deficient chromosome. This indicates that the *dhd*-proximal border of this mini-domain is essential for establishment and/or maintenance of H3K27me3. Furthermore, H3K27me3 signal was absent within the mini-domain in *dhd^{l5}::pW8-dhd^{WT}* rescue ovaries (Fig5-C,D and FigS7), suggesting that the 5'-most 2.8kbp of the domain are also insufficient to establish and/or maintain H3K27me3. This result further confirms that *dhd* can be expressed at high levels in the absence of H3K27me3, consistent with results from *E(z)* KD ovaries (FigS4-A). Finally, the H3K27me3 mini-domain was also completely absent in *dhd^{l5}::pW8-dhd^{ADRE}* ovaries (Fig5-C), indicating that the DRE motifs are required for *dhd* expression independently of H3K27me3.

Based on these observations, we hypothesized that the complete mini-domain sequence, including both border regulatory elements, might be necessary for restoring heterochromatin marks. We thus constructed a transgene containing the full domain sequence of the *dhd* heterochromatin domain (*pW8-dhd^{FD}*)(Fig5-A), inserted at the same genomic location as the *pW8-dhd^{WT}* transgene. Interestingly, this transgene

restored *dhd* expression (Fig5-B) and rescued fertility (Table 1) but was unable to restore H3K27me3 (Fig5-D and FigS7). These results indicate that the border-to-border mini-domain is not autonomous and suggest that its genomic location impacts its chromatin configuration.

We next focused on the H3K9me3 mark. In contrast to H3K27me3, Cut&Run analysis in *dhd^{l5}* mutants showed that H3K9me3 was lost at the *dhd*-proximal half of the domain, while this mark was maintained at the *dhd*-distal part (Fig5-C). Cut&Run-qPCR using primers across the *dhd* domain revealed that H3K9me3 was not restored at the *dhd*-proximal part of the H3K9me3 domain in *dhd^{l5}::pW8-dhd^{WT}* or *dhd^{l5}::pW8-dhd^{FD}* ovaries (Fig5-D and FigS7). Together, these results indicate that the border-to-border mini-domain is not autonomous to establish its own heterochromatin configuration, and that *dhd* expression can proceed at near normal levels independently of these marks.

Lid, Sin3A, Snr1 and Mod(mdg4) activate *dhd* independently of its heterochromatin mini-domain.

The fact that the *pW8-dhd^{WT}* transgene restored most of *dhd* expression without re-establishment of the heterochromatin mini-domain at this locus provided an opportunity to clarify the role of our set of *dhd* regulators. KD of *lid* is associated with increased H3K27me3 at the *dhd* mini-domain, suggesting that Lid may operate as an anti-repressor by counteracting heterochromatinization of the locus. However, we have previously found that *dhd* expression is not re-established in *lid* KD ovaries carrying a *pW8-dhd^{WT}* rescue transgene [5]. Lid is thus required for *dhd* expression not only at its endogenous locus but also from the rescue transgene not decorated by H3K27me3

(Fig5). Therefore, Lid activates *dhd* independently of heterochromatin, suggesting that it does not operate strictly as an anti-repressor.

To discriminate between anti-repressive or activating roles of Sin3A, Mod(mdg4) and Snr1, we generated flies combining a *dhd^{J5}* deficiency, the *pW8-dhd^{WT}* transgene and an shRNA targeting *lid*, *Sin3a*, *Snr1* or *mod(mdg4)*, driven in germ cells by a *nos-Gal4* driver (Fig6-A). We confirmed by RT-qPCR that knockdowns were still efficient when using this driver (Fig-S8-A). Remarkably, all of these flies were almost completely sterile, and showed strong downregulation of *dhd* revealed by RT-qPCR (Fig6-B and Table1). Using Cut&Run-qPCR at the *dhd* locus, we further confirmed that these knockdowns had no effect on H3K27me3, which remained depleted in all conditions (Fig6-C and FigS8-B). Lid, Sin3A, Snr1 and Mod(mdg4) therefore stimulate *dhd* transcription in the absence of its heterochromatin mini-domain.

Our results indicate that the *dhd* heterochromatin mini-domain does not play a repressive role in ovaries, but do not exclude that it might maintain *dhd* silent in other tissues. RT-qPCR analysis on dissected ovaries, testes and male and female carcasses from transgenic lines expressing *dhd^{J5};pW8-dhd^{WT}* revealed *dhd* expression uniquely from ovaries (Fig6-D). Because this transgene rescues *dhd* expression without restoring heterochromatin marks, these results suggest that the *dhd* heterochromatin mini-domain is not essential to repress ectopic *dhd* expression in adults. We note, however, that we cannot exclude that *dhd* was weakly and/or transiently expressed in certain cell types in these conditions, or that the rescue transgene could accumulate repressive marks in tissues other than ovaries.

Discussion

The ovarian hyperactivation of *dhd*

Here, we sought to understand how the genomic and epigenomic environments of *dhd* contributed to its remarkable regulation, its expression being both among the highest in *Drosophila*, and absolutely specific to adult ovaries [5,9]. Lid, Sin3A, Snr1 and Mod(mdg4) all shared a critical and rather specific role in ensuring *dhd* expression. Yet, these four broadly expressed proteins play multiple roles other than *dhd* regulation. For example, transcriptomic analyses following individual depletion of Lid, Sin3A or Snr1 in S2 cells, wing discs or pupae shows activation or repression of hundreds of targets [19,22,56]. ChIP-seq data further indicates that Mod(mdg4), Sin3A and Lid each target several thousand sites in the genome [24,26,57,58]. Consistently, our RNA-seq analyses did reveal that each of these knockdowns were associated to up- or down-regulation of 407 to 2020 genes in ovaries, with *dhd* being in every case among the most strongly dependent on these factors. We propose that *dhd* is a hypersensitive gene that reacts radically to epigenome imbalances.

The key question is therefore what is the formula for *dhd* ovarian hyperactivation. One reasonable hypothesis was that *dhd* could be highly regulated by distal enhancers. This would be notably consistent with the previously described role of Mod(mdg4) in organizing 3D contacts between regulatory elements and promoters [26]. It would also be consistent with recent findings that H3K27me3 micro-domains may reflect such contacts [59]. However, no interaction between *dhd* and any other locus can be found in Hi-C data, and our rescue transgene experiments show that a small, ectopic genomic segment almost fully recapitulates its expression, arguing that

the genomic and epigenomic environment at the endogenous *dhd* region may play only a minor role in its ovary-specific hyperactivation.

We indeed found a key regulatory element containing a tandem DRE motif, known to recruit the DREF core promoter factor. The minimal DRE motif (TATCGATA) is found in thousands of gene promoters [60], while multiple genes were individually shown to require this motif for proper activation. These include genes with ovarian expression, and, accordingly, DREF mutations cause oogenesis defects and female sterility [55]. In contrast, the particular tandem DRE motif in the *dhd* regulatory sequence is uncommon, being only found in 9 other gene promoters. Yet, among these 9 genes, only 4 displayed an expression bias in ovaries, and none were nearly as highly transcribed as *dhd*. Therefore, this motif does not seem to be autonomously sufficient for ovarian hyperexpression.

Another unusual feature of *dhd* is its surrounding heterochromatin mini-domain bearing both H3K27me3 and H3K9me3 marks, as well as H3K4me3. The co-occurrence of these active and repressive modifications at an ensemble of developmentally regulated genes in mammals led to the concept of bivalent promoters [61]. It is speculated that such promoters may be poised for rapid activation or repression upon differentiation. In *Drosophila*, bivalent chromatin is associated with genes that can be strongly activated in a tissue-specific manner [62,63]. Our experiments showed that *dhd* is expressed at ~60-70% of its normal levels in *E(z)* KD ovaries, as well as in rescue transgenes - conditions in which the H3K27me3/H3K9me3 mini-domain is impaired. We thus cannot exclude that these heterochromatin marks play a positive role in *dhd* activation to ensure its transcription at maximum capacity, perhaps via establishment of a bivalent configuration. We also note that our whole-tissue experiments leave the

possibility open that these histone modifications may decorate *dhd* in different cell types and/or at different times during gametogenesis.

Altogether, we uncovered multiple unusual genomic and epigenomic characteristics at the *dhd* locus, but failed to identify any single feature that was truly defining. The dramatic regulation of *dhd* may rely not on any individual trait but rather on a unique combination of such rare features. Further work will be needed to elucidate how these different components may together achieve ovarian hyperexpression.

A non-canonical chromatin domain

The unique properties of *dhd* led us to uncover interesting features of its epigenomic regulators. First, *dhd* is embedded in an H3K27me3 mini-domain flanked by regulatory elements. A recent report suggested that H3K27me3 domain borders may be established independently of PREs or border elements, provided that an immediately neighboring active gene instead delimits H3K27me3 spreading [47]. The case of *dhd* is however peculiar in that the H3K27me3 domain border overlaps with this highly active gene, a scenario that was not found in other domains. The coincidence of H3K27me3 and H3K9me3 is also uncommon and, in fact, we could not find any other such dual domain in ovaries. Our result favors the view that this heterochromatin domain does not silence *dhd* expression. Nonetheless, H3K9me3 was always maintained at the *dhd*-distal portion of the domain. We therefore cannot exclude that this mark represses *dhd* neighbors.

An intriguing question is then how this heterochromatin mini-domain is formed. We found that a transgene containing the full mini-domain sequence is unable to restore H3K27me3 or H3K9me3, suggesting that genomic location of this domain is a

critical determinant of *dhd* heterochromatin. From this perspective, the border elements may act as weak PREs, as described in other contexts [64]. Interestingly, we also showed that H3K9me3 can be partially maintained at the distal part of the domain in *dhd^{l5}* mutants while H3K27me3 is completely lost. This indicates that these marks are not necessarily inter-dependent at this locus, and H3K9me3 may benefit from additional mechanisms ensuring its deposition. While not much is known on highly localized euchromatic deposition of H3K9me3, Smolko and colleagues suggested that Setdb1-dependent accumulation of H3K9me3 at certain target genes is dependent on the RNA-binding protein Sxl [45]. At the *dhd* domain, different mechanisms could thus ensure H3K27me3 and H3K9me deposition, both of which would depend on the endogenous genomic location.

Along these lines, another recent study reported the existence of H3K27me3 micro-domains (typically 2-8 nucleosomes wide) that depend on 3D contacts with larger H3K27me3 domains, mediated, in particular, by BEAF-32 and CP190 [59]. The *dhd* mini-domain is wider and much more strongly enriched in H3K27me3 than typical micro-domains. Nonetheless, our data is consistent with a model whereby H3K27me3 could be deposited via such looping interactions. First, BEAF-32 and CP190 are indeed found at the border elements of the *dhd* mini-domain. Second, this mini-domain does not feature internal PREs and border elements are only weakly if at all bound by Polycomb proteins, arguing against an autonomous recruitment of E(z). Finally, a deletion of the BEAF-32/CP190-bearing regulatory element in the *dhd^{l5}* mutant, or its displacement to an ectopic genomic location in rescue transgenes both abrogate H3K27me3 deposition. Consistent with such a model, data from Heurteau *et al.* show a modest reduction of H3K27me3 enrichment at the *dhd* mini-domain upon BEAF-32 depletion. Of note, BEAF-32 was also previously shown to facilitate H3K9me3

deposition at sites featuring multiple instances of the CGATA motif, analogous to those found at the *dhd* promoter [65]. Other studies found that ATCGAT motifs recognized by BEAF-32, also found at the *dhd* promoter, are more broadly enriched at the promoters of Lid-activated genes [66], which is the case of *dhd*. Thus, it is possible that a BEAF-32-mediated looping mechanism is responsible for H3K27me3 enrichment at the *dhd* mini-domain. However, our results also show that this mark is not strictly required to repress nor to activate *dhd* in adults, and that Lid, Sin3A, Snr1 and Mod(mdg4) activate *dhd* independently of it.

Scrutiny of *dhd* regulation further uncovered how its four regulators have convergent yet distinct roles. This was particularly intriguing for Lid and Sin3A, which can be found in a co-repressor complex [20], at odds with their positive impact on *dhd*. Indeed, their dual depletion in cultured cells causes the misregulation of hundreds of genes [19]. Interestingly, in that study, only 55 out of 849 affected genes were similarly impacted by individual and dual knockdowns, indicating that Lid and Sin3A functionally cooperate only at a minor subset of their common targets. This seems to be the case at the *dhd* locus, where individual KD of these factors caused an equally catastrophic collapse of transcriptional activity, suggesting a cooperative activity. Yet, Lid, but not Sin3A, acted as a negative regulator of H3K27me3 at the *dhd* locus, revealing at least partially independent functions. In contrast, Sin3A, but not Lid, controlled the stability of regulatory elements associated with this H3K27me3, not only at the *dhd* domain but also genome-wide.

Our results further show a critical role for Mod(mdg4) as a transcriptional activator. In cell lines, ChIP-seq experiments specifically mapping the insulating Mod(mdg4)^{67.2} isoform or total Mod(mdg4) showed that additional isoforms are recruited to DNA [24]. Isoforms other than the 67.2 were found in particular at gene

promoters in ovaries and female heads [4]. Such is the case at the *dhd* promoter, where total Mod(mdg4) is found but not the 67.2 isoform (FigS2-B). Non-insulating roles of Mod(mdg4) were previously discussed in the context of the Polycomb-repressed Bithorax complex where the close binding of Mod(mdg4) to *Abd-B* transcription start sites suggested a role in transcription activation [26]. While we cannot rule out indirect effects, these observations argue that an activating isoform of Mod(mdg4) operates directly at the *dhd* promoter. In agreement, Mod(mdg4) appears to be essential to activate *dhd* within its H3K27me3 mini-domain, seemingly by stabilizing the *dhd* promoter regulatory element, although its function is equally essential in the absence of heterochromatin marks in the *dhd* transgenic rescue construct.

The Snr1-containing Brahma complex is required for activation of target genes in *Drosophila in vivo*, notably during immune responses [60] and tissue regeneration [56]. In ovaries, while Snr1 has a global impact on nuclear integrity and architecture, previous immunostaining experiments interestingly showed that this factor is only expressed during a restricted time in early oogenesis [22]. This underlines the fact that *dhd* may be dynamically regulated during oogenesis, with different regulatory components intervening at particular times. Considering that *Snr1* KD causes a disruption of the *dhd* promoter-proximal regulatory element associated with H3K27me3, this would suggest that its associated DNA-binding transcription factors also intervene during a restricted time in oogenesis. A precise dissection of the timing of *dhd* transcription, and determining whether these factors target *dhd* directly and simultaneously, would be essential to understand the cascade of events leading to its massive expression.

The case of *dhd* indeed illustrates the complexity of understanding the chromatin landscape at cell type-specific genes, when the starting material is a complex tissue. In

this context, the Cut&Run analysis implemented in our study allowed us to reveal the co-occupancy of H3K27me3 nucleosomes and associated transcription factors. While this approach cannot identify the cell of origin of each individual DNA molecule, it can be used to make important deductions on the combinatorial co-occupancy on DNA of different chromatin components. This approach joins other recent methods comparable in their principle, namely the DNA methyl-transferase single-molecule footprinting (dSMF) method [67] and the low-salt antibody-targeted tagmentation (CUTAC) approach [68]. Together with single-cell methodologies, these approaches hold the potential to begin uncovering complex epigenomic regulation processes, such as that of *dhd*, that were until recently inaccessible.

Acknowledgements

We thank Dr. Kami Ahmad for sharing reagents and advice on Cut&Run. We are grateful to the TRiP projects and Bloomington Drosophila Stock Center for fly stocks. We acknowledge the contribution of SFR Biosciences (UAR3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) facilities PLATIM and Arthro-tools. Sequencing was performed by the GenomEast platform, a member of the 'France Génomique' consortium. We would like to thank Raphaëlle Dubruille and Francesca Palladino for critical reading of the manuscript. We also warmly thank them, as well as Lucas Waltzer and all members of the Loppin lab for their input throughout this project. This work was supported by an Association pour la Recherche sur le Cancer (ARC) Foundation grant (PJA20191209671) to BL. DTC was supported by a PhD fellowship from the Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation and from the ARC (ARCD0C42020020001727).

Competing Interests Statement:

The authors declare no competing interests.

Author Contributions

Conceptualization: Daniela Torres-Campana, Benjamin Loppin, Guillermo A. Orsi.

Funding acquisition: Benjamin Loppin, Guillermo A. Orsi.

Investigation: Daniela Torres-Campana, Béatrice Horard, Sandrine Denaud, Benjamin Loppin, Guillermo A. Orsi.

Data curation and formal analysis: Daniela Torres-Campana, Gérard Benoit, Guillermo A. Orsi.

Methodology: Daniela Torres-Campana, Guillermo A. Orsi

Supervision: Béatrice Horard, Benjamin Loppin, Guillermo A. Orsi.

Validation: Daniela Torres-Campana, Guillermo A. Orsi.

Writing – original draft: Daniela Torres-Campana, Guillermo A. Orsi.

Writing – review & editing: Daniela Torres-Campana, Béatrice Horard, Gérard Benoit, Benjamin Loppin and Guillermo A. Orsi.

Materials & Methods

Drosophila strains

Flies were raised at 25°C on standard medium. The following stocks were obtained from the Bloomington *Drosophila* Stock Center (simplified genotypes are given):

P{TRiP.HMS00849}attP2 (*mod(mdg4)* shRNA; #33907), *P{TRiP.HMS00363}attP2* (*Snr1* shRNA; #32372), *P{TRiP.GL00612}attP40* (*lid* shRNA; #36652), *P{TRiP.GLV21071}attP2* (*lid* shRNA; #35706), *P{TRiP.HMS00359}attP2* (*Sin3a* shRNA; #32368), *P{TRiP.HMS00066}attP2* (*E(z)* shRNA; #33659), *P{y[+t7.7]=CaryP}attP2* (Control line for TRiP RNAi lines; #B36303), *P{otu-GAL4::VP16.R}1*; *P{GAL4-nos.NGT}40*; *P{GAL4::VP16-nos.UTR}MVD1* (Maternal Triple Driver or “MTD-Gal4”; #31777), *P{GAL4::VP16-nos.UTR}MVD1* (“nos-Gal4”; #4937). Other stocks are: *w¹¹¹⁸*, *Df(1)J5/FM7c* (Salz et al., 1994), *P[Mst35Ba-EGFP]* (Manier et al., 2010), *pW8-dhd^{WT}* (Tirmarche et al., 2016). TRiP lines target all predicted isoforms of their respective target genes. “Control” in shRNA experiments refers to the offspring of the control line for TRiP lines crossed with the MTD-Gal4 line.

For the *pW8-dhd^{ΔDRE}* mutant, two fragments were amplified by PCR from *w¹¹¹⁸* genomic DNA using the primers ΔDRE-1-for/ ΔDRE-1-rev and ΔDRE-2-for/ ΔDRE-2-rev (Table S1). PCR products were assembled and cloned into the *pW8-dhd^{WT}* vector (Tirmarche et al., 2016) previously digested by KpnI and BamHI using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB, #E5520S). *dhd^{ΔDRE}* transgene was integrated in the *PBac{attP-3B}VK00031* platform (62E1) using PhiC31-mediated transformation (Bischof et al., 2007) and flies were generated by The Best Gene (TheBestGene.com).

Germline knock-down and fertility tests

To obtain *KD* females, virgin shRNA transgenic females were mass crossed with transgenic Gal4 males at 25°C and females of the desired genotype were recovered in the F1 progeny. All RNAi experiments were carried at 25°C. To measure fertility, virgin

females of different genotypes were mated to males in a 1:1 ratio and placed for 2 days at 25°C. They were then transferred to a new vial and allowed to lay eggs for 24 hours. Embryos were counted and then let to develop for at least 36 hours at 25°C. Unhatched embryos were counted to determine hatching rates.

Gene expression analysis by RT-QPCR

Total RNA was extracted from ovaries of 3-day-old females using the NucleoSpin RNA isolation kit (Macherey-Nagel), following the instructions of the manufacturer. 1µg of total RNA was reverse transcribed using the SuperScript II Reverse Transcriptase kit (Invitrogen) with oligo (dT) primers. RT-qPCR reactions were performed in duplicates as described previously (Torres-Campana et al., 2020). Primer sets used are provided in Table S1. Statistical tests were performed using GraphPad Prism version 9.2.0 for Mac OS (GraphPad Software).

Immunofluorescence and imaging

Early (0–30 min) embryos laid by females of the indicated genotypes were collected on agar plates. Embryos were dechorionated in bleach, fixed in a 1:1 heptane:methanol mixture and stored at -20°C. Embryos were washed three times (10 min each) with PBS 0.1%, Triton X-100 (PBS-T) and then incubated with primary antibodies in the same buffer on a wheel overnight at 4°C. They were then washed three times (20 min each) with PBS-T. Incubations with secondary antibodies were performed identically. Embryos were mounted in DAKO mounting medium containing DAPI.

Ovaries were dissected in PBS-T and fixed at room temperature in 4% formaldehyde in PBS for 25 minutes. Immunofluorescence was performed as for embryos. Ovaries were then mounted as described above.

Antibodies used are provided in Table S2. Images were acquired on an LSM 800 confocal microscope (Carl Zeiss). Images were processed with Zen imaging software (Carl Zeiss) and ImageJ software.

Western blotting

Ovaries from 30 females were collected and homogenized in lysis buffer (20mM HEPES pH7.9, 100mM KCl, 0.1mM EDTA, 0.1mM EGTA, 5% Glycerol, 0.05% Igepal and protease inhibitors (Roche)). Protein extracts were cleared by centrifugation and purified with Pierce GST Spin Purification Kit (ThermoFisher Scientific, #16106). Western analysis was performed using standard procedures and used antibodies and concentrations are presented in Table S2.

Ovarian RNA sequencing and analysis

Samples were processed as previously described (Torres-Campana et al., 2020).

Sequencing was completed on two biological replicates of the following genotypes:

mod(mdg4) KD (*MTD-Gal4>shRNA mod(mdg4)*), i.e

$P\{w[+mC] = otu-GAL4::VP16.R\}1, w[*]/y[1] sc[*] v[1]; P\{w[+mC] = GAL4-nos.NGT\}40/+;$

$P\{w[+mC] = GAL4::VP16-nos.UTR\}CG6325[MVD1]/P\{y[+t7.7] v[+t1.8] = TRiP.$

HMS00849} attP2

Snr1 KD (*MTD-Gal4>shRNA Snr1*), i.e

$P\{w[+mC] = otu-GAL4::VP16.R\}1, w[*]/y[1] sc[*] v[1]; P\{w[+mC] = GAL4-nos.NGT\}40/+;$

$P\{w [+mC] = GAL4::VP16-nos.UTR\}CG6325[MVD1]/P\{y[+t7.7] v[+t1.8] = TRiP.$

$HMS00363\}attP2$

Chromatin profiling by CUT&RUN

Cut&Run in *Drosophila* tissues was previously described [37]. Briefly, ovaries from 3-day-old flies were dissected in Wash+ Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.9 mM spermidine, 0.1% BSA with cOmplete protease inhibitor, Roche) and were bound to BioMag Plus Concanavalin-A-conjugated magnetic beads (ConA beads, Polysciences, Inc). Tissues were then permeabilized for 10min in db+ Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.9 mM spermidine, 2 mM EDTA, 0.1% BSA, 0.05% digitonin and protease inhibitors). Samples were then incubated with gentle rocking overnight at 4°C with primary antibody solution in db+ buffer (see Table S2 for antibody concentrations). Protein A fused to micrococcal nuclease (p-AMNase) was added in db+ buffer and samples were incubated with rotation at room temperature for 1 hour. Cleavage was done in WashCa+ buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.9 mM spermidine, 0.1% BSA, 2 mM CaCl₂ with and protease inhibitors) at 0° for 30 minutes. Digestion was stopped with addition of 2XSTOP Buffer (200mM NaCl, 20mM EDTA, 4mM EGTA, 62.5µg/mL RNaseA). Samples were incubated at 37°C for 30 min to digest RNA and release DNA fragments. Cleaved DNA was then recovered with Ampure XP beads (Beckman Coulter) immediately after protease treatment. Antibodies used for CUT&RUN are listed in Table S2. Retrieved DNA was used either for qPCR or for library preparation followed by deep sequencing. Sequencing libraries for each sample were synthesized using Diagenode MicroPlex Library Preparation kit according to supplier recommendations (version 2.02.15) and were sequenced on Illumina Hiseq 4000

sequencer as Paired-End 100 base reads following Illumina's instructions (GenomEast platform, IGBM, Strasbourg, France). Image analysis and base calling were performed using RTA 2.7.7 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover.

Cut&Run-qPCR

0,1 ng of retrieved DNA in Cut&Run were used as template in a real time quantitative PCR assay using SYBR Premix Ex Taq II (Tli RNaseH Plus) (Takara). All qPCR reactions were performed in duplicates using Bio-Rad CFX-96 Connect system with the following conditions: 95°C for 30s followed by 40 cycles of denaturation at 95°C for 5s and annealing and extension at 59°C for 30s. As a normalization control, we processed ovary samples from each studied genotype as for Cut&Run, except the antibody and pA-MNase incubation steps were omitted and instead we incubated tissue with 10U of Micrococcal Nuclease for 30min at 37°C (ThermoFisher Scientific, #88216). Fold change in histone mark enrichment was determined relative to this whole MNase control and relative to the *Sas10* gene, which was depleted in the histone marks tested in this study. Primer sets used are provided in Table S1.

Sequencing data processing

Paired-end reads were mapped to the release 6 of the *D. melanogaster* genome using Bowtie2 (v. 2.4.2). To compare samples with identical readcount for genome coverage quantifications, we employed Downsample SAM/BAM (Galaxy Version 2.18.2.1). To obtain short fragment datasets for DNA regulatory element identification, peak calling and visualization, we selected fragments shorter than 120 bp from SAM files. These

were typically a small minority of all fragments, as our Cut&Run datasets were largely dominated by nucleosome-sized fragments (150-250bp). We therefore could separately analyze modification-bearing nucleosome coverage (for which the complete Cut&Run dataset, “All fragments”, was a good approximation) or putative regulatory element coverage (<120bp fragments). For genomic track visualization, we used bamCoverage from deepTools 2.0 (Galaxy Version 3.3.2.0.0) to calculate read coverage per 25bp bin for all fragments or 10bp bin for short fragments, with paired-end extension. Peak calling was done on sorted short fragments (<120 bp) with MACS2 (v. 2.1.1.20160309) with the following parameters: -nomodel, -p-value =0.0001, -keep-dup=all and the rest by default. To establish a high-confidence short fragment peak list we retained peaks that were present in biological replicates from the control genotype. Genome browser views screenshots were produced with the IGV software, for Cut&Run we used a 25bp bin for all fragments and a 10bp bin for short fragments (<120bp). For the midpoint-plot of fragment sizes around short fragment peaks, the length of each fragment was plotted as a function of the distance from the fragment midpoint to the summit of the peak identified by MACS2. For signal quantification at the *dhd* locus , normalized read counts were counted within the domain (Coordinates [5,312,054-5,317,465]) or within border element peaks (5' element: [5,312,120-5,312,211], 3' element: [5,317,300-5,317,447])

Heatmaps were generated with RStudio (RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL) and the packages ‘gplots’ (v.3.1.1) and ‘plyr’ (Wickham, 2011).

Data Availability:

Original sequencing data from this publication have been deposited to the Gene Expression Omnibus with identifiers GSE174263 (RNA-seq) and GSE174250 (Cut&Run).

Additional sequencing data used in this study are available from GEO under the following accession numbers: GSE151981 (ATAC-seq), GSE37444 and GSE146993 (H3K27me3 ChIP-seq), GSE99027 (H3K9me3 ChIP-seq), GSE36393 (Mod(mdg4) ChIP-seq), GSE62904 (CP190, Beaf-32 and Dref ChIP-seq) and GSE24521 (Polycomb and Polyhomeotic ChIP-seq).

Motif scanning

Motif scanning on the *pW8-dhd^{WT}* transgene sequence was done with FIMO (v. 5.3.3) [69] using the flyreg v.2 motif database with default parameters.

Figure Legends

Figure 1. Mod(mdg4) and Snr1 are required for *dhd* expression.

A—Maternal Mod(mdg4) and Snr1 are required for protamine removal and sperm nuclear decompaction at fertilization. Top: Confocal images of pronuclear apposition in eggs from Control (*MTD>+*), *dhd^{l5}*, *mod(mdg4)* KD or *Snr1* KD females mated with transgenic *ProtA::GFP* males. The sperm nucleus in *dhd^{l5}*, *mod(mdg4)* KD and *Snr1* KD eggs retains ProtA::GFP (green) and has a needle-shape morphology. Bottom: zoom on the sperm nucleus. Scale bars: 5µm.

B— *dhd* is strongly downregulated in *mod(mdg4)* KD and *Snr1* KD ovaries. RNA-seq normalized reads per gene (in RPKM) are shown for *mod(mdg4)* KD vs Control (top) and *Snr1* KD vs Control (bottom). Genes downregulated (green) or upregulated (red) in KD ovaries are highlighted.

C— Genome Browser view of Control, *dhd^{l5}*, *lid* KD, *Sin3a* KD, *mod(mdg4)* KD and *Snr1* KD ovarian RNA-seq signal at the *dhd* region showing dramatic downregulation in all KD conditions. Note that the Control track is represented at two different scales to fit

the high read count for *dhd* (top track) or the low read count for its neighboring genes (bottom track).

D— The DHD protein is undetectable in KD ovaries. Western blot analysis using an anti-DHD antibody on ovary extracts of the indicated genotypes. Alpha-tubulin is used as a loading control.

Figure 2. A single Cut&Run experiment maps both histone modifications and their associated regulatory elements.

A— Schematic overview of the Cut&Run procedure for dissected *Drosophila* ovaries. After tissue permeabilization and antibody targeting, ProteinA-MNase cleaves nearby exposed DNA allowing the solubilization and retrieval of both nucleosomal particles carrying the targeted histone modification and DNA particles occupied by transcription factors in the immediate vicinity.

B—Cut&Run reveals nucleosomes and transcription factor binding sites. Mid-point plot of ovarian H3K27me3 Cut&Run data centered at peaks identified by MACS2 from short fragments (<120bp) in the same experiment. This plot represents all paired-end sequenced fragments as their middle point coordinate in the X-axis, and their size in the Y-axis, revealing a class of clustered short fragments (50-130bp) flanked by nucleosome-sized fragments (>140bp).

C— H3K27me3 Cut&Run at the bithorax complex (BX-C) in *Drosophila* ovaries reveals its regulatory architecture. Genome browser track displaying all Cut&Run fragments and <120 bp fragments separately. Multiple well-described Polycomb Response Elements (PRE) and insulators within the Bithorax complex detected as short fragment peaks are indicated (arrows).

D— Cut&Run re-discovers regulatory elements associated with H3K27me3 genome-wide. Upper panels: short fragment peaks read density heatmaps of ovarian H3K27me3 Cut&Run (all fragments and <120bp fragments), ATAC-seq (from S2 cells, [35]), CP190 ChIP-seq (from Kc cells, [38]) and Polycomb ChIP-seq (Pc, from S2 cells, [37]) plotted at ±1kb around peak summit. Data is sorted by the ratio of H3K27me3 Cut&Run total reads at the 3' versus 5' flanks to reveal short fragment peaks at the borders or within H3K27me3 domains (dashed lines). Lower panels: average profiles corresponding to the top heatmaps, distinguishing 5' border peaks, 3' border peaks and peaks embedded within domains. Cut&Run short fragment peaks are enriched for ATAC-seq signal as well as CP190 (particularly at border peaks) and Polycomb (particularly at middle peaks).

Figure 3. *dhd* is embedded in an H3K27me3/H3K9me3 mini-domain flanked by regulatory elements.

A— The *dhd* region features an H3K27me3 and H3K9me3 mini-domain. Genome browser snapshots showing the distribution of all fragments and <120 bp fragments in the *dhd* region, revealing that *dhd* lies within a ~5450bp heterochromatin mini-domain

flanked by border regulatory elements associated with both H3K27me3 and H3K9me3. Arrows indicate direction of transcription.

B-The *dhd* mini-domain is highly enriched in H3K27me3 relative to its size. Scatter plot of our 278 ovarian H3K27me3 domains identified in Cut&Run, representing their average read counts normalized to domain size in the Y-axis versus domain size in the X-axis. C- Effect of the KDs on H3K27me3 enrichment genome-wide. Average normalized counts of H3K27me3 Cut&Run (all fragments) in H3K27me3 domains (plotted as meta-domains and including ± 3 kb from domain borders) in Control (*MTD>+*), *lid* KD, *Sin3a* KD, *mod(mdg4)* KD, *Snr1* KD and *E(z)* KD (arrows).

D- *lid* KD, but not *Sin3a*, *Snr1* or *mod(mdg4)*, impacts H3K27me3 enrichment at the *dhd* mini-domain. Left: genome browser plots of normalized H3K27me3 Cut&Run signal (all fragments) at the *dhd* genomic region in Control and KD ovaries. Right: Quantification of normalized read counts for the same samples. Data in this figure is for one representative replicate: other replicates are shown in FigS5-A.

E- Weak impact of KDs on heterochromatic marks at *dhd*. H3K27me3 and H3K9me3 Cut&Run-qPCR in Control and KD ovaries using the *Sas10* gene as negative control and *Ubx* and *CG12239* as positive controls for H3K27me3 and H3K9me3 respectively. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability from a representative replicate. Data in this figure is for one representative replicate: other replicates are shown in FigS5-B. P values indicate one-way ANOVA with Dunnett's multiple comparisons test to a control (* P < 0.0001; n.s = not significant).

Figure 4. Sin3A, Snr1 and Mod(mdg4) control the regulatory architecture of the *dhd* H3K27me3 mini-domain

A- *Sin3a* and *Snr1* KD, but not *lid* or *mod(mdg4)*, impact H3K27me3-associated regulatory elements genome-wide. Left: H3K27me3 Cut&Run <120 bp fragments normalized counts in Control and KD ovaries, plotted at ± 1 kb around the summit of short fragment peaks. Right: Heatmaps displaying H3K27me3 Cut&Run short fragment peaks normalized read counts ± 1 kb around peak center in Control and KD ovaries.

B- Sin3A, Snr1 and Mod(mdg4), but not Lid, impact the organization of regulatory elements at the borders of the *dhd* H3K27me3 mini-domain. Left: genome browser plots of normalized Control H3K27me3 Cut&Run signal (all fragments, top) and of normalized signal from <120bp fragments retrieved in H3K27me3 Cut&Run in Control and KD ovaries. Right: Quantification of <120bp fragments normalized read counts for the same samples. 5' and 3' border elements are plotted separately. Data in this figure is for one representative replicate: other replicates are shown in FigS4-C.

C- Table recapitulating the effect of the different KDs on H3K27me3 and H3K27me3-associated regulatory elements (H3K27me3-RE) genome-wide and at the *dhd* locus. “ \approx ” indicates modest or no change, “ \nearrow ” indicates an increase and “ \searrow ” a decrease in average read counts compared to Control. “?” indicates inability to conclude.

Figure 5. The *dhd* promoter-proximal DRE motifs are required for its expression.

A- Schematic representation of the genotypes studied in this figure. Upper panels: genomic browser views recapitulating the Control distribution of H3K27me3 Cut&Run signal (all fragments and <120bp fragments) as well as RNA-seq signal from Figures 1 and 2 at the *dhd* locus and showing lack of signal at the transgene insertion locus in the absence of any transgenic construct. Middle panel: schematic representation of the genomic composition of *w¹¹¹⁸* (reference strain), mutant and rescue flies, indicating the status of the *dhd* locus and the composition of the rescue transgene. Dashed lines indicate the targeted region by primer couples (primers R and 1-5) used for RT- and Cut&Run-qPCR in panels B and D. Bottom panel: sequence of the *dhd* promoter at the endogenous location (left) and in the Δ DRE mutant transgene where the 14bp containing the DRE motifs were deleted.

B- The DRE motifs at the *dhd* promoter are necessary for its expression. RT-qPCR quantification of *dhd* mRNA levels in ovaries from wild-type *w¹¹¹⁸* flies, *dhd^{l5}* mutants or *dhd^{l5}* mutants carrying either a WT (*pW8-dhd^{WT}*) or a mutant (*pW8-dhd ^{Δ DRE}*) or a full domain (*pW8-dhd^{FD}*) transgene (measured using the R primers, normalized to *rp49* and relative to expression in *w¹¹¹⁸*). Data from biological duplicates analyzed in technical duplicates are presented as mean \pm SEM. P values indicate one-way ANOVA with Tukey's multiple comparisons test (* P < 0.0001; n.s = not significant).

C- The *dhd* heterochromatin domain is affected in *dhd*-containing transgenic constructs. Genome browser plots of normalized ovarian H3K27me3 and H3K9me3 Cut&Run signal at the *dhd* genomic region in the indicated genotypes. The H3K27me3 domain is abolished in all transgenic rescues. The H3K9me3 domain is partly affected in the *dhd^{l5}* mutant, being lost the *dhd*-proximal end but maintained in the *dhd*-distal end. Dashed lines and numbers (1 to 5) indicate the targeted region by primer couples used for qPCR in panel D.

D- A transgene containing the full *dhd* domain does not restore heterochromatin marks. H3K27me3 and H3K9me3 Cut&Run-qPCR in Control, *dhd^{l5}*, *dhd^{l5};;pW8-dhd^{WT}* and *dhd^{l5};;pW8-dhd^{FD}* ovaries. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability from a representative replicate. Data in this figure is for one representative replicate: other replicates are shown in FigS7.

Figure 6. Lid, Sin3A, Mod(mdg4) and Snr1 are necessary for *dhd* expression in the absence of its heterochromatin domain.

A-Schematic representation of the genomic composition of *w¹¹¹⁸* (reference strain) and mutant flies carrying a rescue transgene and shRNA constructs controlled by the female germline specific nanos-Gal4 driver, respectively inserted at the platforms *attP 3B-62E1* and *attP 2-68A4*.

B-The rescue transgene does not restore *dhd* expression in KD ovaries. RT-qPCR quantification of *dhd* mRNA levels in ovaries of the indicated genotypes (normalized to *rp49* and relative to expression in *w¹¹¹⁸* ovaries). Data from biological duplicates analyzed in technical duplicates are presented as mean \pm SEM. P value indicates one-way ANOVA with Tukey's multiple comparisons test (* P < 0.0001),

C-H3K27me3 is absent from the *dhd* rescue transgene. H3K27me3 Cut&Run-qPCR in the indicated genotypes. The *Sas10* gene was used as negative control and *Ubx* as positive control. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability from a representative replicate. Error bars show technical variability from a representative replicate. Data in this figure is for one representative replicate: other replicates are shown in FigS8.

D- *dhd* is not ectopically expressed in adult tissues in the absence of its heterochromatin domain. RT-qPCR quantification of *dhd* mRNA levels in dissected ovaries or corresponding female carcasses as well as testes or corresponding male carcasses, in all indicated genotypes (normalized to *rp49* and relative to expression in ovaries in *w¹¹¹⁸*). Data from biological duplicates analyzed in technical duplicates are presented as mean \pm SEM. P value indicates one-way ANOVA with Tukey's multiple comparisons test (* P < 0.0001).

REFERENCES

1. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011;21: 381–395. doi:10.1038/cr.2011.22
2. Hammond CM, Strømme CB, Huang H, Patel DJ, Groth A. Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol.* 2017;18: 141–158. doi:10.1038/nrm.2016.159
3. Tyagi M, Imam N, Verma K, Patel AK. Chromatin remodelers: We are the drivers!! *Nucleus.* 2016;7: 388–404. doi:10.1080/19491034.2016.1211217
4. Melnikova LS, Georgiev PG, Golovnin AK. The Functions and Mechanisms of Action of Insulators in the Genomes of Higher Eukaryotes. *Acta Naturae.* 2020;12: 15–33. doi:10.32607/actanaturae.11144
5. Torres-Campana D, Kimura S, Orsi GA, Horard B, Benoit G, Loppin B. The Lid/KDM5 histone demethylase complex activates a critical effector of the oocyte-to-zygote transition. Bosco G, editor. *PLoS Genet.* 2020;16: e1008543. doi:10.1371/journal.pgen.1008543
6. Spradling A. *The development of Drosophila melanogaster.* 1993.
7. Salz HK, Flickinger TW, Mittendorf E, Pellicena-Palle A, Petschek JP, Albrecht EB. The *Drosophila* maternal effect locus *deadhead* encodes a thioredoxin homolog required for female meiosis and early embryonic development. *Genetics.* 1994;136: 1075–1086. doi:10.1093/genetics/136.3.1075

8. Svensson MJ, Chen JD, Pirrotta V, Larsson J. The ThioredoxinT and deadhead gene pair encode testis- and ovary-specific thioredoxins in *Drosophila melanogaster*. *Chromosoma*. 2003;112: 133–143. doi:10.1007/s00412-003-0253-5
9. Tirmarche S, Kimura S, Dubruille R, Horard B, Loppin B. Unlocking sperm chromatin at fertilization requires a dedicated egg thioredoxin in *Drosophila*. *Nat Commun*. 2016;7: 13539. doi:10.1038/ncomms13539
10. Petrova B, Liu K, Tian C, Kitaoka M, Freinkman E, Yang J, et al. Dynamic redox balance directs the oocyte-to-embryo transition via developmentally controlled reactive cysteine changes. *Proc Natl Acad Sci USA*. 2018;115: E7978–E7986. doi:10.1073/pnas.1807918115
11. Freier R, Aragón E, Bagiński B, Pluta R, Martin-Malpartida P, Ruiz L, et al. Structures of the germline-specific Deadhead and thioredoxin T proteins from *Drosophila melanogaster* reveal unique features among thioredoxins. *IUCrJ*. 2021;8: 281–294. doi:10.1107/S2052252521000221
12. Emelyanov AV, Fyodorov DV. Thioredoxin-dependent disulfide bond reduction is required for protamine eviction from sperm chromatin. *Genes Dev*. 2016;30: 2651–2656. doi:10.1101/gad.290916.116
13. Dingwall AK, Beek SJ, McCallum CM, Tamkun JW, Kalpana GV, Goff SP, et al. The *Drosophila* *snr1* and *brm* proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *MBoC*. 1995;6: 777–791. doi:10.1091/mbc.6.7.777
14. Gerasimova TI, Gdula DA, Gerasimov DV, Simonova O, Corces VG. A *drosophila* protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell*. 1995;82: 587–597. doi:10.1016/0092-8674(95)90031-4
15. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*. 2017;6: e21856. doi:10.7554/eLife.21856
16. Hirose F, Yamaguchi M, Handa H, Inomata Y, Matsukage A. Novel 8-base pair sequence (*Drosophila* DNA replication-related element) and specific binding factor involved in the expression of *Drosophila* genes for DNA polymerase alpha and proliferating cell nuclear antigen. *J Biol Chem*. 1993;268: 2092–2099.
17. Spain MM, Caruso JA, Swaminathan A, Pile LA. *Drosophila* SIN3 isoforms interact with distinct proteins and have unique biological functions. *J Biol Chem*. 2010;285: 27457–27467. doi:10.1074/jbc.M110.130245
18. Das TK, Sangodkar J, Negre N, Narla G, Cagan RL. Sin3a acts through a multi-gene module to regulate invasion in *Drosophila* and human tumors. *Oncogene*. 2013;32: 3184–3197. doi:10.1038/onc.2012.326
19. Gajan A, Barnes VL, Liu M, Saha N, Pile LA. The histone demethylase dKDM5/LID interacts with the SIN3 histone deacetylase complex and shares functional

- similarities with SIN3. *Epigenetics & Chromatin*. 2016;9: 4. doi:10.1186/s13072-016-0053-9
20. Moshkin YM, Kan TW, Goodfellow H, Bezstarosti K, Maeda RK, Pilyugin M, et al. Histone Chaperones ASF1 and NAP1 Differentially Modulate Removal of Active Histone Marks by LID-RPD3 Complexes during NOTCH Silencing. *Molecular Cell*. 2009;35: 782–793. doi:10.1016/j.molcel.2009.07.020
 21. Ni J-Q, Zhou R, Czech B, Liu L-P, Holderbaum L, Yang-Zhou D, et al. A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nat Methods*. 2011;8: 405–407. doi:10.1038/nmeth.1592
 22. Zraly CB, Marendra DR, Nanchal R, Cavalli G, Muchardt C, Dingwall AK. SNR1 is an essential subunit in a subset of *Drosophila* brm complexes, targeting specific functions during development. *Dev Biol*. 2003;253: 291–308. doi:10.1016/s0012-1606(02)00011-8
 23. Dorn R, Reuter G, Loewendorf A. Transgene analysis proves mRNA trans-splicing at the complex *mod(mdg4)* locus in *Drosophila*. *Proc Natl Acad Sci U S A*. 2001;98: 9724–9729. doi:10.1073/pnas.151268698
 24. Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi JE, Corces VG. *Drosophila* CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res*. 2012;22: 2176–2187. doi:10.1101/gr.136788.111
 25. Büchner K, Roth P, Schotta G, Krauss V, Saumweber H, Reuter G, et al. Genetic and Molecular Complexity of the Position Effect Variegation Modifier *mod(mdg4)* in *Drosophila*. *Genetics*. 2000;155: 141–157. doi:10.1093/genetics/155.1.141
 26. Savitsky M, Kim M, Kravchuk O, Schwartz YB. Distinct Roles of Chromatin Insulator Proteins in Control of the *Drosophila* Bithorax Complex. *Genetics*. 2016;202: 601–617. doi:10.1534/genetics.115.179309
 27. Manier MK, Belote JM, Berben KS, Novikov D, Stuart WT, Pitnick S. Resolving mechanisms of competitive fertilization success in *Drosophila melanogaster*. *Science*. 2010;328: 354–357. doi:10.1126/science.1187096
 28. Henikoff JG, Belsky JA, Krassovsky K, MacAlpine DM, Henikoff S. Epigenome characterization at single base-pair resolution. *Proceedings of the National Academy of Sciences*. 2011;108: 18318–18323. doi:10.1073/pnas.1110731108
 29. Kasinathan S, Orsi GA, Zentner GE, Ahmad K, Henikoff S. High-resolution mapping of transcription factor binding sites on native chromatin. *Nat Methods*. 2014;11: 203–209. doi:10.1038/nmeth.2766
 30. Orsi GA, Kasinathan S, Hughes KT, Saminadin-Peter S, Henikoff S, Ahmad K. High-resolution mapping defines the cooperative architecture of Polycomb response elements. *Genome Research*. 2014;24: 809–820. doi:10.1101/gr.163642.113

31. Ramachandran S, Ahmad K, Henikoff S. Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates. *Molecular Cell*. 2017;68: 1038-1053.e4. doi:10.1016/j.molcel.2017.11.015
32. Cleard F, Wolle D, Taverner AM, Aoki T, Deshpande G, Andolfatto P, et al. Different Evolutionary Strategies To Conserve Chromatin Boundary Function in the Bithorax Complex. *Genetics*. 2017;205: 589–603. doi:10.1534/genetics.116.195586
33. Kyrchanova O, Mogila V, Wolle D, Deshpande G, Parshikov A, Cléard F, et al. Functional Dissection of the Blocking and Bypass Activities of the Fab-8 Boundary in the *Drosophila* Bithorax Complex. Sánchez-Herrero E, editor. *PLoS Genet*. 2016;12: e1006188. doi:10.1371/journal.pgen.1006188
34. Ahmad K, Spens AE. Separate Polycomb Response Elements control chromatin state and activation of the vestigial gene. Cavalli G, editor. *PLoS Genet*. 2019;15: e1007877. doi:10.1371/journal.pgen.1007877
35. Jain SU, Rashoff AQ, Krabbenhoft SD, Hoelper D, Do TJ, Gibson TJ, et al. H3 K27M and EZHIP Impede H3K27-Methylation Spreading by Inhibiting Allosterically Stimulated PRC2. *Mol Cell*. 2020;80: 726-735.e7. doi:10.1016/j.molcel.2020.09.028
36. Minoux M, Holwerda S, Vitobello A, Kitazawa T, Kohler H, Stadler MB, et al. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science*. 2017;355: eaal2913. doi:10.1126/science.aal2913
37. Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, Paro R. Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res*. 2011;21: 216–226. doi:10.1101/gr.114348.110
38. Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong C-T, et al. Widespread Rearrangement of 3D Chromatin Organization Underlies Polycomb-Mediated Stress-Induced Silencing. *Molecular Cell*. 2015;58: 216–231. doi:10.1016/j.molcel.2015.02.023
39. Bartkuhn M, Straub T, Herold M, Herrmann M, Rathke C, Saumweber H, et al. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J*. 2009;28: 877–888. doi:10.1038/emboj.2009.34
40. Colmenares SU, Swenson JM, Langley SA, Kennedy C, Costes SV, Karpen GH. *Drosophila* Histone Demethylase KDM4A Has Enzymatic and Non-enzymatic Roles in Controlling Heterochromatin Integrity. *Developmental Cell*. 2017;42: 156-169.e5. doi:10.1016/j.devcel.2017.06.014
41. Kassis JA, Kennison JA, Tamkun JW. Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics*. 2017;206: 1699–1725. doi:10.1534/genetics.115.185116
42. Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell*. 2017;171: 34–57. doi:10.1016/j.cell.2017.08.002

43. Brower-Toland B, Riddle NC, Jiang H, Huisinga KL, Elgin SCR. Multiple SET methyltransferases are required to maintain normal heterochromatin domains in the genome of *Drosophila melanogaster*. *Genetics*. 2009;181: 1303–1319. doi:10.1534/genetics.108.100271
44. Ninova M, Fejes Tóth K, Aravin AA. The control of gene expression and cell identity by H3K9 trimethylation. *Development*. 2019;146: dev181180. doi:10.1242/dev.181180
45. Smolko AE, Shapiro-Kulnane L, Salz HK. The H3K9 methyltransferase SETDB1 maintains female identity in *Drosophila* germ cells. *Nat Commun*. 2018;9: 4155. doi:10.1038/s41467-018-06697-x
46. Kadoch C, Williams RT, Calarco JP, Miller EL, Weber CM, Braun SMG, et al. Dynamics of BAF–Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat Genet*. 2017;49: 213–222. doi:10.1038/ng.3734
47. De S, Gehred ND, Fujioka M, Chan FW, Jaynes JB, Kassis JA. Defining the Boundaries of Polycomb Domains in *Drosophila*. *Genetics*. 2020;216: 689–700. doi:10.1534/genetics.120.303642
48. Iovino N, Ciabrelli F, Cavalli G. PRC2 controls *Drosophila* oocyte cell fate by repressing cell cycle genes. *Dev Cell*. 2013;26: 431–439. doi:10.1016/j.devcel.2013.06.021
49. Phillips MD, Shearn A. Mutations in polycombetic, a *Drosophila* polycomb-group gene, cause a wide range of maternal and zygotic phenotypes. *Genetics*. 1990;125: 91–101.
50. DeLuca SZ, Ghildiyal M, Pang L-Y, Spradling AC. Differentiating *Drosophila* female germ cells initiate Polycomb silencing by regulating PRC2-interacting proteins. *eLife*. 2020;9: e56922. doi:10.7554/eLife.56922
51. Bergman CM, Carlson JW, Celniker SE. *Drosophila* DNase I footprint database: a systematic genome annotation of transcription factor binding sites in the fruitfly, *Drosophila melanogaster*. *Bioinformatics*. 2005;21: 1747–1749. doi:10.1093/bioinformatics/bti173
52. Daniel Avery Pollard. *Drosophila* Sequence Specific Transcription Factor Binding Site Matrices. Available: <http://www.danielpollard.com/matrices.html>
53. Zhao K, Hart CM, Laemmli UK. Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell*. 1995;81: 879–889. doi:10.1016/0092-8674(95)90008-X
54. Roy S, Gilbert MK, Hart CM. Characterization of *BEAF* Mutations Isolated by Homologous Recombination in *Drosophila*. *Genetics*. 2007;176: 801–813. doi:10.1534/genetics.106.068056

55. Park J-S, Choi Y-J, Thao DTP, Kim Y-S, Yamaguchi M, Yoo M-A. DREF is involved in the steroidogenesis via regulation of shadow gene. *Am J Cancer Res.* 2012;2: 714–725.
56. Tian Y, Smith-Bolton RK. Regulation of growth and cell fate during tissue regeneration by the two SWI/SNF chromatin-remodeling complexes of *Drosophila*. Duronio R, editor. *Genetics.* 2021;217: 1–16. doi:10.1093/genetics/iyaa028
57. Liu X, Secombe J. The Histone Demethylase KDM5 Activates Gene Expression by Recognizing Chromatin Context through Its PHD Reader Motif. *Cell Reports.* 2015;13: 2219–2231. doi:10.1016/j.celrep.2015.11.007
58. Saha N, Liu M, Gajan A, Pile LA. Genome-wide studies reveal novel and distinct biological pathways regulated by SIN3 isoforms. *BMC Genomics.* 2016;17: 111. doi:10.1186/s12864-016-2428-5
59. Heurteau A, Perrois C, Depierre D, Fosseprez O, Humbert J, Schaak S, et al. Insulator-based loops mediate the spreading of H3K27me3 over distant microdomains repressing euchromatin genes. *Genome Biol.* 2020;21: 193. doi:10.1186/s13059-020-02106-z
60. Gurudatta B, Yang J, Van Bortle K, Donlin-Asp P, Corces V. Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell Cycle.* 2013;12: 1605–1615. doi:10.4161/cc.24742
61. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell.* 2006;125: 315–326. doi:10.1016/j.cell.2006.02.041
62. Akmammedov A, Geigges M, Paro R. Bivalency in *Drosophila* embryos is associated with strong inducibility of Polycomb target genes. *Fly (Austin).* 2019;13: 42–50. doi:10.1080/19336934.2019.1619438
63. Schertel C, Albarca M, Rockel-Bauer C, Kelley NW, Bischof J, Hens K, et al. A large-scale, in vivo transcription factor screen defines bivalent chromatin as a key property of regulatory factors mediating *Drosophila* wing development. *Genome Res.* 2015;25: 514–523. doi:10.1101/gr.181305.114
64. De S, Mitra A, Cheng Y, Pfeifer K, Kassis JA. Formation of a Polycomb-Domain in the Absence of Strong Polycomb Response Elements. Bantignies F, editor. *PLoS Genet.* 2016;12: e1006200. doi:10.1371/journal.pgen.1006200
65. Emberly E, Blattes R, Schuettengruber B, Hennion M, Jiang N, Hart CM, et al. BEAF Regulates Cell-Cycle Genes through the Controlled Deposition of H3K9 Methylation Marks into Its Conserved Dual-Core Binding Sites. Misteli T, editor. *PLoS Biology.* 2008;6: e327. doi:10.1371/journal.pbio.0060327
66. Zamurrad S, Hatch HAM, Drelon C, Belalcazar HM, Secombe J. A *Drosophila* Model of Intellectual Disability Caused by Mutations in the Histone Demethylase KDM5. *Cell Rep.* 2018;22: 2359–2369. doi:10.1016/j.celrep.2018.02.018

67. Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. *Molecular Cell*. 2017;67: 411-422.e4. doi:10.1016/j.molcel.2017.06.027
68. Henikoff S, Henikoff JG, Kaya-Okur HS, Ahmad K. Efficient chromatin accessibility mapping *in situ* by nucleosome-tethered tagmentation. *Genomics*; 2020 Apr. doi:10.1101/2020.04.15.043083
69. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics*. 2011;27: 1017–1018. doi:10.1093/bioinformatics/btr064

Supporting Information Captions:

Figure S1. *mod(mdg4)* KD and *Snr1* KD downregulate *dhd*

A-*mod(mdg4)* KD and *Snr1* KD are efficient in the female germline. Left: RT-qPCR quantification of *mod(mdg4)* (top) or *Snr1* (bottom) mRNA levels in Control and KD ovaries (normalized to *rp49* and relative to expression in Control ovaries). Data from biological duplicates analyzed in technical duplicates are presented as mean \pm SEM. **Right:** Quantification of *mod(mdg4)* (top) and *Snr1*(bottom) counts in RNA-seq data from Fig1B. Both duplicates are shown.

B-*lid*, *Sin3a*, *Snr1* and *mod(mdg4)* KDs downregulate *dhd* but do not significantly affect its neighboring genes. Quantification of counts in RNA-seq data for *dhd* and its neighboring genes in Control, *lid*, *Sin3a*, *Snr1* and *mod(mdg4)* KD show that low-expressing genes in the *dhd* region are not or only modestly impacted by the KDs. Both duplicates are shown.

C-Limited overlap in the effects of *mod(mdg4)* and *Snr1* KDs. Hierarchical clustering of sample distance heatmap of RNA-seq samples.

D-Principal component analysis for RNA-seq samples.

E- *lid*, *Sin3a*, *mod(mdg4)* and *Snr1* KD severely downregulate *dhd* expression. RT-qPCR quantification of *dhd* mRNA levels in ovaries of indicated genotypes (normalized to *rp49* and relative to expression in Control ovaries). Two different shRNA constructs (val21 and val22) against *lid* were tested. Data from biological duplicates analyzed in technical duplicates are presented as mean \pm SEM.

Figure S2. Cut&Run is consistent with ChIP-seq data.

A—Histone modification profiles at the *dhd* region in cultured embryonic cells and ovaries. ChIP-seq data showing the active mark H3K4me3 (yellow) [5], and the repressive marks H3K9me3 (green) [40,45] and H3K27me3 (blue) (accession number GSE146993, [50]).

B—Short fragment peaks align with known regulatory elements. Genome browser views of the bithorax complex (BX-C) (left) and the *dhd* region (right). Display of H3K27me3 Cut&Run (from Control ovaries, all fragments and <120bp fragments), ATAC-seq (from S2 cells, [35]), CP190 ChIP-seq (from Kc cells, [38]), Mod(*mdg4*) (all isoforms) and Mod(*mdg4*)67.2 isoform ChIP-seq (from Kc cells, [24]) Polycomb (Pc) and Polyhomeotic (Ph) ChIP-seq (from S2 cells, [37]). Cut&Run short fragments largely overlap with peaks from the other tracks displayed.

Figure S3. Cut&Run maps H3K9me3 in ovaries.

A- Cut&Run shows the expected enrichment of H3K9me3 at pericentromeric heterochromatin. Genome browser views of H3K9me3 ChIP-seq [45] and Cut&Run signal in all chromosomes.

B- Cut&Run detects a previously identified H3K9me3 peak over a testis-specific TSS [45]. Genome browser view of *phf7* and neighboring genes. Blue arrow indicates testis-specific TSS and magenta arrow indicates ovary-specific TSS.

Figure S4. Whole-ovary experiments yield signal from both somatic follicle cells and germline cells.

A-*E(z)* KD does not severely affect *dhd* expression. RT-qPCR quantification of *dhd* mRNA levels in Control and *E(z)* KD ovaries (normalized to *rp49* and relative to expression in Control ovaries). Data from biological duplicates analyzed in technical duplicates are presented as mean ± SEM.

B- *E(z)* KD and *Snr1* KD affect H3K27me3 levels in nurse cells. Confocal images of representative egg chambers in Control, *E(z)* KD, *lid* KD, *Sin3a* KD, *mod(mdg4)* KD and *Snr1* KD. In control ovaries, H3K27me3 staining marks somatic follicle cell nuclei, the karyosome and germline nurse cell nuclei. In *E(z)* KD ovaries the karyosome and nurse cells loose staining of the histone mark but follicle cells are marked normally. No notable change is observed in *lid* KD, *Sin3a* KD or *mod(mdg4)* KD while in *Snr1* KD nurse cells staining is less intense. Scale bar 10µm.

C- Cut&Run in whole ovaries captures signal from both somatic and germline cells. Genome browser views of H3K27me3 Cut&Run signal in Control and *E(z)* KD ovaries and H3K27me3 ChIP-seq from FACS sorted nurse cells and somatic follicle cells [50]. Upper panels show representative loci enriched for the mark solely in nurse cells (germline) and absent in *E(z)* KD ovaries. Lower panels show H3K27me3 domains where the signal comes almost exclusively from follicle cells and is not significantly affected in the germline *E(z)* KD.

Figure S5. Cut&Run is reproducible among replicates.

A—H3K27me3 Cut&Run signal at the *dhd* locus from Control and KD ovaries. Left: Dotplot showing normalized read counts of H3K27me3 Cut&Run at the *dhd* domain from independent biological triplicates of the indicated genotypes (duplicates for *E(z)* KD).

B— Cut&Run qPCR yields reproducible data among replicates. Biological replicates from H3K27me3 and H3K9me3 Cut&Run-qPCR in Control and KD ovaries shown in Fig3-E. The *Sas10* gene was used as negative control and *Ubx* and *CG12239* as positive controls for H3K27me3 and H3K9me3 respectively. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability.

C— *Sin3a* KD, *Snr1* KD and *mod(mdg4)* KD affect the stability of the H3K27me3-associated regulatory elements at the *dhd* mini-domain. Dotplot showing normalized read counts of H3K27me3 Cut&Run <120bp fragments at *dhd* regulatory elements from independent biological triplicates of the indicated genotypes (duplicates for *E(z)* KD). 5' and 3' border elements are plotted separately.

Figure S6. Dref and Beaf-32 are found at *dhd* regulatory elements.

Genome browser views of ovarian H3K27me3 Cut&Run (all fragments and <120bp fragments) and Dref and Beaf-32 CHIP-seq (from Kc cells, [38]). <120 bp fragment peaks at the *dhd* domain borders align with DREF and Beaf-32 peaks.

Figure S7. Cut&Run qPCR generates reproducible data from transgenes

Biological replicates of H3K27me3 and H3K9me3 Cut&Run-qPCR in Control, *dhd^{Δ5}*, *dhd^{Δ5};pW8-dhd^{WT}* and *dhd^{Δ5};pW8-dhd^{FD}* ovaries shown in Fig5-D. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability.

Figure S8. In *lid* KD, *Sin3a* KD, *mod(mdg4)* KD and *Snr1* KD rescue flies, H3K27me3 is absent from the *dhd* rescue transgene

A— *lid* KD, *Sin3a* KD, *mod(mdg4)* KD and *Snr1* KD are efficient in the female germline of rescue flies. From left to right: RT-qPCR quantification of *lid*, *Sin3a*, *mod(mdg4)* and *Snr1* mRNA levels in ovaries of the indicated genotypes (normalized to *rp49* and relative to expression in *w¹¹¹⁸* ovaries). Data from biological duplicates analyzed in technical duplicates are presented as mean ± SEM.

B— The *dhd* rescue transgene does not restore H3K27me3 in KD flies. Biological replicates of H3K27me3 Cut&Run-qPCR in the indicated genotypes shown in Fig6-C. The *Sas10* gene was used as negative control and *Ubx* as positive control. Fold enrichment was calculated relative to *Sas10*. Error bars show technical variability.

Table S1. List of primers used in this paper.

Table S2. List of antibodies used in this paper.

Figure 1

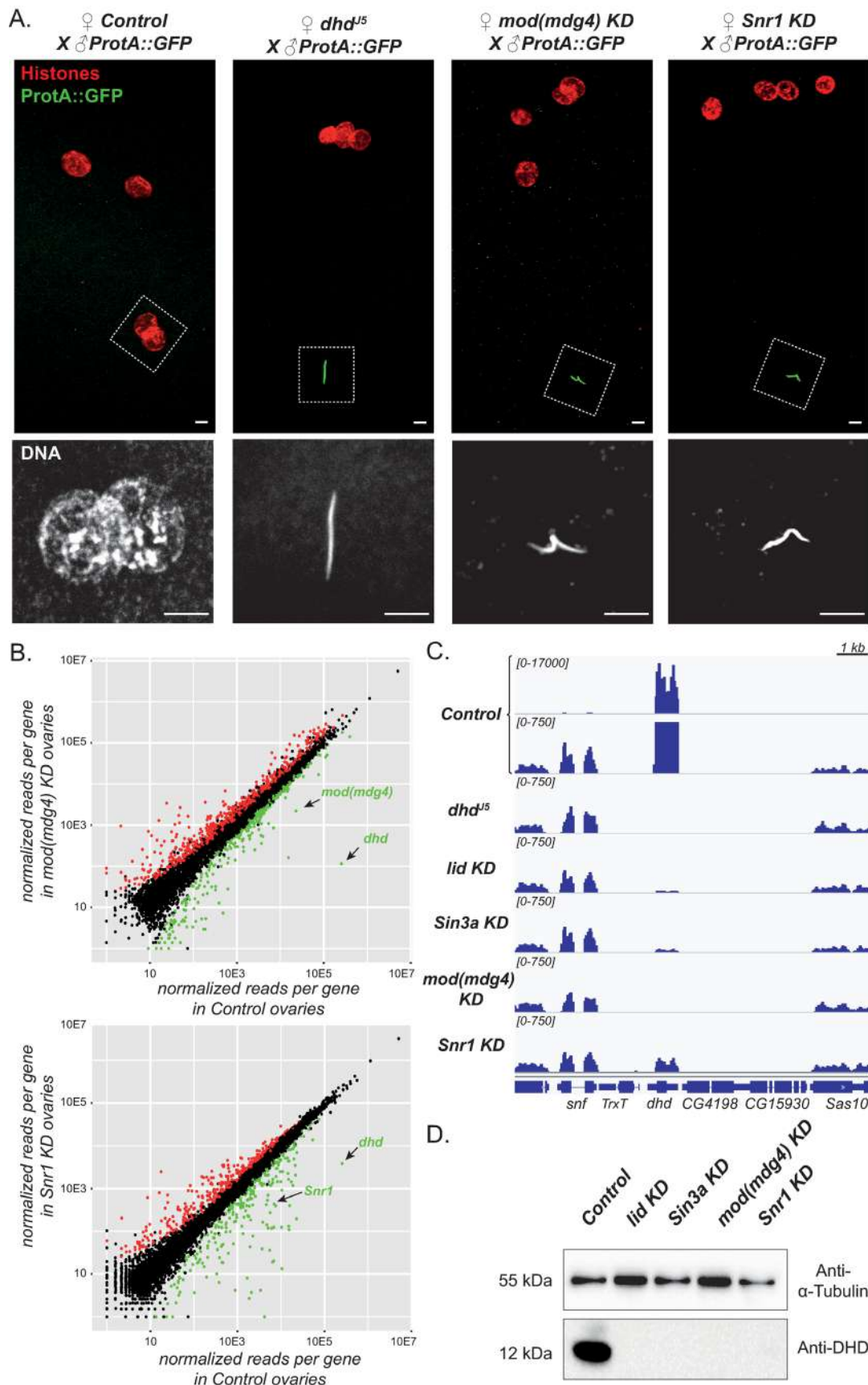


Figure 2

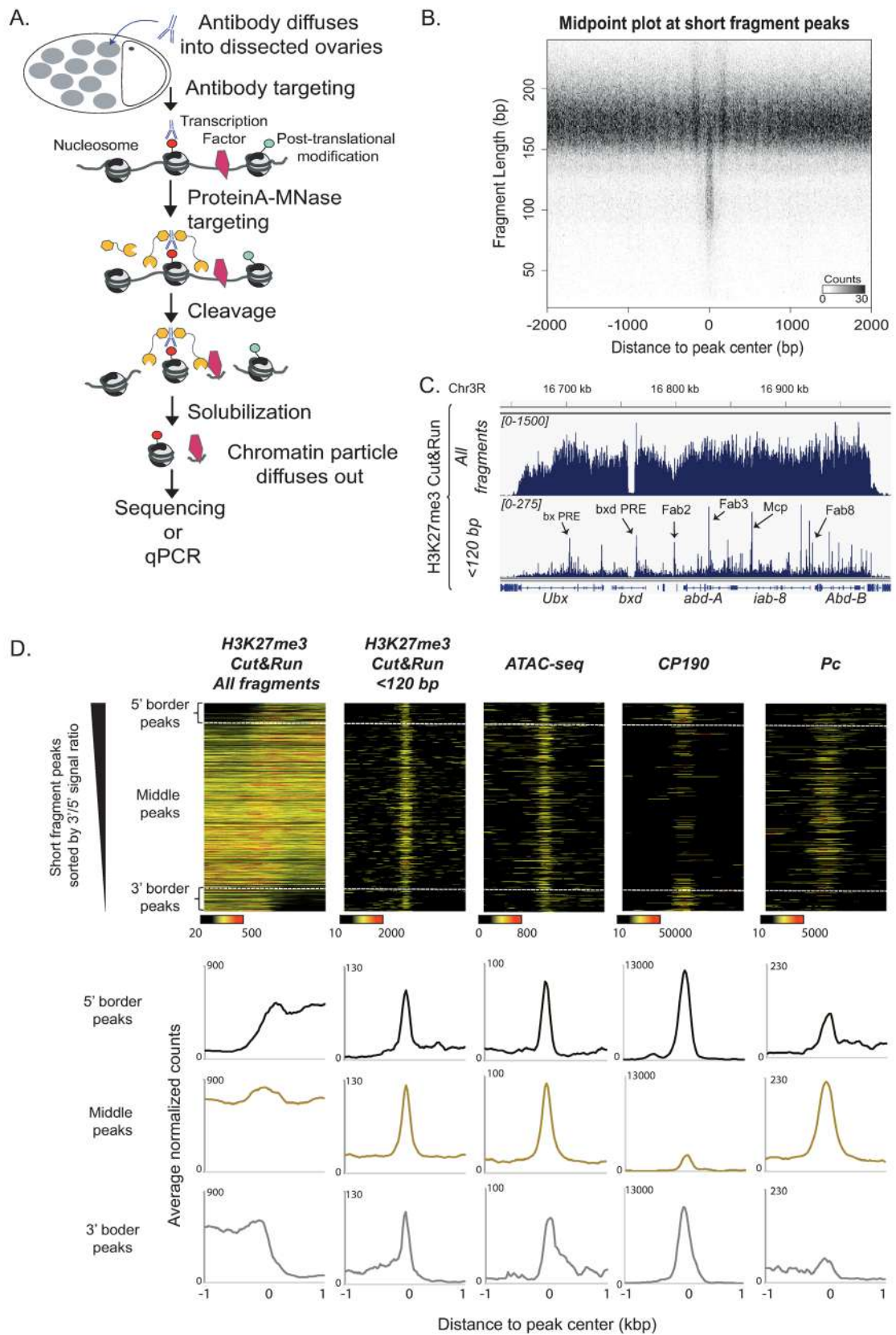


Figure 3

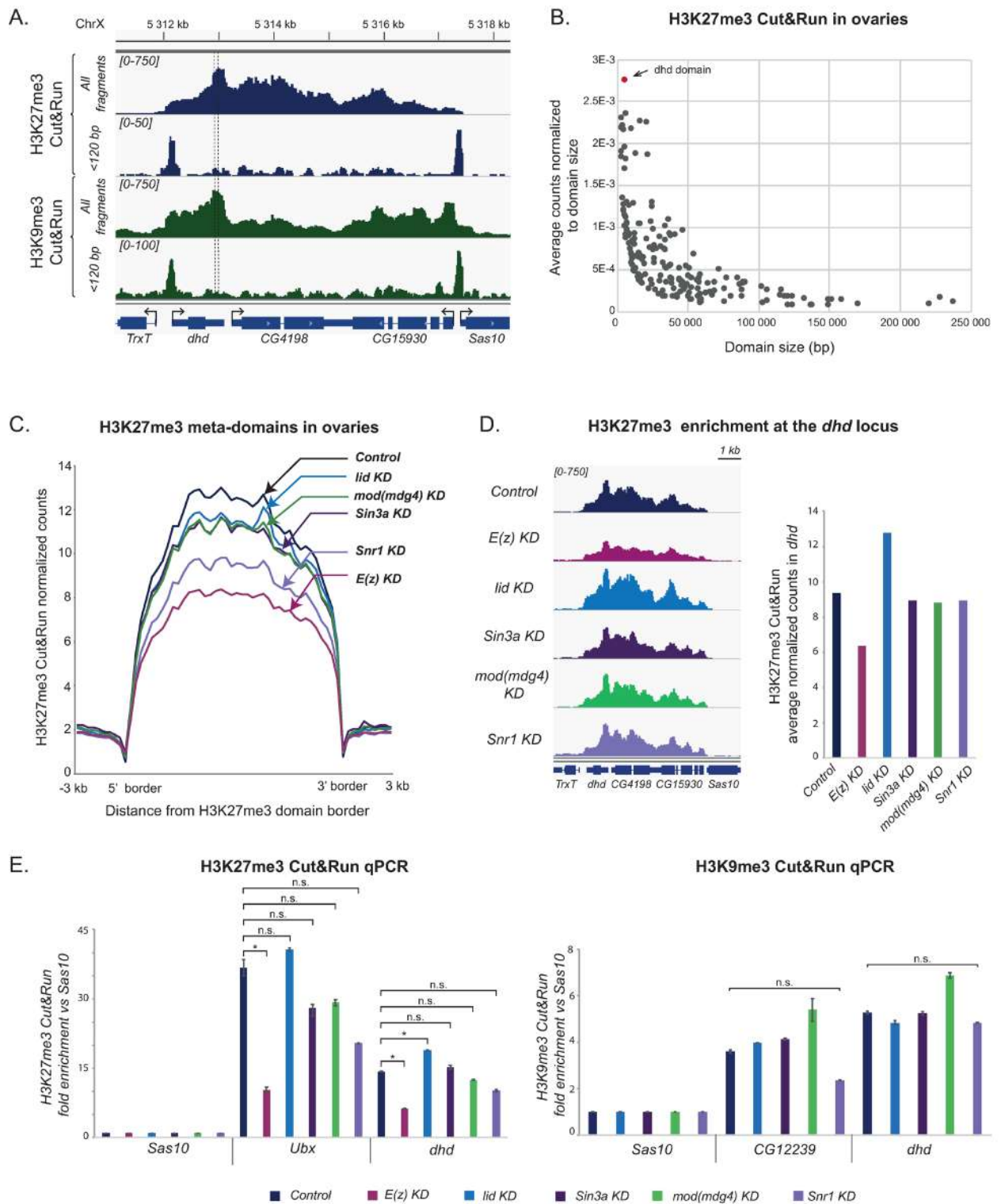


Figure 4

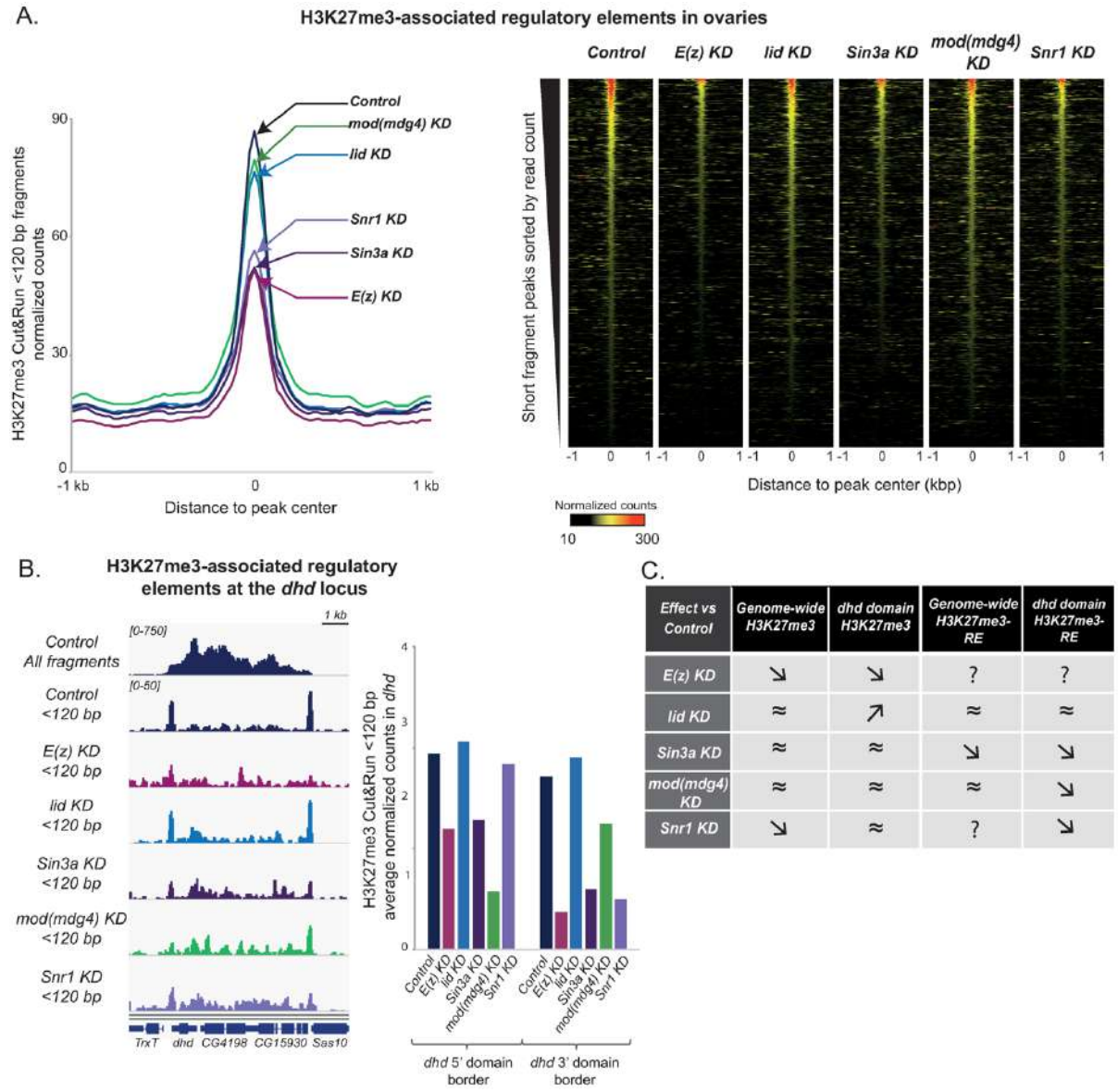


Figure 5

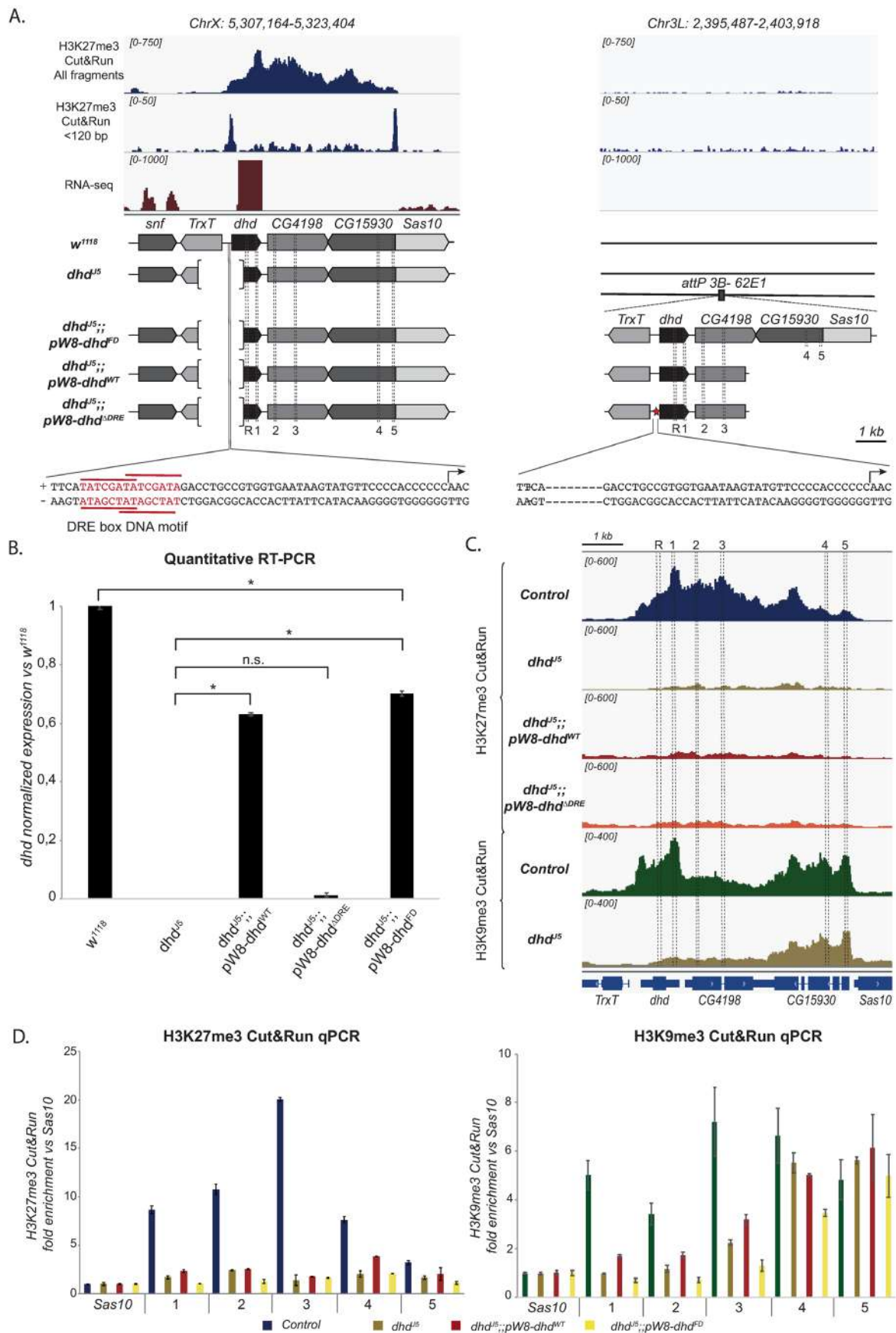


Figure 6

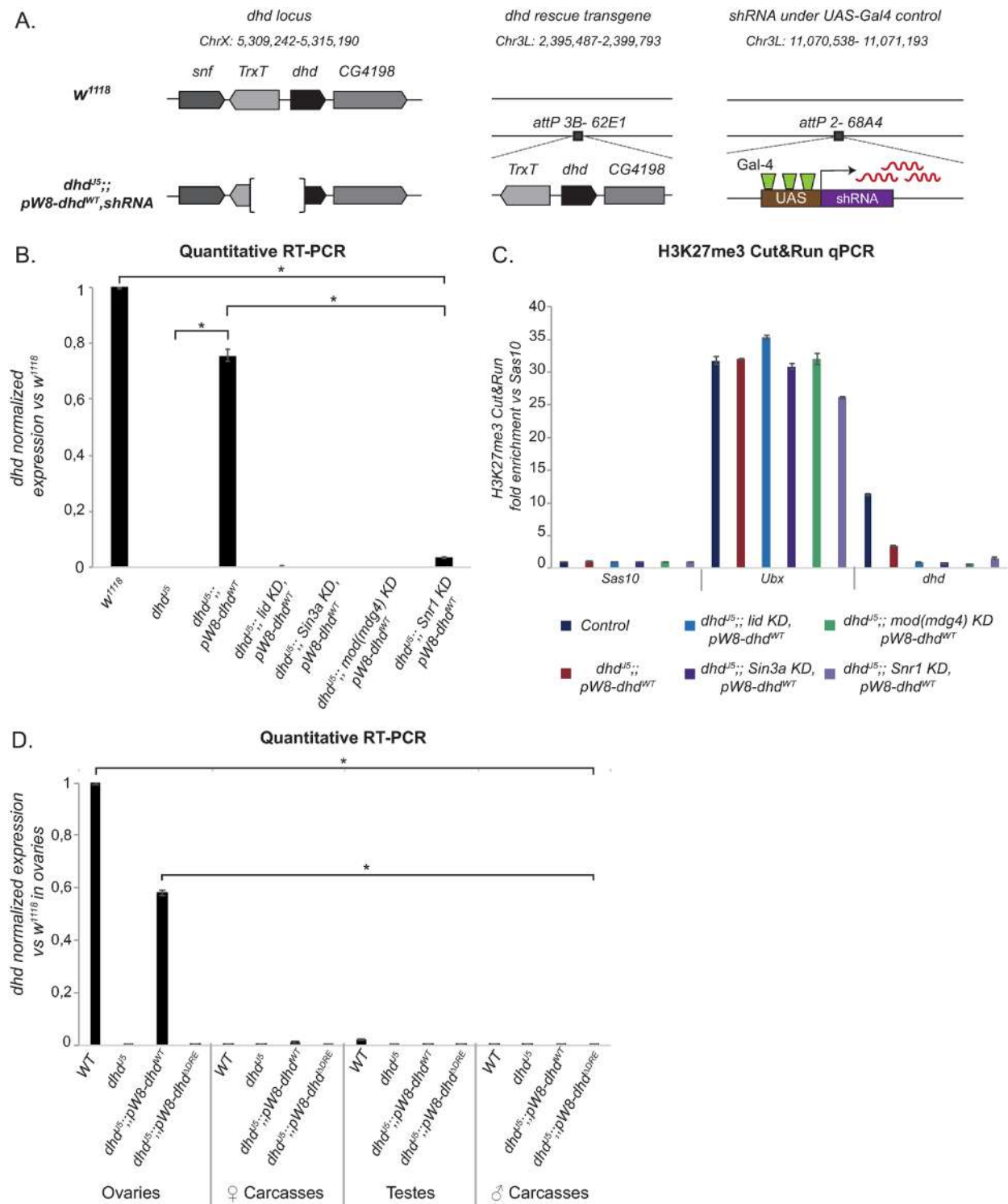


Figure S1

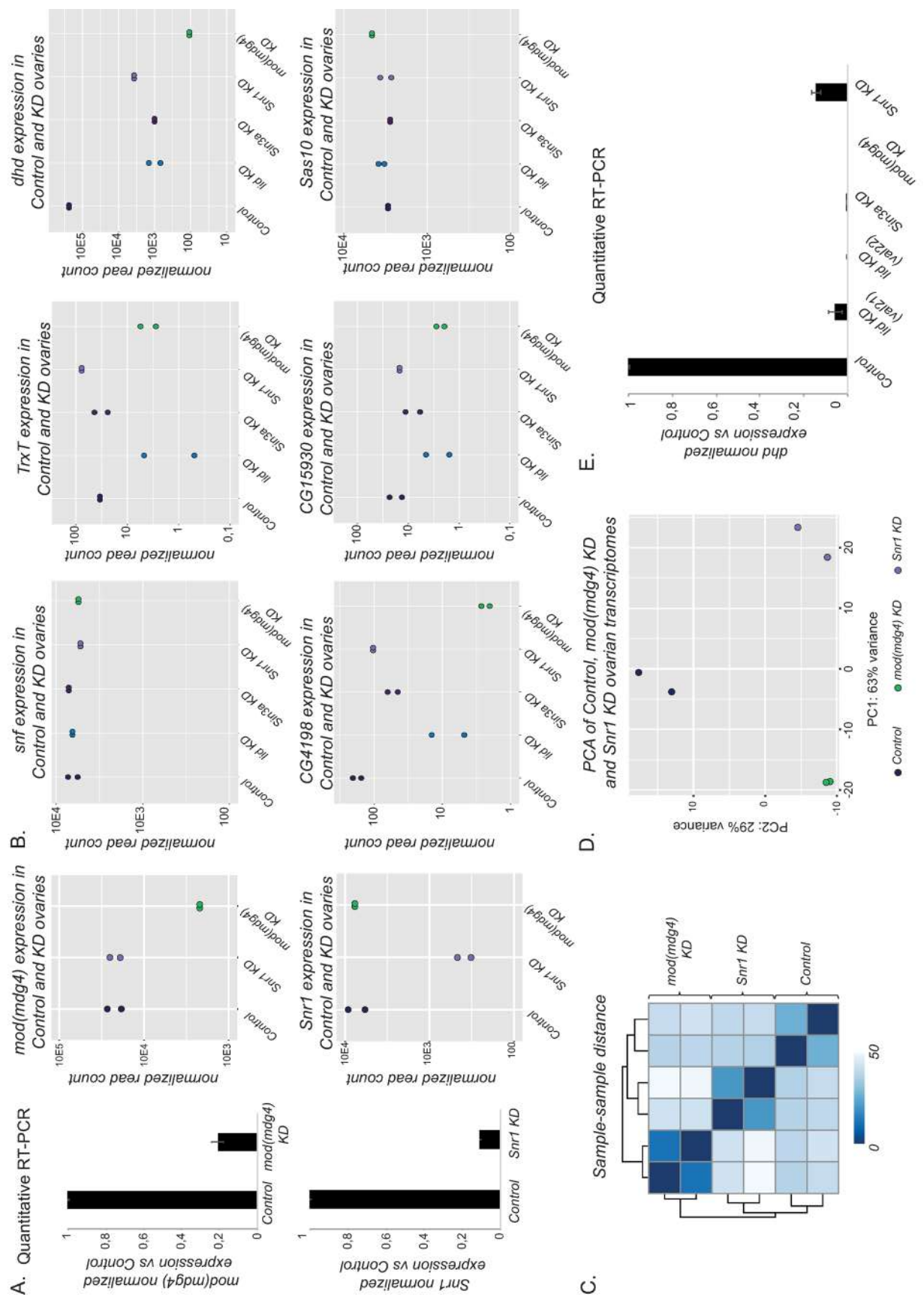
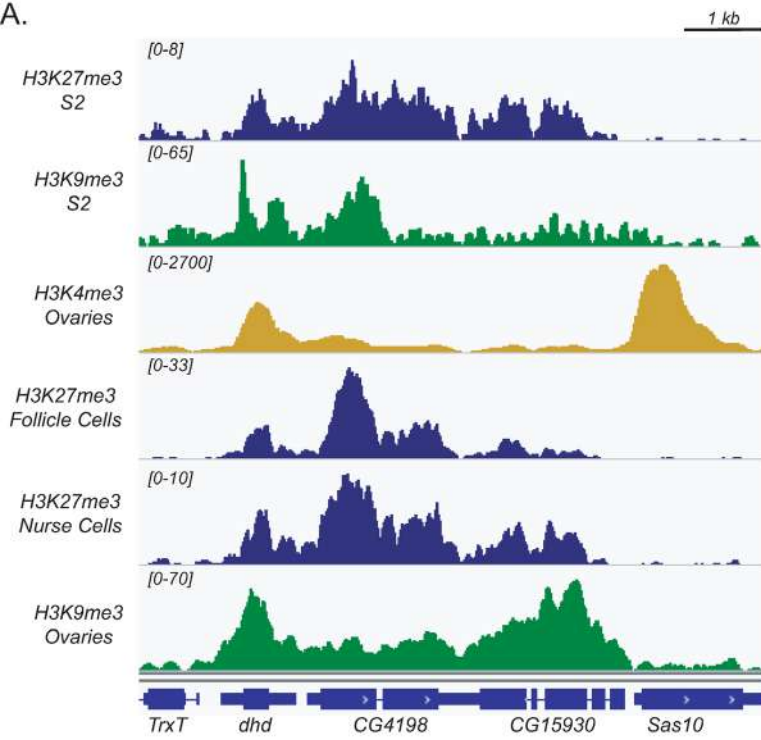


Figure S2

A.



B.

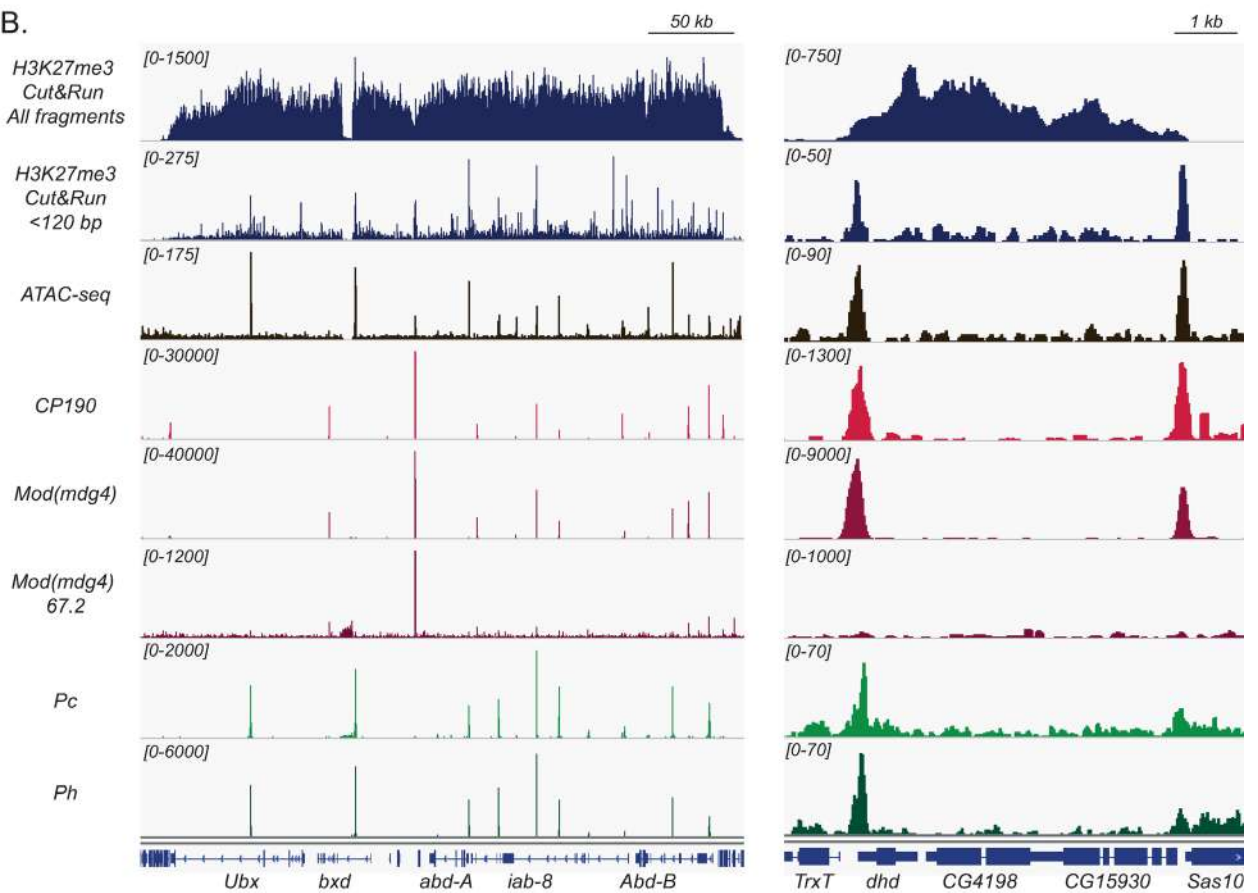


Figure S3

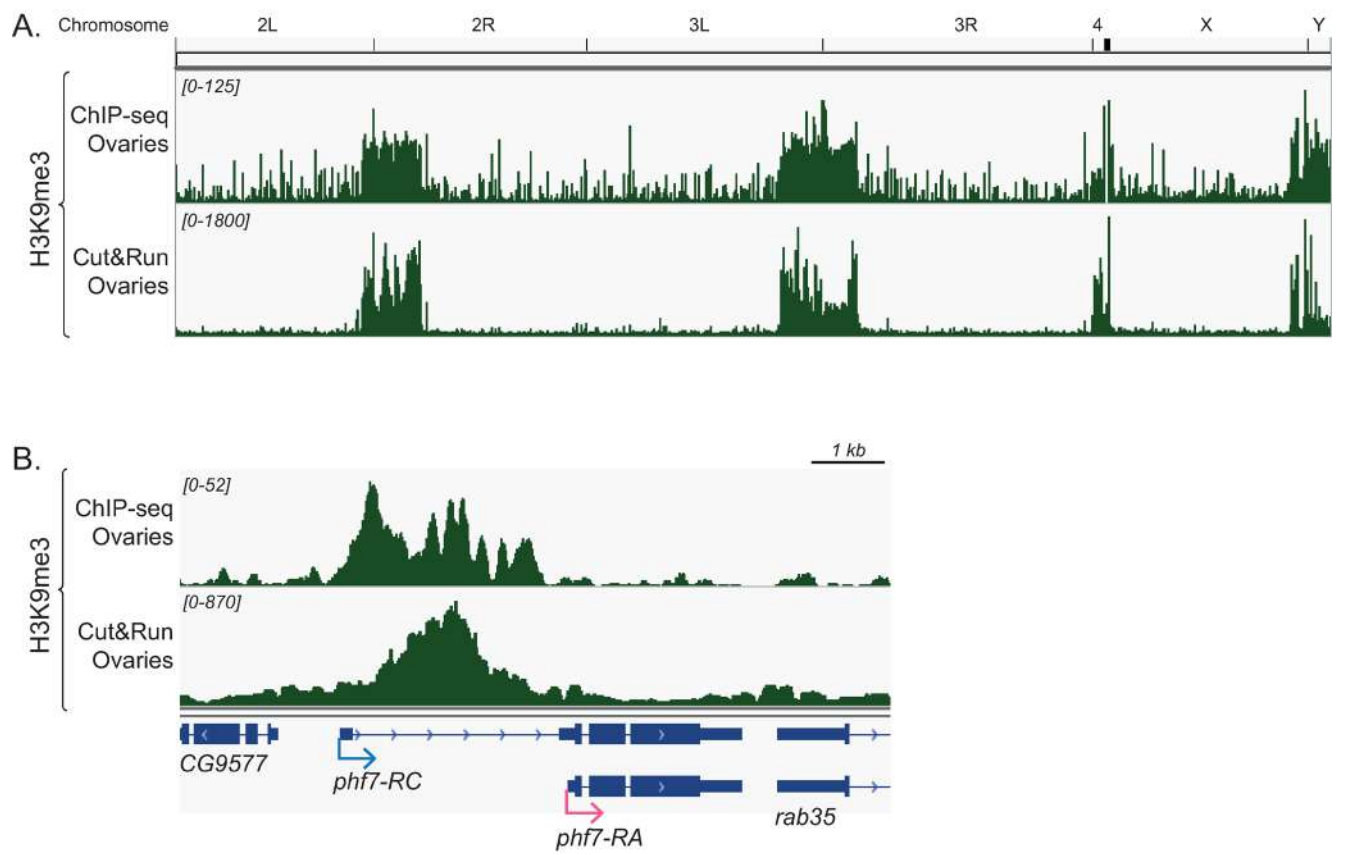


Figure S4

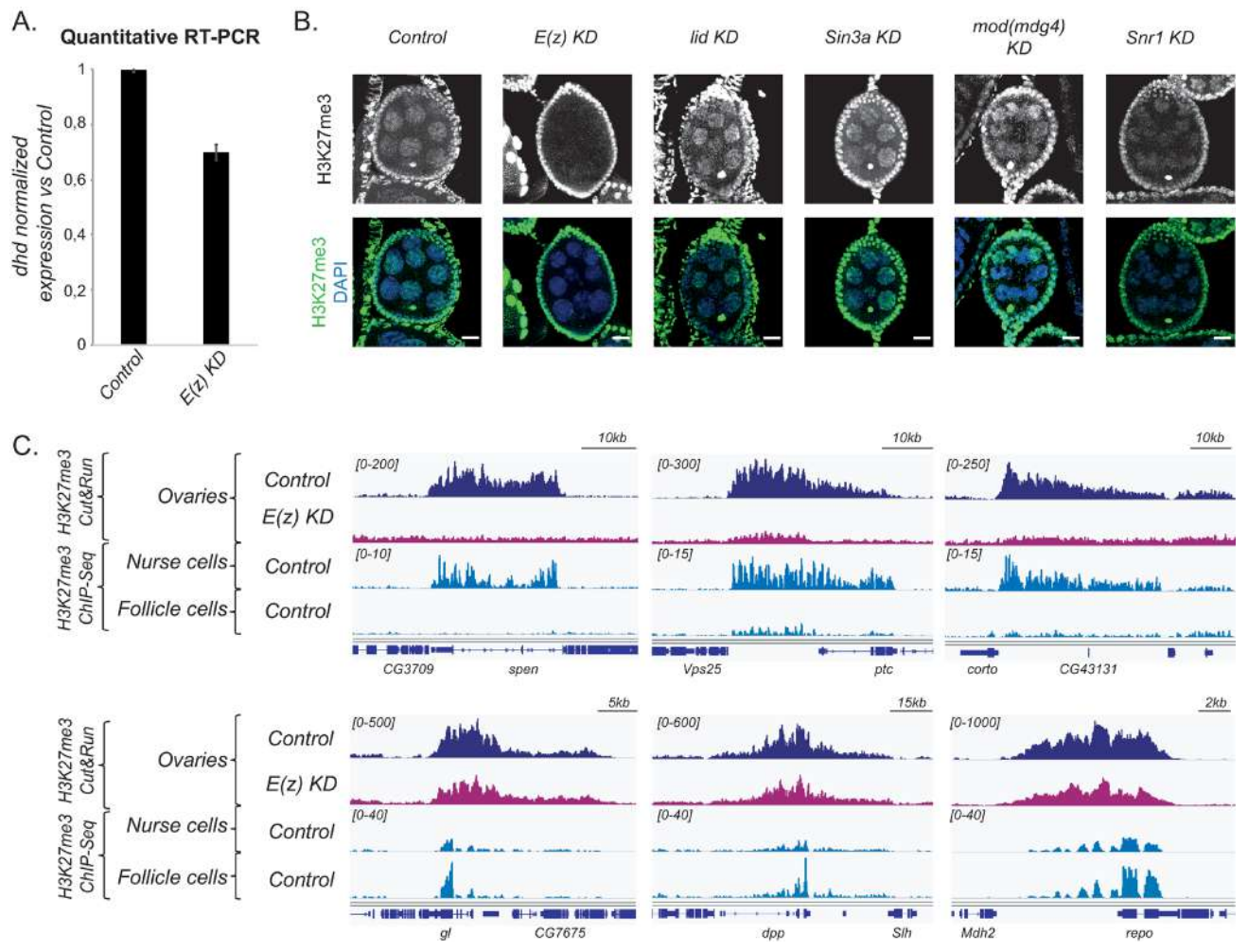
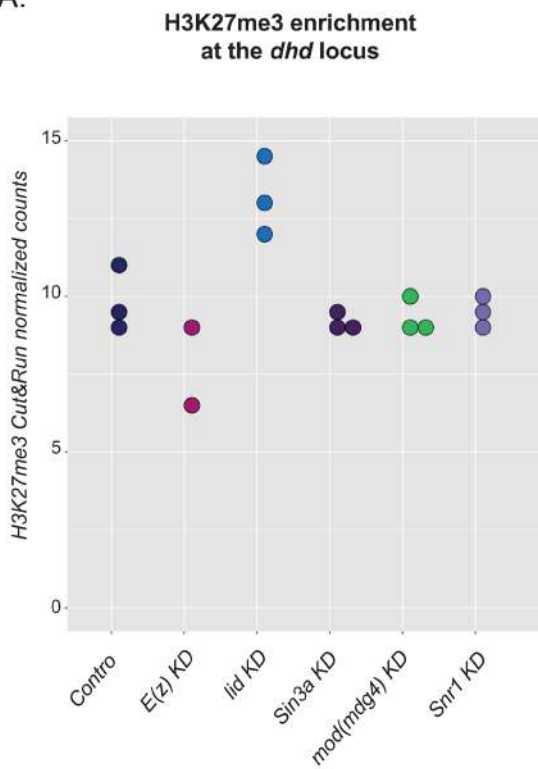
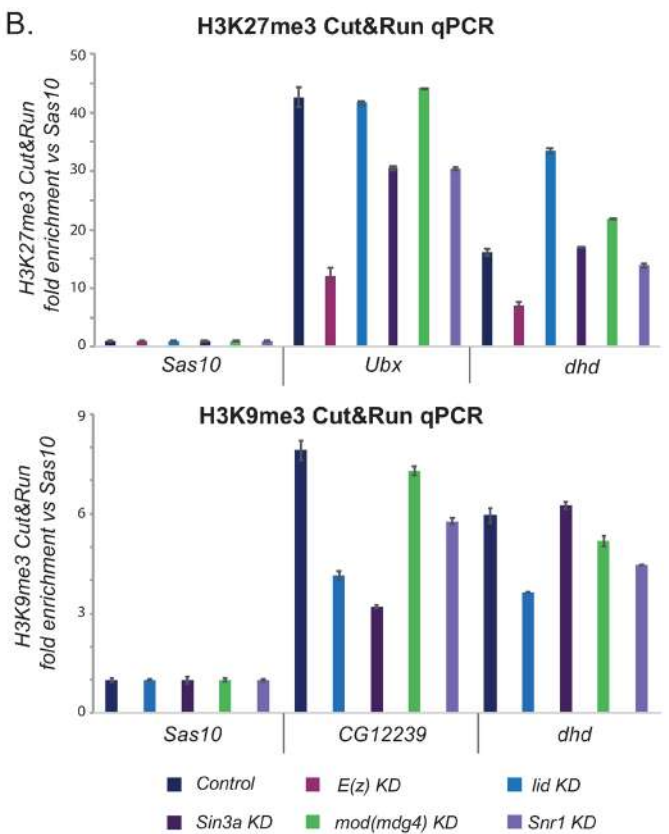


Figure S5

A.



B.



C.

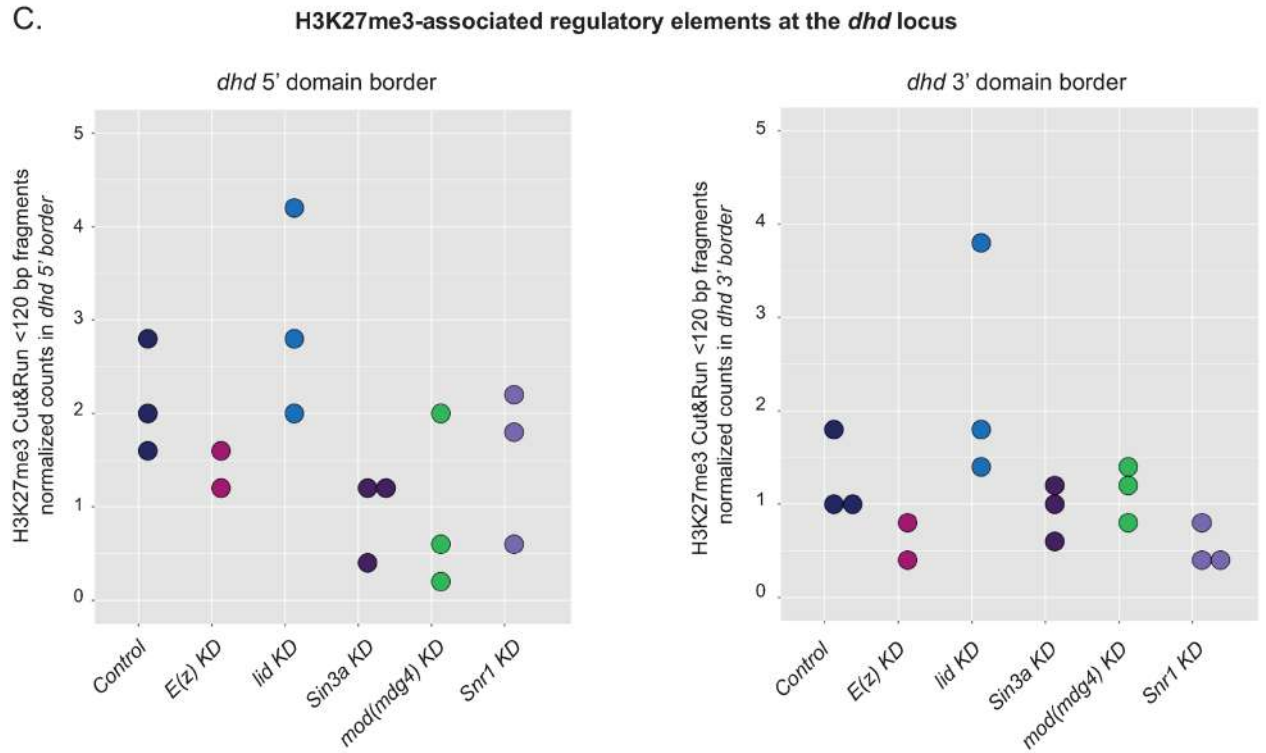


Figure S6

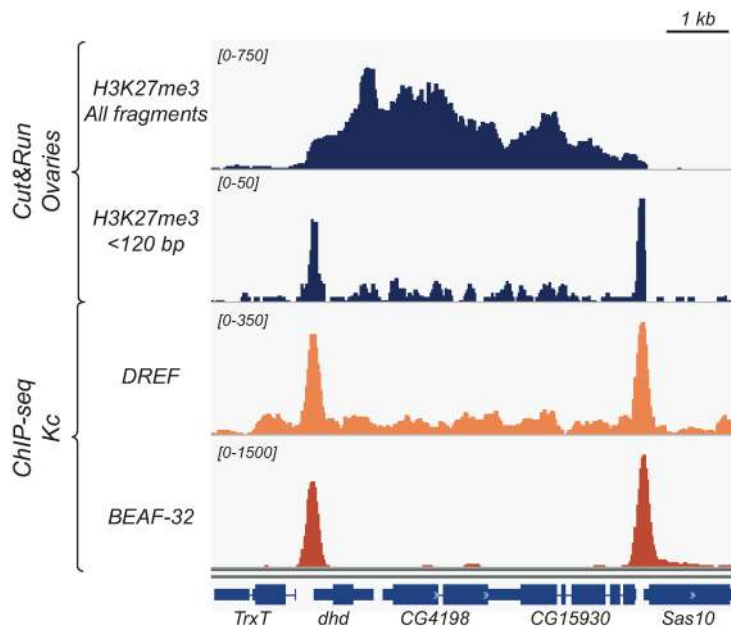


Figure S7

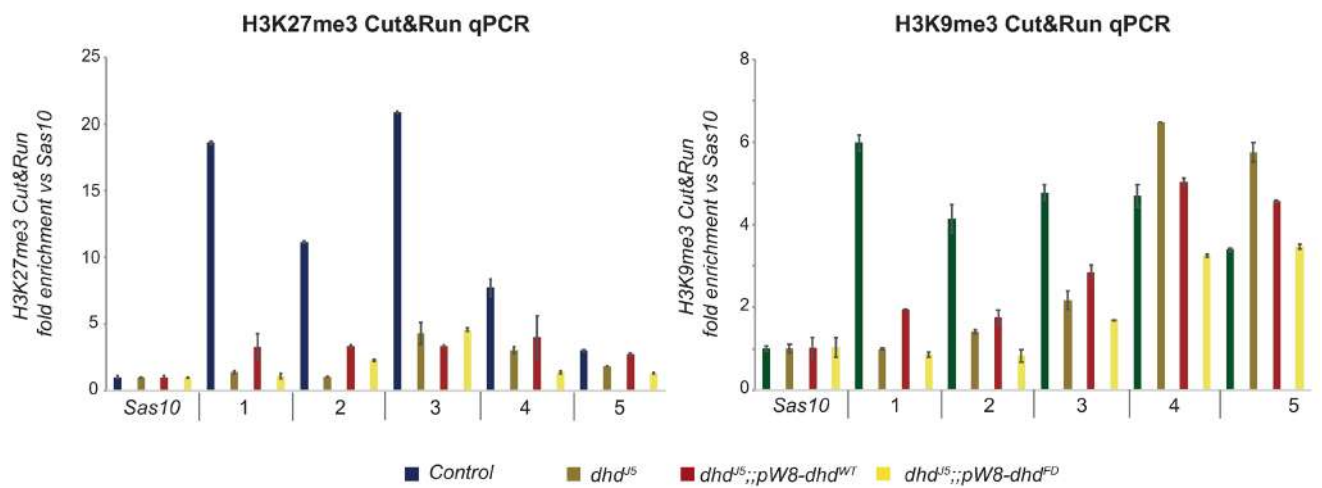


Figure S8

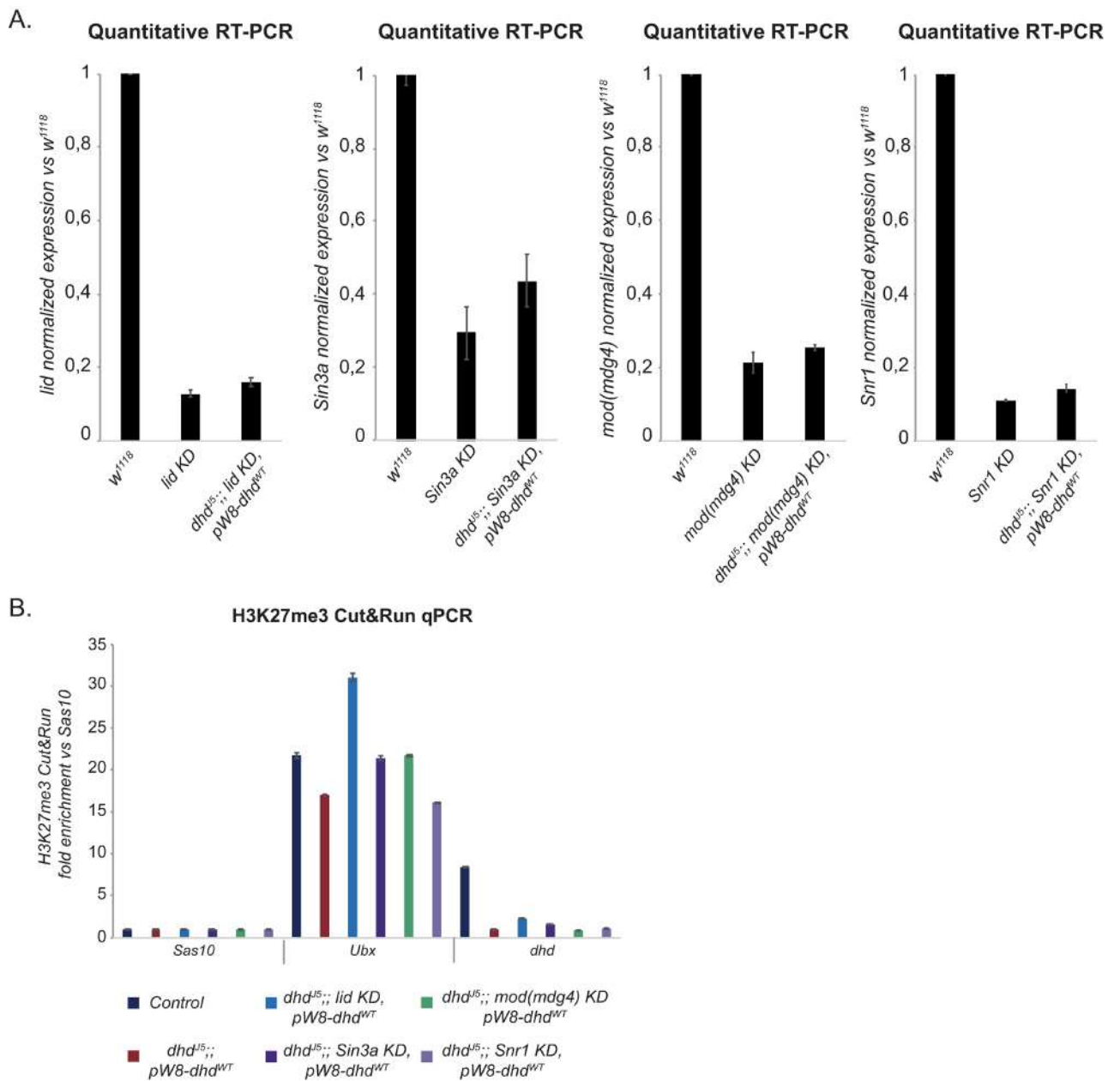


Table S1. List of primers used in this paper.

Primer	Sequence (5' → 3')	Purpose
dhd-for	TCTATGCGACATGGTGTGGT	RT-qPCR
dhd-rev	TCCACATCGATCTTGAGCAC	RT-qPCR
lid-for	ATTGGTTTCACGAGGATTGC	RT-qPCR
lid-rev	CATAGCCACTTGGGTCGATT	RT-qPCR
Sin3a-for	CGACAAATGGGTATCGTTCC	RT-qPCR
Sin3a-rev	GACCAGGTCCAGCTCGAAT	RT-qPCR
mod(mdg4)-for	CAACAGATCACCGTGCAAAC	RT-qPCR
mod(mdg4)-rev	GTTCAGATTTTCGTGGGCAAT	RT-qPCR
Snr1-for	TCAGCTCCCACATCTTAGCC	RT-qPCR
Snr1-rev	ACTGGCGTATCGGAAGTGTT	RT-qPCR
Rp49-for	AAGATCGTGAAGAAGCGCAC	RT-qPCR
Rp49-rev	GATACTGTCCCTTGAAGCGG	RT-qPCR
Ubx-for	AAAATTCCTCGGCCTGATTC	Cut&Run-qPCR
Ubx-rev	AAAAATGGGCGTAGCTCAGA	Cut&Run-qPCR
Sas10-for	AGGAGGAGCAGCAGGATGT	Cut&Run-qPCR
Sas10-rev	AGATCGCGGTCATCGTCTT	Cut&Run-qPCR
CG12239-for	AGATGAGGGACGAAGGATTG	Cut&Run-qPCR
CG12239-rev	TTCCTCGGTACGTTTCAGCTT	Cut&Run-qPCR
dhd-3UTR-for	GTTTAGCTTGTAAGCGCGAGA	Cut&Run-qPCR
dhd-3UTR-rev	ATATCATCTGGTCACTGCTGTTG	Cut&Run-qPCR
CG4198-1-for	GAGCAAGAAGAATGGCCAAC	Cut&Run-qPCR
CG4198-1-rev	CGATCGTTGAACTCCTGGAT	Cut&Run-qPCR
CG4198-2-for	GTGTCCATAGCGCAGCAG	Cut&Run-qPCR
CG4198-2-rev	AAGGTGAACATAACCCACAA	Cut&Run-qPCR
CG15930-1-for	ACCTGGACATCGGCTACATC	Cut&Run-qPCR
CG15930-1-rev	ACCATGTGCGAATTTTCGAT	Cut&Run-qPCR

CG15930-2-for	TTATTCCGCATTTTGGCACT	Cut&Run-qPCR
CG15930-2-for	GGAAGAAGCCGAGGATAACA	Cut&Run-qPCR
ΔDRE-1-for	GAGCAGCAGCCGAATTCGGTACCCCATATCCCTCCCATATCC	ΔDRE-transgene
ΔDRE-1-rev	CACGGCAGGTCTGAATATTATTATCGTAATTATGGCAAAATAATGGC	ΔDRE-transgene
ΔDRE-2-for	GATAATAATATTAGACCTGCCGTGGTGAATAAG	ΔDRE-transgene
ΔDRE-2-rev	CACAGGGATGCCACCCGGGATCCGCTAATGGAATCGCAATCGT	ΔDRE-transgene
FD-for	GATTGCGATTCCATTAGCGAATGCGGACGATATGCCAACGC	FD-transgene
FD-rev	GTCACAGGGATGCCACCCGGGATCCACAAAGAGAAAACTGTTGTA A	FD-transgene

Table S2. List of antibodies used in this paper.

Antibody	Host animal	Dilution	Experiment	Company (Catalog#)
Anti-Histones	mouse, monoclonal	1:1000	Immunofluorescence	Merck (#F152.C25.WJJ)
Anti-GFP	mouse, monoclonal	1:200	Immunofluorescence	Roche (#118144600001)
Anti-H3K27me3	rabbit, polyclonal	1:500	Immunofluorescence	Merck (#07-449)
Anti-H3K27me3	rabbit, monoclonal	1:100	Cut&Run	Cell Signalling Technology (#9733)
Anti-H3K9me3	rabbit, polyclonal	1:50	Cut&Run	Abcam (#8898)
Anti-DHD	rabbit, polyclonal	1:1000	Western Blot	<i>Tirmarche et al., 2016</i>
Anti- α -tubulin	mouse, monoclonal	1:500	Western Blot	Merck (#T9026)

III. Additional Results and Discussion

The oocyte-to-zygote transition involves a series of complex nuclear and cellular events, including completion of meiosis, selective translation of maternal RNAs and formation of male and female pronuclei. These events are mainly controlled by maternal factors present in the oocyte at the moment of fertilization. For example, the maternal HIRA histone chaperone complex is essential for the assembly of paternal chromatin at fertilization (Bonney et al., 2007), hinting thus on the importance of chromatin factors involved in zygote formation. This opened the question of what other maternal chromatin factors are required for the integration of paternal chromosomes into the zygote. To answer this, we performed a female germline specific shRNA genetic screen. Remarkably, among the 8 genes causing paternal chromosome loss (among 380 targets tested), 3 belonged to the HIRA complex and the rest were found to abolish *dhd* expression on a rather specific manner. This result revealed a complex network of epigenomic effectors dedicated to the hyperactivation of this small gene. I thus investigated the molecular mechanisms underlying this regulation.

1. Is there an “ovarian hyperactivation code”?

My work led me to discover a series of epigenomic effectors that play general roles in gene regulation but are critical for *dhd* expression in ovaries. I focused thus on the mechanisms mediated by: (i) the H3K4me3 demethylase Lid, (ii) the deacetylase complex scaffold Sin3A, (iii) the Snr1 subunit of the remodeler Swi/Snf complex and (iv) the chromatin factor Mod(mdg4). Available data in other cell contexts showed that depletion of Lid/dKDM5, Sin3A and Snr1 affects hundreds or even thousands of genes (Gajan et al., 2016; Liu and Secombe, 2015; Saha et al., 2016; Xie et al., 2017; Zamurrad et al., 2018). Genome-wide impact on transcription after for Mod(mdg4) depletion has not been studied. Nonetheless, this factor has also been involved in gene regulation (Gerasimova et al., 1995; Savitsky et al., 2016). Strikingly our RNA-seq data analysis showed that *dhd* is hypersensitive to the loss of any of these factors, suggesting that these ubiquitous regulators of gene expression have a specificity for *dhd* in ovaries.

A hypothesis we considered to explain *dhd* hyperactivation was the control of *dhd* by specific distal regulatory elements and/or its positioning in a genomic environment favorable to high transcription levels. There is more and more evidence that links the 3D folding of the genome to transcription. Indeed, the eukaryotic genome is constrained in the nucleus and it has been observed that it is partitioned in domains with similar transcriptional state (Szabo et al., 2019). An emerging hypothesis is that this can favor gene activation by creating compartments containing the necessary factors to favor transcription (Cho et al., 2018; Cramer, 2019; Sabari et al., 2018). Many factors are required for gene transcription, this could be thus a way to deliver quickly and in sufficient amount the necessary proteins at sites of high transcription activity. At a lower scale, eukaryotic genomes, including *Drosophila*'s are organized in self-interacting regions: TADs (Schwartz and Cavalli, 2017; Szabo et al., 2019). This means that DNA sequences within a TAD physically interact with each other more frequently than with sequences outside the TAD. It was reasonable therefore to consider that the 3D organization of the genome could favor *dhd* interaction with specific regulatory regions. However, using available Hi-C data in embryonic fly cells we did not identify any particular contacts at the *dhd* locus (Fig24). The expression of *dhd* is strictly restricted to ovaries, it is thus possible that specific contacts are only established in the germline at a precise developmental stage and are thus undetectable in other cells. Nevertheless, our *pW8-attB-dhd^{WT}* transgene inserted in the chromosome III is capable of restoring *dhd* expression to high levels, arguing against *dhd* endogenous genomic environment as a key feature for its hyperactivation. This suggests that, at least in some specific cases, a gene can be massively expressed quite independently of its genomic context and putative 3D contacts.

Chromosome X: 5,153,253 - 5,261,457

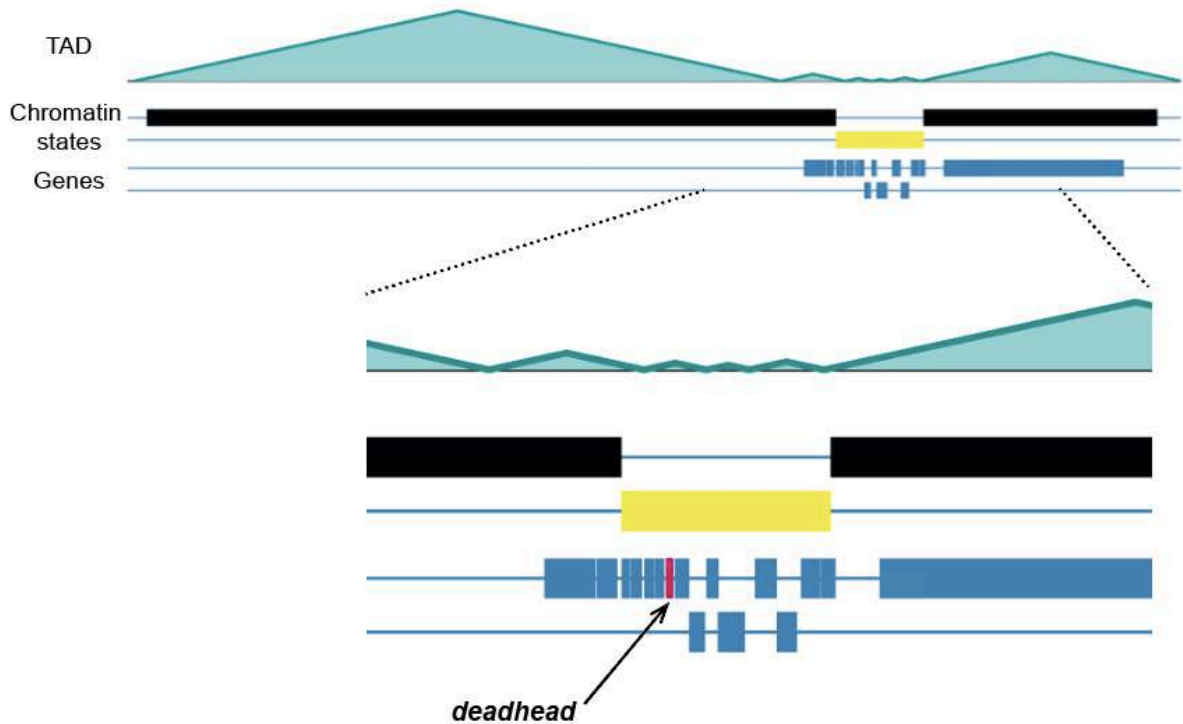


Fig24: *dhd* does not show contacts with other loci in cultured cells.

Representation of Hi-C datasets from *Drosophila* Kc embryonic cell using the ChoroGenome Navigator (Ramírez et al., 2018), showing TADs, chromatin state and genes at the indicated coordinates and a zoom on the *dhd* region. Yellow corresponds to active chromatin and black to inactive chromatin.

To explain *dhd* hyperactivation we can also look at the gene itself. Indeed, *dhd* gene architecture, short and intronless, could contribute to explain its hyperactivation. The process of eliminating introns is energy- and time-consuming. Therefore, *dhd* structure could provide an advantage in transcription and translation efficiency, resulting in its high expression. In humans, shorter gene length is associated with high expression, smaller proteins, and little intronic content. It was hypothesized that, due to the great levels of expression in smaller genes, there is a selective pressure to maximize protein synthesis efficiency (Urrutia, 2003). Also, the main fraction of intronless genes correspond to those coding for histones and G protein-coupled receptors and both of this gene families are abundantly expressed (Bryson-Richardson et al., 2004; Doenecke and Albig, 2005; Louhichi et al., 2011). Genes with longer transcripts are mostly associated with functions in the early development stages, while genes with smaller transcripts have important roles in more general functions (Lopes et al., 2021). A correlation between gene length and function has therefore been established. Shorter genes tend to be associated with housekeeping functions, however *dhd* is required only at a very

specific time and contrary to housekeeping genes, it does not need to be constitutively expressed. While *dhd* length could confer an advantage for high expression, the latter must be precisely regulated. Notably, the *dhd* transcribed region would correspond to ~4 nucleosomes. Nucleosomes represent a barrier for RNA Pol II progression, requiring thus a coordinated dance of nucleosomes sliding, evicting, assembling, among others. It is then possible that this low number of nucleosomes favors transcriptional efficiency.

Promoter-enhancer communication is a key point in transcriptional regulation. Both of these elements are bound by transcription factors recognizing specific DNA motifs. Motif scanning at the *dhd* promoter identified a tandem DRE motif. Remarkably, only 4 other gene promoters in the genome presented this motif with ovary-bias expression but none had high levels of expression, similar to *dhd*. DRE motifs are enriched at core promoters of highly expressed housekeeping genes (FitzGerald et al., 2006; Ohler et al., 2002). Additionally, DRE-binding factor Dref, preferentially binds to and activates housekeeping enhancers that are located closely to ubiquitously expressed genes with specificity to its core promoter, and suggests that the DRE motif is required and sufficient for housekeeping enhancer function (Zabidi et al., 2015). Although we were not able to identify any obvious candidate enhancer element in the vicinity of the *dhd* gene region we cannot exclude that this tandem DRE motif plays an important role in *dhd* hyperactivation by recruiting a still unknown enhancer. Nonetheless, this motif by its own, is not sufficient for ovarian hyperactivation.

As mentioned before, *dhd* is a short gene, it was then reasonable to consider that in order to optimize this retrained space, regulatory elements could be present within the coding sequence (CDS). During my PhD I used a rescue transgene, *pW8-attB-dhd^{WT}*, capable of restoring *dhd* expression to ~70% of its endogenous level. I also had at my disposal a transgenic line with a construct identical to the *pW8-attB-dhd^{WT}* transgene except that *dhd* CDS had been replaced by the similar in size CDS of the ubiquitous thioredoxin *Trx2* (Fig25-A) (Tirmarche, 2016). I thus sought to establish if the *pW8-attB-dhd>Trx2* transgene was able to induce similar *Trx2* levels to those found for *dhd* with the *pW8-attB-dhd^{WT}* transgene. Along with a master student, we tackled this question by measuring by RT-qPCR in ovaries transcripts emerging either from the *pW8-attB-dhd^{WT}* or the *pW8-attB-dhd>Trx2* transgene in a *dhd^{J5}* null mutant context. For this, we used primers targeting *dhd* 3'UTR region allowing us to measure endogenous *dhd* expression in the control as well as transcripts originating from both transgenes. We used the same method as described in (Torres-Campana et al.,

2020). Our preliminary results indicated that as previously observed, the *dhd* transgene is able to restore high *dhd* expression. However, transcripts emerging from the *pW8-attB-dhd>Trx2* transgene were weakly expressed. Indeed, they represented ~10% of the endogenous *dhd* levels and ~15% of the *dhd* levels retrieved with the *pW8-attB-dhd^{WT}* transgene (Fig25-B). Replacing the CDS of *dhd* by the CDS of *Trx2* does not suffice to induce high transcript levels of *Trx2*, comparable to those of *dhd*. These results suggest that *dhd* CDS is important to achieve high levels of transcription.

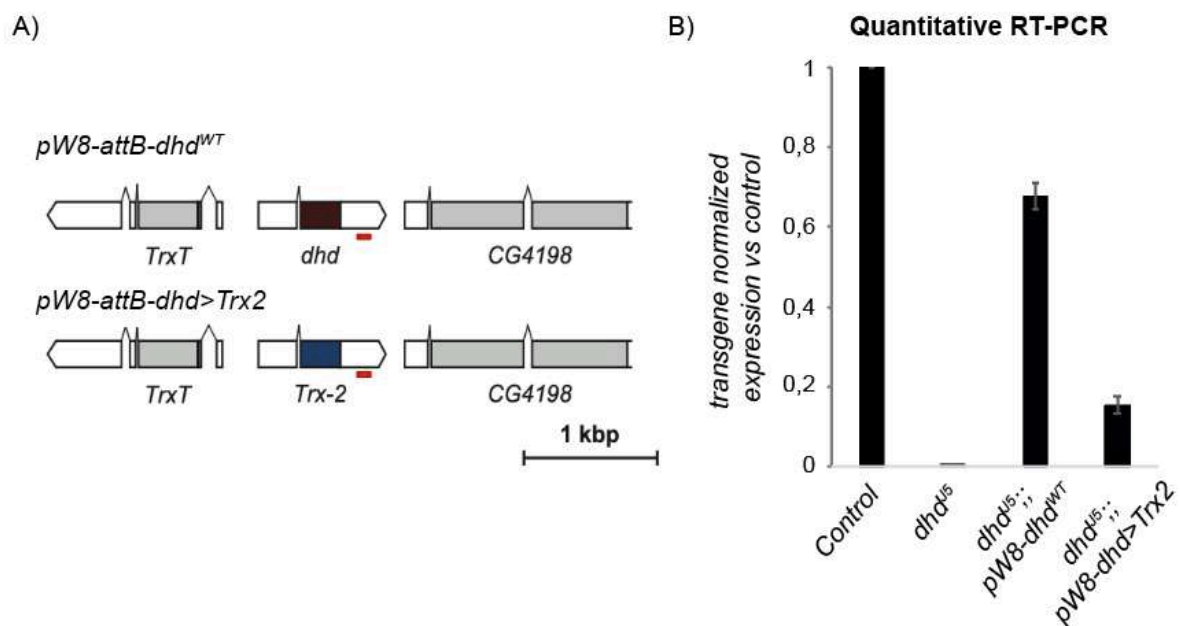


Fig 25- Replacement of *dhd* coding sequence does not replicate its highly expressed levels

(A) Schematic representation of the *pW8-attB-dhd^{WT}* and *pW8-attB-dhd>Trx2* transgenes. The *pW8-attB-dhd^{WT}* contains a genomic region of 4.3 kb containing the *dhd* gene. In the *pW8-attB-dhd>Trx2* transgene, the coding sequence of *dhd* (brown) was replaced by the one of *Trx2* (blue), to place it under the control of the *dhd* promoter. The red line indicates the region covered by the primers used in (B) (Tirmarche, 2016). (B) RT-qPCR quantification of *pW8-attB-dhd^{WT}* and *pW8-attB-dhd>Trx2* mRNA levels in ovaries from control flies, *dhd^{J5}* mutants or *dhd^{J5}* mutants carrying one of the two transgenes (normalized to *Rp49* and relative to expression in control). Data from biological duplicates analyzed in technical duplicates are presented as mean ± SEM.

How can a sequence be optimized for high transcription? Codon usage bias, the preference for certain synonymous codons, is a feature of eukaryotic and prokaryotic genomes that plays an important role in gene expression levels. It has been observed that this is, at least in part,

due to codon usage influence on translation rates (Duret and Mouchiroud, 1999; Weinberg et al., 2016; Yu et al., 2015; Zhao et al., 2017). Recent studies have also found a correlation between codon usage on mRNA levels due to changes in transcription in *Neurospora* (Zhou et al., 2016), suggesting that codon optimization could also play a role at the transcriptional level. It is thus reasonable, to hypothesize that *dhd* sequence contributes to its hyperactivation by this mechanism. It could thus be interesting to determine if *dhd* shows codon usage bias and if so, assess the impact on expression level when such codons are replaced by synonym codons. Also, the comparison of codon usage between *dhd* and other highly transcribed genes could shed light into the optimization of the CDS for transcriptional efficiency.

Finally, the peculiar genomic landscape of *dhd* is also worth considering in the recipe for hyperactivation. I found that *dhd* is embedded within a heterochromatic domain harboring both H3K27me3 and H3K9me3 while also being enriched for H3K4me3. The type of promoters carrying both active and repressive marks *i.e.*, bivalent promoters are associated to developmental genes in mammals (Bernstein et al., 2006; Lesch et al., 2013). When carrying both marks genes are repressed or weakly transcribed and upon cell differentiation, these bivalent regions undergo either full activation or stable silencing (Gaertner et al., 2012; Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007). Recently, this kind of domains were identified in *Drosophila* embryos at a handful of *Hox* genes (Akmammedov et al., 2019). If these domains represent promoters poised for activation or repression has not been investigated yet. However, our *dhd* rescue transgene is expressed to high levels even in the absence of both heterochromatic marks indicating that this epigenetic configuration is not necessary for *dhd* massive expression. Importantly, our data comes from whole ovaries, that is a mix of germ cells and somatic cells at various stages of differentiation. Due to technical limitations in our study, we cannot exclude that these marks are not cell-type and oogenesis-stage-specific.

In conclusion, I uncovered a series of specific features at the *dhd* gene. However, none of them alone seems to define an “ovarian hyperactivation code”. It is more likely that the ensemble of the particular characteristics of *dhd* are necessary to achieve its massive expression.

2. An unusual heterochromatic domain

Both H3K9me3 and H3K27me3 are commonly found in large extended domains however we uncovered an H3K27me3/H3K9me3 mini-domain flanked by regulatory elements.

How can this be achieved? In *Drosophila*, insulators and active transcription have been found important for Polycomb/H3K27me3 domain boundaries (De et al., 2020; Fujioka et al., 2013). CP190 is a protein common to most insulators (Ahanger et al., 2013) and in embryonic fly cells it binds both borders of the *dhd* H3K27me3 domain. The study by De et al., reported different categories of Polycomb boundaries in fly larvae (De et al., 2020). 22% of the boundaries analyzed corresponded to an insulator and an active promoter next to the Polycomb domain, transcribing away from the domain. This is the case observed at the *Sas10*-proximal boundary of the mini-domain we identified. However, the scenario of a gene transcribing towards the H3K27me3 domain, as is the case at the *dhd*-proximal border, was not observed in this study. Although the case of *dhd* transcribing within the heterochromatic domain is peculiar, both borders of the domain show active transcription and insulator proteins. It is thus possible that the presence of these two elements helps define these boundaries.

One of the features of heterochromatic H3K9me3 domains it's their ability to spread. However, we and others observe in the female germline a highly localized deposition of H3K9me3 [Smolko et al., 2018]. To target H3K9me3 deposition, RNA-based mechanisms or DNA-binding proteins mechanisms have been described. In their study, Smolko and colleagues found that H3K9me3 deposition by the histone methyltransferase dSETDB1 depended on Sxl, an RNA binding protein. It is thus possible that an RNA-based mechanism is at play. Of the three *Drosophila* enzymes known to methylate H3K9, only dSETDB1 is required for germline development. It is thus possible that this methyltransferase achieves a localized deposition of H3K9me3 through specific features and interactions with determined partners such as Sxl. In some cases, H3K9me3 is permissive to – or even required for – transcription (Lu et al., 2000; Smolko et al., 2018; Wakimoto and Hearn, 1990). Active heterochromatic genes have H3K9me3 on their bodies but not their TSSs, it is tempting to speculate that H3K9me3 is repressive only at promoter regions, while being compatible with, or even promoting, transcription on gene bodies.

Both H3K27me3 and H3K9me3 are known as important to mediate gene repression. The role they play at the *dhd* locus is thus intriguing. First, as suggested by the *dhd* rescue transgene, these marks are not necessary to achieve high levels of expression. Next, we did

not detect an ectopic expression of *dhd* in adult tissues when these marks were absent, arguing thus against a repressive role outside of ovarian tissue. Nonetheless, we cannot exclude a repressive role at other developmental stages.

H3K27me3 and H3K9me3 are compatible with transcription (Ahmad and Spens, 2019; Lu et al., 2000; Smolko et al., 2018; Wakimoto and Hearn, 1990) we can thus wonder if they could actually play a role in promoting transcription of *dhd*. It is also possible that this mini-domain has a role in repressing the other two genes present within it, *CG4198*, which is testis-specific, and *CG15930*, which shows low expression in larval imaginal discs (<http://flybase.org>). Of note, these two hypotheses are not mutually exclusive. Further studies will be necessary to elucidate the role of this heterochromatic mini-domain and shed light on the roles played by these marks on transcription regulation.

3. Different roles for epigenomic regulators

Lid, Sin3A, Snr1 and Mod(mdg4) are essential for *dhd* expression. They regulate *dhd* without significantly impacting H3K27me3 enrichment in the *dhd* domain, except for Lid who slightly counteracts it. This suggests that depletion of epigenomic effectors can alter transcription without major changes in the surrounding epigenomic landscape. In agreement, a recent study reported that depletion of the H3K4me1/2 demethylase dLsd1 in the female germline results in both up and down-regulation of gene expression in similar proportions, with a subset of misregulated genes directly bound by dLsd1 (Lepessant et al., 2020). In this study, they assessed histone marks in fly embryonic cells at four dLsd1-bound genes, two up-regulated and two down-regulated. The mark targeted by dLsd1, H3K4me2, increased upon its depletion. Interestingly, none of the other marks tested, H3K4me3, H3K9me2 and H3K27me3 were significantly affected. However, contrary to the case of *dhd*, low-transcribed genes in ovaries were enriched for the H3K27me3 mark while active ones were not, highlighting the enigmatic role of H3K27me3 at the *dhd* locus.

Snr1 is part of the *Drosophila* Swi/Snf remodeler complex, known to oppose Polycomb repression (Kassis et al., 2017; Kingston and Tamkun, 2014; Schuettengruber et al., 2017). In mammalian cells it has been found that mSWI/SNF can actively remove PRC complexes from chromatin thereby counteracting its repression (Kadoch et al., 2017). Surprisingly, we found that upon *snr1* knockdown there is a genome-wide decrease on H3K27me3, suggesting thus that Snr1 could play a role favoring this mark. Interestingly, a

recent study described a role for the mSWI/SNF complex in promoting polycomb repression (Weber et al., 2021). Their model proposes that in the absence of mSWI/SNF, PRC complexes accumulate where mSWI/SNF normally evicts them, thereby causing their redistribution away from heavily occupied sites and causing derepression at these sites. This result illustrates a mechanism where control of one locus can be exerted by effectively maintaining a certain amount of soluble repressive complexes in the cell. It would thus be interesting to measure the composition of soluble regulatory complexes in our knockdown conditions. This could help identify a putative regulator of *dhd* that becomes altered upon our knockdown conditions.

Another level of complexity in the study of the molecular mechanisms promoted by a single factor, comes from the fact that the latter can have different isoforms with different effects. Therefore, the impact on transcription can vary. This is particularly relevant for Mod(mdg4) since ~ 30 isoforms have been identified to date with a common N-Terminal domain (Büchner et al., 2000). Little is known about the impact of Mod(mdg4) depletion on genome-wide transcription. Nonetheless, it has been found that it promotes *Hox* gene activation (Gerasimova et al., 1995; Savitsky et al., 2016). Isoforms other than the insulator isoform 67.2 were found at gene promoters (Melnikova et al., 2019), as is the case for *dhd*, indicating potential targets regulated by specific Mod(mdg4) isoforms. We can wonder if there is a functional significance in having a single gene encoding for such a large number of isoforms. Gabler and colleagues compared *mod(mdg4)* orthologous loci from *D. melanogaster* and *D. virilis* and found an evolutionary conservation of all Mod(mdg4) isoforms (Gabler et al., 2005). The large number of isoforms is thus functionally important in both *Drosophila* species. Additionally, they observed the conservation of the unique C-terminal regions pointing towards a functional differentiation between single isoforms. It would therefore be interesting to assess if, as in *D. melanogaster*, in other species, Mod(mdg4) isoforms can play distinct roles such as insulation and activation.

Emerging evidence, including my study, suggest that the influence of epigenomic effectors on transcriptional outcomes cannot just be categorized as activating or repressing. Most of these factors have pivotal roles depending on cellular context, interactors and targets.

4. An approach to profile regulatory architecture of chromatin domains.

Due to heterogenous cell populations in complex tissues, the study of the chromatin landscape at cell-type specific genes is a major challenge. Great efforts have thus been deployed in the development of single-cell profiling technologies (Lee et al., 2020), opening the possibility to profile chromatin in isolated single cells. Clustering analysis can then be used to retrospectively assign identities to individual cells and give an insight into the epigenomic features of individual cell types. Yet, these approaches suffer from low per-cell throughput, leading to relatively low resolution, high false-negative rates and therefore complicating the interpretations.

Alternatively, the dual-enzyme single-molecule footprinting (dSMF) allows to measure protein-DNA contacts at the level of single molecules (Krebs et al., 2017). This method uses both GpC and CpG DNA methyltransferases to methylate exposed DNA *in vivo*, followed by bisulfite long-read sequencing. With this approach, the occupation of chromatin by nucleosomes, transcription factors and polymerases can be resolved on individual DNA molecules at the scale of several kilobases. While this approach cannot identify the cell of origin of each individual DNA molecule, it can be used to make deductions on the co-occupancy on DNA of different chromatin components. Analogously, the Cut&Run analysis I implemented revealed the co-occupancy on DNA of H3K27me3 nucleosomes and transcription factors at their associated insulators and PREs. This was possible because in Cut&Run, antibody-bound MNase does not restrict its activity to its sole protein target but can rather cleave DNA in spatial proximity of the affected nucleosome.

Similarly, a recent study described the CUTAC method (Henikoff et al., 2020). Authors showed that lowering salt concentrations during antibody tethering in Cut&Tag can reveal H3K4me3-associated DNA regulatory elements (*i.e.*, associated with active chromatin). A relative advantage to our strategy is that histone modifications and regulatory elements can be mapped simultaneously, under a single experimental condition. Yet, compared to CUTAC, our approach presents two likely caveats. First, DNA regulatory inference from Cut&Run is likely to be more difficult to apply to lower input samples, as small cleavage fragments are quite rare compared to nucleosome-sized fragments (1 in 10

approximately). Second, DNA regulatory elements which transcription factor footprint is at least as large as a nucleosome are likely to be missed.

Further exploration of how combination of these approaches may help circumvent the issues associated to each approach. For example, a recent study has combined high-resolution nuclease footprinting with single-molecule methylation profiling (MNase-seq, ORGANIC ChIP, CUT&RUN, and dSMF) to study transcription factor cooperativity at active enhancers on a genome-wide scale in *Drosophila* embryonic cells. This shows how the combination of different approaches can yield valuable insight on the chromatin regulatory landscape. These are thus promising methods, as they hold the potential to universally map the regulatory landscape around histone modifications, which are universal, opening the possibility to implement them in any tissue or species.

General Conclusion

In almost every animal, the first stages of life depend almost exclusively on maternal products contained in the mature oocyte. Preparation of a competent egg represents an extreme cellular differentiation process where complex transcriptional programs must be tightly regulated. Emerging evidence has highlighted the importance of epigenetic mechanisms in the establishment of the molecular basis of the crucial and complex oocyte-to-zygote transition.

Indeed, the work I presented focused on the intricate regulatory network of epigenomic effectors to achieve hyperactivation of the oocyte-to-zygote effector: the maternal thioredoxin Deadhead. Interestingly, the *dhd* gene can achieve massive levels of transcription with a short regulatory sequence, surrounded by silenced genes, without apparent particular 3D contacts and embedded in a heterochromatic mini-domain. While in different cellular contexts, *dhd* regulators, Lid, Sin3A, Snr1 and Mod(mdg4) have broad effects on transcription, in ovaries, the *dhd* gene represents a hypersensitive target to the loss of any of these factors. The example of *dhd* is thus an illustration of how the activation of a single gene in the right place, at the right time and in the right amount, requires a unique recipe involving a series of factors, that at first glance might not be expected to cooperate.

This study also raises the question of the role of non-canonical heterochromatin domains. Although *dhd* was embedded in an H3K27me3/H3K9me3 mini-domain, these marks were not necessary for its expression in ovaries, nor for its repression in adult tissues. Further studies will be needed to understand the role of these marks, usually mediating transcriptional repression, at actively transcribed genes.

Altogether, my work identified a series of chromatin factors necessary for the establishment of specific transcriptional programs during oogenesis. Importantly, it showed how mechanisms underlying chromatin-based gene regulation are highly dependent on cellular context, interactors of effectors and targeted genes.

References:

- Adam M, Robert F, Laroche M, Gaudreau L. H2A.Z Is Required for Global Chromatin Integrity and for Recruitment of RNA Polymerase II under Specific Conditions. *Mol Cell Biol* 2001;21:6270–9. <https://doi.org/10.1128/MCB.21.18.6270-6279.2001>.
- Agarwal N, Ansari A. Enhancement of Transcription by a Splicing-Competent Intron Is Dependent on Promoter Directionality. *PLoS Genet* 2016;12:e1006047. <https://doi.org/10.1371/journal.pgen.1006047>.
- Ahanger SH, Shouche YS, Mishra RK. Functional sub-division of the *Drosophila* genome via chromatin looping: The emerging importance of CP190. *Nucleus* 2013;4:115–22. <https://doi.org/10.4161/nucl.23389>.
- Ahmad K, Henikoff S. The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly. *Molecular Cell* 2002;9:1191–200. [https://doi.org/10.1016/S1097-2765\(02\)00542-7](https://doi.org/10.1016/S1097-2765(02)00542-7).
- Ahmad K, Spens AE. Separate Polycomb Response Elements control chromatin state and activation of the vestigial gene. *PLoS Genet* 2019;15:e1007877. <https://doi.org/10.1371/journal.pgen.1007877>.
- Ahn SH, Kim M, Buratowski S. Phosphorylation of Serine 2 within the RNA Polymerase II C-Terminal Domain Couples Transcription and 3' End Processing. *Molecular Cell* 2004;13:67–76. [https://doi.org/10.1016/S1097-2765\(03\)00492-1](https://doi.org/10.1016/S1097-2765(03)00492-1).
- Akkouche A, Mugat B, Barckmann B, Varela-Chavez C, Li B, Raffel R, et al. Piwi Is Required during *Drosophila* Embryogenesis to License Dual-Strand piRNA Clusters for Transposon Repression in Adult Ovaries. *Molecular Cell* 2017;66:411–419.e4. <https://doi.org/10.1016/j.molcel.2017.03.017>.
- Akmammedov A, Geigges M, Paro R. Bivalency in *Drosophila* embryos is associated with strong inducibility of Polycomb target genes. *Fly (Austin)* 2019;13:42–50. <https://doi.org/10.1080/19336934.2019.1619438>.
- Albert B, Léger-Silvestre I, Normand C, Ostermaier MK, Pérez-Fernández J, Panov KI, et al. RNA polymerase I-specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle. *Journal of Cell Biology* 2011;192:277–93. <https://doi.org/10.1083/jcb.201006040>.
- Allshire RC, Madhani HD. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* 2018;19:229–44. <https://doi.org/10.1038/nrm.2017.119>.

Andersen PR, Tirian L, Vunjak M, Brennecke J. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* 2017;549:54–9. <https://doi.org/10.1038/nature23482>.

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Current Biology* 2001;11:1017–27. [https://doi.org/10.1016/S0960-9822\(01\)00299-8](https://doi.org/10.1016/S0960-9822(01)00299-8).

Ardehali MB, Mei A, Zobeck KL, Caron M, Lis JT, Kusch T. *Drosophila* Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription: dSet1 is the major H3K4 trimethyltransferase. *The EMBO Journal* 2011;30:2817–28. <https://doi.org/10.1038/emboj.2011.194>.

Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences* 1991;88:10148–52. <https://doi.org/10.1073/pnas.88.22.10148>.

Arimbasseri AG, Maraia RJ. RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends in Biochemical Sciences* 2016;41:546–59. <https://doi.org/10.1016/j.tibs.2016.03.003>.

Arnér ESJ, Holmgren A. Physiological functions of thioredoxin and thioredoxin reductase: Thioredoxin and thioredoxin reductase. *European Journal of Biochemistry* 2000;267:6102–9. <https://doi.org/10.1046/j.1432-1327.2000.01701.x>.

Arnold CD, Gerlach D, Stelzer C, Boryn LM, Rath M, Stark A. Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. *Science* 2013;339:1074–7. <https://doi.org/10.1126/science.1232542>.

Avilés-Pagán EE, Kang ASW, Orr-Weaver TL. Identification of New Regulators of the Oocyte-to-Embryo Transition in *Drosophila*. *G3* 2020;10:2989–98. <https://doi.org/10.1534/g3.120.401415>.

Avilés-Pagán EE, Orr-Weaver TL. Activating embryonic development in *Drosophila*. *Seminars in Cell & Developmental Biology* 2018;84:100–10. <https://doi.org/10.1016/j.semcdb.2018.02.019>.

Ayyanathan K. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes & Development* 2003;17:1855–69. <https://doi.org/10.1101/gad.1102803>.

Banerji J, Rusconi S, Schaffner W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* 1981;27:299–308. [https://doi.org/10.1016/0092-8674\(81\)90413-X](https://doi.org/10.1016/0092-8674(81)90413-X).

Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res* 2011;21:381–95. <https://doi.org/10.1038/cr.2011.22>.

Bastock R, St Johnston D. *Drosophila* oogenesis. *Curr Biol* 2008;18:R1082-1087. <https://doi.org/10.1016/j.cub.2008.09.011>.

Becker PB, Workman JL. Nucleosome Remodeling and Epigenetics. *Cold Spring Harbor Perspectives in Biology* 2013;5:a017905–a017905. <https://doi.org/10.1101/cshperspect.a017905>.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse. *Cell* 2005;120:169–81. <https://doi.org/10.1016/j.cell.2005.01.001>.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006;125:315–26. <https://doi.org/10.1016/j.cell.2006.02.041>.

Bernstein BE, Tong JK, Schreiber SL. Genomewide studies of histone deacetylase function in yeast. *Proceedings of the National Academy of Sciences* 2000;97:13708–13. <https://doi.org/10.1073/pnas.250477697>.

Beuchle D, Struhl G, Muller J. Polycomb group proteins and heritable silencing of *Drosophila* Hox genes. *Development* 2001;128:993–1004. <https://doi.org/10.1242/dev.128.6.993>.

Biswas S, Rao CM. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *European Journal of Pharmacology* 2018;837:8–24. <https://doi.org/10.1016/j.ejphar.2018.08.021>.

Blastyák A, Mishra RK, Karch F, Gyurkovics H. Efficient and Specific Targeting of Polycomb Group Proteins Requires Cooperative Interaction between Grainyhead and Pleiohomeotic. *Mol Cell Biol* 2006;26:1434–44. <https://doi.org/10.1128/MCB.26.4.1434-1444.2006>.

Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* 2016;529:418–22. <https://doi.org/10.1038/nature16496>.

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, et al. Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes. *PLoS Biol* 2005;3:e157. <https://doi.org/10.1371/journal.pbio.0030157>.

Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 2017;171:557-572.e24. <https://doi.org/10.1016/j.cell.2017.09.043>.

Bonnefoy E, Orsi GA, Couble P, Loppin B. The Essential Role of *Drosophila* HIRA for De Novo Assembly of Paternal Chromatin at Fertilization. *PLoS Genet* 2007;3:e182. <https://doi.org/10.1371/journal.pgen.0030182>.

Bowman SK, Deaton AM, Domingues H, Wang PI, Sadreyev RI, Kingston RE, et al. H3K27 modifications define segmental regulatory domains in the *Drosophila* bithorax complex. *ELife* 2014;3:e02833. <https://doi.org/10.7554/eLife.02833>.

Brent MM, Anand R, Marmorstein R. Structural Basis for DNA Recognition by FoxO1 and Its Regulation by Posttranslational Modification. *Structure* 2008;16:1407–16. <https://doi.org/10.1016/j.str.2008.06.013>.

Brown JL, Kassis JA. Spps, a *Drosophila* Sp1/KLF family member, binds to PREs and is required for PRE activity late in development. *Development* 2010;137:2597–602. <https://doi.org/10.1242/dev.047761>.

Bryson-Richardson RJ, Logan DW, Currie PD, Jackson IJ. Large-scale analysis of gene structure in rhodopsin-like GPCRs: evidence for widespread loss of an ancient intron. *Gene* 2004;338:15–23. <https://doi.org/10.1016/j.gene.2004.05.001>.

Buchenau P, Hodgson J, Strutt H, Arndt-Jovin DJ. The Distribution of Polycomb-Group Proteins During Cell Division and Development in *Drosophila* Embryos: Impact on Models for Silencing. *Journal of Cell Biology* 1998;141:469–81. <https://doi.org/10.1083/jcb.141.2.469>.

Büchner K, Roth P, Schotta G, Krauss V, Saumweber H, Reuter G, et al. Genetic and Molecular Complexity of the Position Effect Variegation Modifier *mod(mdg4)* in *Drosophila*. *Genetics* 2000;155:141–57. <https://doi.org/10.1093/genetics/155.1.141>.

Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, et al. A transcription factor-based mechanism for mouse heterochromatin formation. *Nat Struct Mol Biol* 2012;19:1023–30. <https://doi.org/10.1038/nsmb.2382>.

Buratowski S. Progression through the RNA Polymerase II CTD Cycle. *Molecular Cell* 2009;36:541–6. <https://doi.org/10.1016/j.molcel.2009.10.019>.

Burke TW, Kadonaga JT. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes & Development* 1996;10:711–24. <https://doi.org/10.1101/gad.10.6.711>.

Burton A, Brochard V, Galan C, Ruiz-Morales ER, Rovira Q, Rodriguez-Terrones D, et al. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat Cell Biol* 2020;22:767–78. <https://doi.org/10.1038/s41556-020-0536-6>.

Cairns BR. The logic of chromatin architecture and remodelling at promoters. *Nature* 2009;461:193–8. <https://doi.org/10.1038/nature08450>.

Cammas F. Association of the transcriptional corepressor TIF1 with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes & Development* 2004;18:2147–60. <https://doi.org/10.1101/gad.302904>.

Cano-Rodriguez D, Gjaltema RAF, Jilderda LJ, Jellema P, Dokter-Fokkens J, Ruiters MHJ, et al. Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. *Nat Commun* 2016;7:12284. <https://doi.org/10.1038/ncomms12284>.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* 2002;298:1039–43. <https://doi.org/10.1126/science.1076997>.

Carey M, Li B, Workman JL. RSC Exploits Histone Acetylation to Abrogate the Nucleosomal Block to RNA Polymerase II Elongation. *Molecular Cell* 2006;24:481–7. <https://doi.org/10.1016/j.molcel.2006.09.012>.

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 Methylation by Set2 Directs Deacetylation of Coding Regions by Rpd3S to Suppress Spurious Intragenic Transcription. *Cell* 2005;123:581–92. <https://doi.org/10.1016/j.cell.2005.10.023>.

Cattoni DI, Cardozo Gizzi AM, Georgieva M, Di Stefano M, Valeri A, Chamousset D, et al. Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. *Nat Commun* 2017;8:1753. <https://doi.org/10.1038/s41467-017-01962-x>.

Cesarini E, Mozzetta C, Marullo F, Gregoretti F, Gargiulo A, Columbaro M, et al. Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *Journal of Cell Biology* 2015;211:533–51. <https://doi.org/10.1083/jcb.201504035>.

- Chaubal A, Pile LA. Same agent, different messages: insight into transcriptional regulation by SIN3 isoforms. *Epigenetics & Chromatin* 2018;11:17. <https://doi.org/10.1186/s13072-018-0188-y>.
- Cho W-K, Spille J-H, Hecht M, Lee C, Li C, Grube V, et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 2018;361:412–5. <https://doi.org/10.1126/science.aar4199>.
- Chubb JR, Trcek T, Shenoy SM, Singer RH. Transcriptional Pulsing of a Developmental Gene. *Current Biology* 2006;16:1018–25. <https://doi.org/10.1016/j.cub.2006.03.092>.
- Clapier CR, Cairns BR. The Biology of Chromatin Remodeling Complexes. *Annu Rev Biochem* 2009;78:273–304. <https://doi.org/10.1146/annurev.biochem.77.062706.153223>.
- Claycomb J, Orrweaver T. Developmental gene amplification: insights into DNA replication and gene expression. *Trends in Genetics* 2005;21:149–62. <https://doi.org/10.1016/j.tig.2005.01.009>.
- Cramer P. Organization and regulation of gene transcription. *Nature* 2019;573:45–54. <https://doi.org/10.1038/s41586-019-1517-4>.
- Cubeñas-Potts C, Rowley MJ, Lyu X, Li G, Lei EP, Corces VG. Different enhancer classes in *Drosophila* bind distinct architectural proteins and mediate unique chromatin interactions and 3D architecture. *Nucleic Acids Research* 2017;45:1714–30. <https://doi.org/10.1093/nar/gkw1114>.
- Curry E, Zeller C, Masrour N, Patten DK, Gallon J, Wilhelm-Benartzi CS, et al. Genes Predisposed to DNA Hypermethylation during Acquired Resistance to Chemotherapy Are Identified in Ovarian Tumors by Bivalent Chromatin Domains at Initial Diagnosis. *Cancer Res* 2018;78:1383–91. <https://doi.org/10.1158/0008-5472.CAN-17-1650>.
- Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. *Drosophila* Enhancer of Zeste/ESC Complexes Have a Histone H3 Methyltransferase Activity that Marks Chromosomal Polycomb Sites. *Cell* 2002;111:185–96. [https://doi.org/10.1016/S0092-8674\(02\)00975-3](https://doi.org/10.1016/S0092-8674(02)00975-3).
- van Daal A, White EM, Gorovsky MA, Elgin CR. *Drosophila* has a single copy of the gene encoding a highly conserved histone H2A variant of the H2A. F/Z type. *Nucl Acids Res* 1988;16:7487–97. <https://doi.org/10.1093/nar/16.15.7487>.
- Das TK, Sangodkar J, Negre N, Narla G, Cagan RL. Sin3a acts through a multi-gene module to regulate invasion in *Drosophila* and human tumors. *Oncogene* 2013;32:3184–97. <https://doi.org/10.1038/onc.2012.326>.

- Davidson, E. *Gene Activity in Early Development*. Academic Press, Orlando, FL; 1986.
- De S, Gehred ND, Fujioka M, Chan FW, Jaynes JB, Kassis JA. Defining the Boundaries of Polycomb Domains in *Drosophila*. *Genetics* 2020;216:689–700. <https://doi.org/10.1534/genetics.120.303642>.
- Deal RB, Henikoff JG, Henikoff S. Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones. *Science* 2010;328:1161–4. <https://doi.org/10.1126/science.1186777>.
- Deaton AM, Gómez-Rodríguez M, Mieczkowski J, Tolstorukov MY, Kundu S, Sadreyev RI, et al. Enhancer regions show high histone H3.3 turnover that changes during differentiation. *ELife* 2016;5:e15316. <https://doi.org/10.7554/eLife.15316>.
- Dej KJ, Spradling AC. Nurse cell chromosome reorganization 1999:11.
- Déjardin J, Rappailles A, Cuvier O, Grimaud C, Decoville M, Locker D, et al. Recruitment of *Drosophila* Polycomb group proteins to chromatin by DSP1. *Nature* 2005;434:533–8. <https://doi.org/10.1038/nature03386>.
- Delattre M, Spierer A, Tonka CH, Spierer P. The genomic silencing of position-effect variegation in *Drosophila melanogaster*: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. *J Cell Sci* 2000;113 Pt 23:4253–61.
- DeLuca SZ, Ghildiyal M, Niu W, Pang L-Y, Spradling AC. Differentiating *Drosophila* female germ cells initiate Polycomb silencing by altering PRC2 sampling. *Developmental Biology*; 2020. <https://doi.org/10.1101/2020.02.03.932582>.
- Deng WM, Althausen C, Ruohola-Baker H. Notch-Delta signaling induces a transition from mitotic cell cycle to endocycle in *Drosophila* follicle cells. *Development* 2001;128:4737–46.
- de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, et al. CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell* 2015;60:676–84. <https://doi.org/10.1016/j.molcel.2015.09.023>.
- Di Stefano L, Walker JA, Burgio G, Corona DFV, Mulligan P, Naar AM, et al. Functional antagonism between histone H3K4 demethylases in vivo. *Genes & Development* 2011;25:17–28. <https://doi.org/10.1101/gad.1983711>.
- Dingwall AK, Beek SJ, McCallum CM, Tamkun JW, Kalpana GV, Goff SP, et al. The *Drosophila* snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *MBoC* 1995;6:777–91. <https://doi.org/10.1091/mbc.6.7.777>.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;485:376–80. <https://doi.org/10.1038/nature11082>.

Doenecke D, Albig W. Intronless Genes. In: John Wiley & Sons, Ltd, editor. Encyclopedia of Life Sciences, Chichester: John Wiley & Sons, Ltd; 2005, p. a0005022.
<https://doi.org/10.1038/npg.els.0005022>.

Dorafshan E, Kahn TG, Schwartz YB. Hierarchical recruitment of Polycomb complexes revisited. *Nucleus* 2017;8:496–505. <https://doi.org/10.1080/19491034.2017.1363136>.

Dorn R, Krauss V, Reuter G, Saumweber H. The enhancer of position-effect variegation of *Drosophila*, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *PNAS* 1993;90:11376–80.
<https://doi.org/10.1073/pnas.90.23.11376>.

Douillet D, Sze CC, Ryan C, Piunti A, Shah AP, Ugarenko M, et al. Uncoupling histone H3K4 trimethylation from developmental gene expression via an equilibrium of COMPASS, Polycomb and DNA methylation. *Nat Genet* 2020;52:615–25.
<https://doi.org/10.1038/s41588-020-0618-1>.

Drelon C, Belalcazar HM, Secombe J. The Histone Demethylase KDM5 Is Essential for Larval Growth in *Drosophila*. *Genetics* 2018;genetics.301004.2018.
<https://doi.org/10.1534/genetics.118.301004>.

Drouin S, Laramée L, Jacques P-É, Forest A, Bergeron M, Robert F. DSIF and RNA Polymerase II CTD Phosphorylation Coordinate the Recruitment of Rpd3S to Actively Transcribed Genes. *PLoS Genet* 2010;6:e1001173.
<https://doi.org/10.1371/journal.pgen.1001173>.

Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang Xianglin, et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 2017;547:232–5.
<https://doi.org/10.1038/nature23263>.

Duret L, Mouchiroud D. Expression pattern and, surprisingly, gene length shape codon usage in *Caenorhabditis*, *Drosophila*, and *Arabidopsis*. *Proceedings of the National Academy of Sciences* 1999;96:4482–7. <https://doi.org/10.1073/pnas.96.8.4482>.

Dwyer K, Agarwal N, Pile L, Ansari A. Gene Architecture Facilitates Intron-Mediated Enhancement of Transcription. *Front Mol Biosci* 2021;8:669004.
<https://doi.org/10.3389/fmolb.2021.669004>.

Ecco G, Cassano M, Kauzlaric A, Duc J, Coluccio A, Offner S, et al. Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental Cell* 2016;36:611–23. <https://doi.org/10.1016/j.devcel.2016.02.024>.

Egloff S, Dienstbier M, Murphy S. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends in Genetics* 2012;28:333–41. <https://doi.org/10.1016/j.tig.2012.03.007>.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, et al. Serine-7 of the RNA Polymerase II CTD Is Specifically Required for snRNA Gene Expression. *Science* 2007;318:1777–9. <https://doi.org/10.1126/science.1145989>.

Eichhorn SW, Subtelny AO, Kronja I, Kwasniewski JC, Orr-Weaver TL, Bartel DP. mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in *Drosophila* oocytes and early embryos. *eLife* 2016;5:e16955. <https://doi.org/10.7554/eLife.16955>.

Emelyanov AV, Fyodorov DV. Thioredoxin-dependent disulfide bond reduction is required for protamine eviction from sperm chromatin. *Genes Dev* 2016;30:2651–6. <https://doi.org/10.1101/gad.290916.116>.

Erceg J, Pakozdi T, Marco-Ferrerres R, Ghavi-Helm Y, Girardot C, Bracken AP, et al. Dual functionality of *cis*-regulatory elements as developmental enhancers and Polycomb response elements. *Genes Dev* 2017;31:590–602. <https://doi.org/10.1101/gad.292870.116>.

Eskeland R, Eberharter A, Imhof A. HP1 Binding to Chromatin Methylated at H3K9 Is Enhanced by Auxiliary Factors. *Mol Cell Biol* 2007;27:453–65. <https://doi.org/10.1128/MCB.01576-06>.

Esnault C, Ghavi-Helm Y, Brun S, Soutourina J, Van Berkum N, Boschiero C, et al. Mediator-Dependent Recruitment of TFIIH Modules in Preinitiation Complex. *Molecular Cell* 2008;31:337–46. <https://doi.org/10.1016/j.molcel.2008.06.021>.

Eychenne T, Novikova E, Barrault M-B, Alibert O, Boschiero C, Peixeiro N, et al. Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture. *Genes Dev* 2016;30:2119–32. <https://doi.org/10.1101/gad.285775.116>.

Ferreira H, Flaus A, Owen-Hughes T. Histone Modifications Influence the Action of Snf2 Family Remodelling Enzymes by Different Mechanisms. *Journal of Molecular Biology* 2007;374:563–79. <https://doi.org/10.1016/j.jmb.2007.09.059>.

Fischer T, Cui B, Dhakshnamoorthy J, Zhou M, Rubin C, Zofall M, et al. Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *Proceedings of the National Academy of Sciences* 2009;106:8998–9003. <https://doi.org/10.1073/pnas.0813063106>.

FitzGerald PC, Sturgill D, Shyakhtenko A, Oliver B, Vinson C. Comparative genomics of *Drosophila* and human core promoters. *Genome Biol* 2006;7:R53. <https://doi.org/10.1186/gb-2006-7-7-r53>.

Flavahan WA, Drier Y, Liao BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* 2016;529:110–4. <https://doi.org/10.1038/nature16490>.

Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schöpflin R, et al. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* 2016;538:265–9. <https://doi.org/10.1038/nature19800>.

Freier R, Aragón E, Bagiński B, Pluta R, Martin-Malpartida P, Ruiz L, et al. Structures of the germline-specific Deadhead and thioredoxin T proteins from *Drosophila melanogaster* reveal unique features among thioredoxins. *IUCrJ* 2021;8:281–94. <https://doi.org/10.1107/S2052252521000221>.

Fudenberg G, Abdennur N, Imakaev M, Goloborodko A, Mirny LA. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb Symp Quant Biol* 2017;82:45–55. <https://doi.org/10.1101/sqb.2017.82.034710>.

Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports* 2016;15:2038–49. <https://doi.org/10.1016/j.celrep.2016.04.085>.

Fujioka M, Sun G, Jaynes JB. The *Drosophila* eve Insulator Homie Promotes eve Expression and Protects the Adjacent Gene from Repression by Polycomb Spreading. *PLoS Genet* 2013;9:e1003883. <https://doi.org/10.1371/journal.pgen.1003883>.

Fukaya T, Lim B, Levine M. Enhancer Control of Transcriptional Bursting. *Cell* 2016;166:358–68. <https://doi.org/10.1016/j.cell.2016.05.025>.

Furger A, O’Sullivan JM, Binnie A, Lee BA, Proudfoot NJ. Promoter proximal splice sites enhance transcription. *Genes Dev* 2002;16:2792–9. <https://doi.org/10.1101/gad.983602>.

Gabler M, Volkmar M, Weinlich S, Herbst A, Dobberthien P, Sklarss S, et al. Trans-splicing of the mod(mdg4) Complex Locus Is Conserved Between the Distantly Related Species *Drosophila melanogaster* and *D. virilis*. *Genetics* 2005;169:723–36. <https://doi.org/10.1534/genetics.103.020842>.

Gaertner B, Johnston J, Chen K, Wallaschek N, Paulson A, Garruss AS, et al. Poised RNA Polymerase II Changes over Developmental Time and Prepares Genes for Future Expression. *Cell Reports* 2012;2:1670–83. <https://doi.org/10.1016/j.celrep.2012.11.024>.

Gajan A, Barnes VL, Liu M, Saha N, Pile LA. The histone demethylase dKDM5/LID interacts with the SIN3 histone deacetylase complex and shares functional similarities with SIN3. *Epigenetics & Chromatin* 2016;9:4. <https://doi.org/10.1186/s13072-016-0053-9>.

- Ganapathi M, Srivastava P, Sutar S, Kumar K, Dasgupta D, Pal Singh G, et al. Comparative analysis of chromatin landscape in regulatory regions of human housekeeping and tissue specific genes. *BMC Bioinformatics* 2005;6:126. <https://doi.org/10.1186/1471-2105-6-126>.
- Garcia JF, Dumesic PA, Hartley PD, El-Samad H, Madhani HD. Combinatorial, site-specific requirement for heterochromatic silencing factors in the elimination of nucleosome-free regions. *Genes & Development* 2010;24:1758–71. <https://doi.org/10.1101/gad.1946410>.
- Gates LA, Foulds CE, O'Malley BW. Histone Marks in the 'Driver's Seat': Functional Roles in Steering the Transcription Cycle. *Trends in Biochemical Sciences* 2017;42:977–89. <https://doi.org/10.1016/j.tibs.2017.10.004>.
- Georgiev PG, Gerasimova TI. Novel genes influencing the expression of the yellow locus and *mdg4* (*gypsy*) in *Drosophila melanogaster*. *Molec Gen Genet* 1989;220:121–6. <https://doi.org/10.1007/BF00260865>.
- Gerasimova TI, Corces VG. Polycomb and Trithorax Group Proteins Mediate the Function of a Chromatin Insulator. *Cell* 1998;92:511–21. [https://doi.org/10.1016/S0092-8674\(00\)80944-7](https://doi.org/10.1016/S0092-8674(00)80944-7).
- Gerasimova TI, Gdula DA, Gerasimov DV, Simonova O, Corces VG. A drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell* 1995;82:587–97. [https://doi.org/10.1016/0092-8674\(95\)90031-4](https://doi.org/10.1016/0092-8674(95)90031-4).
- Ghosh SK, Jost D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. *PLoS Comput Biol* 2018;14:e1006159. <https://doi.org/10.1371/journal.pcbi.1006159>.
- Goldberg AD, Banaszynski LA, Noh K-M, Lewis PW, Elsaesser SJ, Stadler S, et al. Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. *Cell* 2010;140:678–91. <https://doi.org/10.1016/j.cell.2010.01.003>.
- Gorgoni B. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: A developmental perspective. *Briefings in Functional Genomics and Proteomics* 2004;3:125–41. <https://doi.org/10.1093/bfgp/3.2.125>.
- Gu B, Lee M. Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci* 2013;3:39. <https://doi.org/10.1186/2045-3701-3-39>.
- Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell* 2007;130:77–88. <https://doi.org/10.1016/j.cell.2007.05.042>.
- Guillemette B, Bataille AR, Gévry N, Adam M, Blanchett

e M, Robert F, et al. Variant Histone H2A.Z Is Globally Localized to the Promoters of Inactive Yeast Genes and Regulates Nucleosome Positioning. *PLoS Biol* 2005;3:e384. <https://doi.org/10.1371/journal.pbio.0030384>.

Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* 2015;162:900–10. <https://doi.org/10.1016/j.cell.2015.07.038>.

Gurudatta BV, Corces VG. Chromatin insulators: lessons from the fly. *Briefings in Functional Genomics and Proteomics* 2009;8:276–82. <https://doi.org/10.1093/bfpg/elp032>.
Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* 2018;19:621–37. <https://doi.org/10.1038/s41580-018-0028-8>.

Hall AW, Battenhouse AM, Shivram H, Morris AR, Cowperthwaite MC, Shpak M, et al. Bivalent Chromatin Domains in Glioblastoma Reveal a Subtype-Specific Signature of Glioma Stem Cells. *Cancer Res* 2018;78:2463–74. <https://doi.org/10.1158/0008-5472.CAN-17-1724>.

Hall IM. Establishment and Maintenance of a Heterochromatin Domain. *Science* 2002;297:2232–7. <https://doi.org/10.1126/science.1076466>.

Hammond CM. Histone chaperone networks shaping chromatin function 2017:18.
Hammond MP, Laird CD. Chromosome structure and DNA replication in nurse and follicle cells of *Drosophila melanogaster*. *Chromosoma* 1985;91:267–78. <https://doi.org/10.1007/BF00328222>.

Harvey AJ, Bidwai AP, Miller LK. Doom, a product of the *Drosophila* mod(mdg4) gene, induces apoptosis and binds to baculovirus inhibitor-of-apoptosis proteins. *Mol Cell Biol* 1997;17:2835–43. <https://doi.org/10.1128/MCB.17.5.2835>.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 2007;39:311–8. <https://doi.org/10.1038/ng1966>.

Henikoff S, Henikoff JG, Kaya-Okur HS, Ahmad K. Efficient chromatin accessibility mapping *in situ* by nucleosome-tethered tagmentation. *Genomics*; 2020. <https://doi.org/10.1101/2020.04.15.043083>.

Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Research* 2008;19:460–9. <https://doi.org/10.1101/gr.087619.108>.

- Henry KW. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes & Development* 2003;17:2648–63. <https://doi.org/10.1101/gad.1144003>.
- Hödl M, Basler K. Transcription in the Absence of Histone H3.2 and H3K4 Methylation. *Current Biology* 2012;22:2253–7. <https://doi.org/10.1016/j.cub.2012.10.008>.
- Hoffmann NA, Jakobi AJ, Moreno-Morcillo M, Glatt S, Kosinski J, Hagen WJH, et al. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* 2015;528:231–6. <https://doi.org/10.1038/nature16143>.
- Horard B, Loppin B. Fécondation: Le noyau spermatique déverrouillé par une thiorédoxine ultraspécialisée. *Med Sci (Paris)* 2017;33:585–7. <https://doi.org/10.1051/medsci/20173306009>.
- Horner VL, Wolfner MF. Transitioning from egg to embryo: Triggers and mechanisms of egg activation. *Dev Dyn* 2008;237:527–44. <https://doi.org/10.1002/dvdy.21454>.
- Hornung G, Bar-Ziv R, Rosin D, Tokuriki N, Tawfik DS, Oren M, et al. Noise-mean relationship in mutated promoters. *Genome Research* 2012;22:2409–17. <https://doi.org/10.1101/gr.139378.112>.
- Howe FS, Fischl H, Murray SC, Mellor J. Is H3K4me3 instructive for transcription activation? *BioEssays* 2017;39:e201600095. <https://doi.org/10.1002/bies.201600095>.
- Hsin J-P, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & Development* 2012;26:2119–37. <https://doi.org/10.1101/gad.200303.112>.
- Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* 2017;169:216-228.e19. <https://doi.org/10.1016/j.cell.2017.03.024>.
- Imbeault M, Helleboid P-Y, Trono D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* 2017;543:550–4. <https://doi.org/10.1038/nature21683>.
- Iovino N, Ciabrelli F, Cavalli G. PRC2 controls *Drosophila* oocyte cell fate by repressing cell cycle genes. *Dev Cell* 2013;26:431–9. <https://doi.org/10.1016/j.devcel.2013.06.021>.
- Ito H, Sato K, Koganezawa M, Ote M, Matsumoto K, Hama C, et al. Fruitless Recruits Two Antagonistic Chromatin Factors to Establish Single-Neuron Sexual Dimorphism. *Cell* 2012;149:1327–38. <https://doi.org/10.1016/j.cell.2012.04.025>.

- Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, et al. PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing. *Molecular Cell* 2007;28:823–37. <https://doi.org/10.1016/j.molcel.2007.11.012>.
- J. Svensson M, Larsson J. Thioredoxin-2 affects lifespan and oxidative stress in *Drosophila*: Thioredoxins in *Drosophila*. *Hereditas* 2007;144:25–32. <https://doi.org/10.1111/j.2007.0018-0661.01990.x>.
- Jacobs SA, Taverna SD, Zhang Y, Briggs SD, Li J, Eissenberg JC, et al. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J* 2001;20:5232–41. <https://doi.org/10.1093/emboj/20.18.5232>.
- Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* 2009;10:161–72. <https://doi.org/10.1038/nrg2522>.
- Jiang Z, Zhang B. On the role of transcription in positioning nucleosomes. *PLoS Comput Biol* 2021;17:e1008556. <https://doi.org/10.1371/journal.pcbi.1008556>.
- Jost D, Carrivain P, Cavalli G, Vaillant C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Research* 2014;42:9553–61. <https://doi.org/10.1093/nar/gku698>.
- Kadamb R, Mittal S, Bansal N, Batra H, Saluja D. Sin3: Insight into its transcription regulatory functions. *European Journal of Cell Biology* 2013;92:237–46. <https://doi.org/10.1016/j.ejcb.2013.09.001>.
- Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* 2015;1:e1500447. <https://doi.org/10.1126/sciadv.1500447>.
- Kadoch C, Williams RT, Calarco JP, Miller EL, Weber CM, Braun SMG, et al. Dynamics of BAF–Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat Genet* 2017;49:213–22. <https://doi.org/10.1038/ng.3734>.
- Kahn TG, Dorafshan E, Schultheis D, Zare A, Stenberg P, Reim I, et al. Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. *Nucleic Acids Res* 2016;gkw701. <https://doi.org/10.1093/nar/gkw701>.
- Kahn TG, Schwartz YB, Dellino GI, Pirrotta V. Polycomb Complexes and the Propagation of the Methylation Mark at the *Drosophila* Ubx Gene. *Journal of Biological Chemistry* 2006;281:29064–75. <https://doi.org/10.1074/jbc.M605430200>.

Kao C-F. Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. *Genes & Development* 2004;18:184–95. <https://doi.org/10.1101/gad.1149604>.

Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, et al. DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric Transcripts in mESCs. *Cell Stem Cell* 2011;8:676–87. <https://doi.org/10.1016/j.stem.2011.04.004>.

Kassis JA, Brown JL. Polycomb Group Response Elements in *Drosophila* and Vertebrates. *Advances in Genetics*, vol. 81, Elsevier; 2013, p. 83–118. <https://doi.org/10.1016/B978-0-12-407677-8.00003-8>.

Kassis JA, Kennison JA, Tamkun JW. Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics* 2017;206:1699–725. <https://doi.org/10.1534/genetics.115.185116>.

Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, et al. 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* 2017;170:367–381.e20. <https://doi.org/10.1016/j.cell.2017.06.029>.

Keller C, Adaixo R, Stunnenberg R, Woolcock KJ, Hiller S, Bühler M. HP1Swi6 Mediates the Recognition and Destruction of Heterochromatic RNA Transcripts. *Molecular Cell* 2012;47:215–27. <https://doi.org/10.1016/j.molcel.2012.05.009>.

King, RC. Ovarian development in *Drosophila melanogaster*. New York. Academic.; 1970. Kingston RE, Tamkun JW. Transcriptional Regulation by Trithorax-Group Proteins. *Cold Spring Harbor Perspectives in Biology* 2014;6:a019349–a019349. <https://doi.org/10.1101/cshperspect.a019349>.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, et al. The *Drosophila* HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters. *Cell* 2009;138:1137–49. <https://doi.org/10.1016/j.cell.2009.07.014>.

Klymenko T, Müller J. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep* 2004;5:373–7. <https://doi.org/10.1038/sj.embor.7400111>.

Kochetov AV, Ponomarenko MP, Frolov AS, Kisselev LL, Kolchanov NA. Prediction of eukaryotic mRNA translational properties. *Bioinformatics* 1999;15:704–12. <https://doi.org/10.1093/bioinformatics/15.7.704>.

Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. *Molecular Cell* 2017;67:411–422.e4. <https://doi.org/10.1016/j.molcel.2017.06.027>.

Kronja I, Whitfield ZJ, Yuan B, Dzek K, Kirkpatrick J, Krijgsveld J, et al. Quantitative proteomics reveals the dynamics of protein changes during *Drosophila* oocyte maturation and the oocyte-to-embryo transition. *Proc Natl Acad Sci USA* 2014;111:16023–8. <https://doi.org/10.1073/pnas.1418657111>.

Kubik S, Bruzzone MJ, Jacquet P, Falcone J-L, Rougemont J, Shore D. Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. *Molecular Cell* 2015;60:422–34. <https://doi.org/10.1016/j.molcel.2015.10.002>.

Kusch T, Mei A, Nguyen C. Histone H3 lysine 4 trimethylation regulates cotranscriptional H2A variant exchange by Tip60 complexes to maximize gene expression. *Proc Natl Acad Sci USA* 2014;111:4850–5. <https://doi.org/10.1073/pnas.1320337111>.

Kutach AK, Kadonaga JT. The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in *Drosophila* Core Promoters. *Mol Cell Biol* 2000;20:4754–64. <https://doi.org/10.1128/MCB.20.13.4754-4764.2000>.

Kuzmichev A. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes & Development* 2002;16:2893–905. <https://doi.org/10.1101/gad.1035902>.

Kyrchanova O, Georgiev P. Chromatin insulators and long-distance interactions in *Drosophila*. *FEBS Letters* 2014;588:8–14. <https://doi.org/10.1016/j.febslet.2013.10.039>.

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 2017;547:236–40. <https://doi.org/10.1038/nature22822>.

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, et al. H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. *Cell* 2013;152:1021–36. <https://doi.org/10.1016/j.cell.2013.01.052>.

Lavarone E, Barbieri CM, Pasini D. Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat Commun* 2019;10:1679. <https://doi.org/10.1038/s41467-019-09624-w>.

Lee C-K, Shibata Y, Rao B, Strahl BD, Lieb JD. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* 2004;36:900–5. <https://doi.org/10.1038/ng1400>.

Lee J, Hyeon DY, Hwang D. Single-cell multiomics: technologies and data analysis methods. *Exp Mol Med* 2020;52:1428–42. <https://doi.org/10.1038/s12276-020-0420-2>.

- Lepesant JMJ, Iampietro C, Galeota E, Augé B, Aguirrenbengoa M, Mercé C, et al. A dual role of dLsd1 in oogenesis: regulating developmental genes and repressing transposons. *Nucleic Acids Research* 2020;48:1206–24. <https://doi.org/10.1093/nar/gkz1142>.
- Lesch BJ, Dokshin GA, Young RA, McCarrey JR, Page DC. A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. *Proceedings of the National Academy of Sciences* 2013;110:16061–6. <https://doi.org/10.1073/pnas.1315204110>.
- Lhoumaud P, Hennion M, Gamot A, Cuddapah S, Queille S, Liang J, et al. Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. *The EMBO Journal* 2014;33:1599–613. <https://doi.org/10.15252/embj.201385965>.
- Li B, Gogol M, Carey M, Lee D, Seidel C, Workman JL. Combined Action of PHD and Chromo Domains Directs the Rpd3S HDAC to Transcribed Chromatin. *Science* 2007;316:1050–4. <https://doi.org/10.1126/science.1139004>.
- Li H, Liefke R, Jiang J, Kurland JV, Tian W, Deng P, et al. Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* 2017;549:287–91. <https://doi.org/10.1038/nature23881>.
- Li L, Greer C, Eisenman RN, Secombe J. Essential Functions of the Histone Demethylase Lid. *PLoS Genetics* 2010;6:e1001221. <https://doi.org/10.1371/journal.pgen.1001221>.
- Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes & Development* 2013;27:2380–96. <https://doi.org/10.1101/gad.229328.113>.
- Lieberman-Aiden E, van Berkum NL, Williams L, Imapkaev M, Ragozcy T, Telling A, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* 2009;326:289–93. <https://doi.org/10.1126/science.1181369>.
- Lim J, Lee M, Son A, Chang H, Kim VN. mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes Dev* 2016;30:1671–82. <https://doi.org/10.1101/gad.284802.116>.
- Liu X, Greer C, Secombe J. KDM5 Interacts with Foxo to Modulate Cellular Levels of Oxidative Stress. *PLoS Genetics* 2014;10:e1004676. <https://doi.org/10.1371/journal.pgen.1004676>.
- Liu X, Secombe J. The Histone Demethylase KDM5 Activates Gene Expression by Recognizing Chromatin Context through Its PHD Reader Motif. *Cell Reports* 2015;13:2219–31. <https://doi.org/10.1016/j.celrep.2015.11.007>.

Lloret-Llinares M, Pérez-Lluch S, Rossell D, Morán T, Ponsa-Cobas J, Auer H, et al. dKDM5/LID regulates H3K4me3 dynamics at the transcription-start site (TSS) of actively transcribed developmental genes. *Nucleic Acids Res* 2012;40:9493–505. <https://doi.org/10.1093/nar/gks773>.

Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, et al. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* 1990;5:1743–53.

Lopes I, Altab G, Raina P, de Magalhães JP. Gene Size Matters: An Analysis of Gene Length in the Human Genome. *Front Genet* 2021;12:559998. <https://doi.org/10.3389/fgene.2021.559998>.

Loppin B, Dubruille R, Horard B. The intimate genetics of *Drosophila* fertilization. *Open Biol* 2015;5:150076. <https://doi.org/10.1098/rsob.150076>.

Loubiere V, Delest A, Thomas A, Bonev B, Schuettengruber B, Sati S, et al. Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development. *Nat Genet* 2016;48:1436–42. <https://doi.org/10.1038/ng.3671>.

Louhichi A, Fourati A, Rebaï A. IGD: A resource for intronless genes in the human genome. *Gene* 2011;488:35–40. <https://doi.org/10.1016/j.gene.2011.08.013>.

Lu BY, Emtage PC, Duyf BJ, Hilliker AJ, Eissenberg JC. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. *Genetics* 2000;155:699–708.

Lucas T, Ferraro T, Roelens B, De Las Heras Chanes J, Walczak AM, Coppey M, et al. Live Imaging of Bicoid-Dependent Transcription in *Drosophila* Embryos. *Current Biology* 2013;23:2135–9. <https://doi.org/10.1016/j.cub.2013.08.053>.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 1997;389:251–60. <https://doi.org/10.1038/38444>.

Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* 2015;161:1012–25. <https://doi.org/10.1016/j.cell.2015.04.004>.

Luzhin AV, Flyamer IM, Khrameeva EE, Ulianov SV, Razin SV, Gavrilov AA. Quantitative differences in TAD border strength underly the TAD hierarchy in *Drosophila* chromosomes. *J Cell Biochem* 2019;120:4494–503. <https://doi.org/10.1002/jcb.27737>.

Lynch MD, Smith AJH, De Gobbi M, Flenley M, Hughes JR, Vernimmen D, et al. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment: An interspecies analysis of chromatin bivalency. *The EMBO Journal* 2012;31:317–29. <https://doi.org/10.1038/emboj.2011.399>.

Mahowald AP, Tiefert M. Fine structural changes in the *Drosophila* oocyte nucleus during a short period of RNA synthesis: An autoradiographic and ultrastructural study of RNA synthesis in the oocyte nucleus of *Drosophila*. *W Roux' Archiv f Entwicklungsmechanik* 1970;165:8–25. <https://doi.org/10.1007/BF00576994>.

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, et al. Specialized piRNA Pathways Act in Germline and Somatic Tissues of the *Drosophila* Ovary. *Cell* 2009;137:522–35. <https://doi.org/10.1016/j.cell.2009.03.040>.

Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 2003;4:223. <https://doi.org/10.1186/gb-2003-4-7-223>.

Margaritis T, Oreal V, Brabers N, Maestroni L, Vitaliano-Prunier A, Benschop JJ, et al. Two Distinct Repressive Mechanisms for Histone 3 Lysine 4 Methylation through Promoting 3'-End Antisense Transcription. *PLoS Genet* 2012;8:e1002952. <https://doi.org/10.1371/journal.pgen.1002952>.

Martire S, Banaszynski LA. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol* 2020;21:522–41. <https://doi.org/10.1038/s41580-020-0262-8>.

Marullo F, Cesarini E, Antonelli L, Gregoretti F, Oliva G, Lanzaolo C. Nucleoplasmic Lamin A/C and Polycomb group of proteins: An evolutionarily conserved interplay. *Nucleus* 2016;7:103–11. <https://doi.org/10.1080/19491034.2016.1157675>.

Matharu N, Ahituv N. Minor Loops in Major Folds: Enhancer–Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. *PLoS Genet* 2015;11:e1005640. <https://doi.org/10.1371/journal.pgen.1005640>.

Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* 2010;464:927–31. <https://doi.org/10.1038/nature08858>.

Matsui T, Segall J, Weil PA, Roeder RG. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 1980;255:11992–6.

- Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, et al. Nucleosome organization in the *Drosophila* genome. *Nature* 2008;453:358–62. <https://doi.org/10.1038/nature06929>.
- Mayer A, Heidemann M, Lidschreiber M, Schrieck A, Sun M, Hintermair C, et al. CTD Tyrosine Phosphorylation Impairs Termination Factor Recruitment to RNA Polymerase II. *Science* 2012;336:1723–5. <https://doi.org/10.1126/science.1219651>.
- Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* 2010;17:1272–8. <https://doi.org/10.1038/nsmb.1903>.
- McClelland S, Shrivastava R, Medh JD. Regulation of Translational Efficiency by Disparate 5' -UTRs of PPAR γ Splice Variants. *PPAR Research* 2009;2009:1–8. <https://doi.org/10.1155/2009/193413>.
- McCracken S., Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, et al. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes & Development* 1997;11:3306–18. <https://doi.org/10.1101/gad.11.24.3306>.
- McCracken Susan, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, et al. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 1997;385:357–61. <https://doi.org/10.1038/385357a0>.
- McDonel P, Costello I, Hendrich B. Keeping things quiet: Roles of NuRD and Sin3 co-repressor complexes during mammalian development. *The International Journal of Biochemistry & Cell Biology* 2009;41:108–16. <https://doi.org/10.1016/j.biocel.2008.07.022>.
- McLaughlin JM, Bratu DP. *Drosophila melanogaster* Oogenesis: An Overview. In: Bratu DP, McNeil GP, editors. *Drosophila Oogenesis*, vol. 1328, New York, NY: Springer New York; 2015, p. 1–20. https://doi.org/10.1007/978-1-4939-2851-4_1.
- Melnikova L, Elizar'ev P, Erokhin M, Molodina V, Chetverina D, Kostyuchenko M, et al. The same domain of Su(Hw) is required for enhancer blocking and direct promoter repression. *Sci Rep* 2019;9:5314. <https://doi.org/10.1038/s41598-019-41761-6>.
- Melnikova L, Juge F, Gruzdeva N, Mazur A, Cavalli G, Georgiev P. Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proceedings of the National Academy of Sciences* 2004;101:14806–11. <https://doi.org/10.1073/pnas.0403959101>.
- Melnikova L, Kostyuchenko M, Molodina V, Parshikov A, Georgiev P, Golovnin A. Multiple interactions are involved in a highly specific association of the Mod(mdg4)-67.2

isoform with the Su(Hw) sites in *Drosophila*. *Open Biol* 2017;7:170150.
<https://doi.org/10.1098/rsob.170150>.

Melnikova LS, Georgiev PG, Golovnin AK. The Functions and Mechanisms of Action of Insulators in the Genomes of Higher Eukaryotes. *Acta Naturae* 2020;12:15–33.
<https://doi.org/10.32607/actanaturae.11144>.

Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, et al. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. *PLoS Genet* 2010;6:e1001244.
<https://doi.org/10.1371/journal.pgen.1001244>.

Messmer S, Franke A, Paro R. Analysis of the functional role of the Polycomb chromo domain in *Drosophila melanogaster*. *Genes & Development* 1992;6:1241–54.
<https://doi.org/10.1101/gad.6.7.1241>.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007;448:553–60.
<https://doi.org/10.1038/nature06008>.

Miller D, Brinkworth M, Iles D. Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. *REPRODUCTION* 2010;139:287–301. <https://doi.org/10.1530/REP-09-0281>.

Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* 2008;10:483–8. <https://doi.org/10.1038/ncb1712>.

Mito Y, Henikoff JG, Henikoff S. Histone Replacement Marks the Boundaries of cis-Regulatory Domains. *Science* 2007;315:1408–11. <https://doi.org/10.1126/science.1134004>.
Mohan M, Herz H-M, Smith ER, Zhang Y, Jackson J, Washburn MP, et al. The COMPASS Family of H3K4 Methylases in *Drosophila*. *Mol Cell Biol* 2011;31:4310–8.
<https://doi.org/10.1128/MCB.06092-11>.

Mohrmann L, Langenberg K, Krijgsveld J, Kal AJ, Heck AJR, Verrijzer CP. Differential Targeting of Two Distinct SWI/SNF-Related *Drosophila* Chromatin-Remodeling Complexes. *MCB* 2004;24:3077–88. <https://doi.org/10.1128/MCB.24.8.3077-3088.2004>.

Moshkin YM, Kan TW, Goodfellow H, Bezstarosti K, Maeda RK, Pilyugin M, et al. Histone Chaperones ASF1 and NAP1 Differentially Modulate Removal of Active Histone Marks by LID-RPD3 Complexes during NOTCH Silencing. *Molecular Cell* 2009;35:782–93.
<https://doi.org/10.1016/j.molcel.2009.07.020>.

Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone Methyltransferase Activity of a *Drosophila* Polycomb Group Repressor Complex. *Cell* 2002;111:197–208. [https://doi.org/10.1016/S0092-8674\(02\)00976-5](https://doi.org/10.1016/S0092-8674(02)00976-5).

Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013;502:59–64. <https://doi.org/10.1038/nature12593>.

Navarro-Costa P, McCarthy A, Prudêncio P, Greer C, Guilgur LG, Becker JD, et al. Early programming of the oocyte epigenome temporally controls late prophase I transcription and chromatin remodelling. *Nat Commun* 2016;7:12331. <https://doi.org/10.1038/ncomms12331>.

Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, et al. A Comprehensive Map of Insulator Elements for the *Drosophila* Genome. *PLoS Genet* 2010;6:e1000814. <https://doi.org/10.1371/journal.pgen.1000814>.

Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 2017;543:72–7. <https://doi.org/10.1038/nature21373>.

Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 2010;463:457–63. <https://doi.org/10.1038/nature08909>.

Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, et al. Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* 2015;161:526–40. <https://doi.org/10.1016/j.cell.2015.03.027>.

Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 2017;169:930-944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>.

Nordman J, Orr-Weaver TL. Regulation of DNA replication during development. *Development* 2012;139:455–64. <https://doi.org/10.1242/dev.061838>.

Nott A. A quantitative analysis of intron effects on mammalian gene expression. *RNA* 2003;9:607–17. <https://doi.org/10.1261/rna.5250403>.

van Oevelen C, Wang J, Asp P, Yan Q, Kaelin WG, Kluger Y, et al. A Role for Mammalian Sin3 in Permanent Gene Silencing. *Molecular Cell* 2008;32:359–70. <https://doi.org/10.1016/j.molcel.2008.10.015>.

Ogienko AA, Fedorova SA, Baricheva EM. Basic aspects of ovarian development in *Drosophila melanogaster*. *Russ J Genet* 2007;43:1120–34. <https://doi.org/10.1134/S1022795407100055>.

Ohler U, Liao G, Niemann H, Rubin GM. Computational analysis of core promoters in the *Drosophila* genome 2002;12.

Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II. *Genes & Development* 1996;10:2657–83. <https://doi.org/10.1101/gad.10.21.2657>.

Orsi GA, Kasinathan S, Hughes KT, Saminadin-Peter S, Henikoff S, Ahmad K. High-resolution mapping defines the cooperative architecture of Polycomb response elements. *Genome Research* 2014;24:809–20. <https://doi.org/10.1101/gr.163642.113>.

Ozata DM, Gainetdinov I, Zoch A, O’Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet* 2019;20:89–108. <https://doi.org/10.1038/s41576-018-0073-3>.

Özdemir, Gambetta. The Role of Insulation in Patterning Gene Expression. *Genes* 2019;10:767. <https://doi.org/10.3390/genes10100767>.

Pai C-Y, Lei EP, Ghosh D, Corces VG. The Centrosomal Protein CP190 Is a Component of the gypsy Chromatin Insulator. *Molecular Cell* 2004;16:737–48. <https://doi.org/10.1016/j.molcel.2004.11.004>.

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, et al. Whole-Genome Analysis of Histone H3 Lysine 4 and Lysine 27 Methylation in Human Embryonic Stem Cells. *Cell Stem Cell* 2007;1:299–312. <https://doi.org/10.1016/j.stem.2007.08.003>.

Parisi M, Nuttall R, Edwards P, Minor J, Naiman D, Lü J, et al. A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. *Genome Biology* 2004;18.

Pascucci T, Perrino J, Mahowald AP, Waring GL. Eggshell Assembly in *Drosophila*: Processing and Localization of Vitelline Membrane and Chorion Proteins. *Developmental Biology* 1996;177:590–8. <https://doi.org/10.1006/dbio.1996.0188>.

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, et al. Histone H2B Monoubiquitination Functions Cooperatively with FACT to Regulate Elongation by RNA Polymerase II. *Cell* 2006;125:703–17. <https://doi.org/10.1016/j.cell.2006.04.029>.

Pengelly AR, Copur O, Jackle H, Herzig A, Muller J. A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb. *Science* 2013;339:698–9. <https://doi.org/10.1126/science.1231382>.

Pengelly AR, Kalb R, Finkl K, Müller J. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev* 2015;29:1487–92. <https://doi.org/10.1101/gad.265439.115>.

Perino M, van Mierlo G, Karemaker ID, van Genesen S, Vermeulen M, Marks H, et al. MTF2 recruits Polycomb Repressive Complex 2 by helical-shape-selective DNA binding. *Nat Genet* 2018;50:1002–10. <https://doi.org/10.1038/s41588-018-0134-8>.

Petrova B, Liu K, Tian C, Kitaoka M, Freinkman E, Yang J, et al. Dynamic redox balance directs the oocyte-to-embryo transition via developmentally controlled reactive cysteine changes. *Proc Natl Acad Sci USA* 2018;115:E7978–86. <https://doi.org/10.1073/pnas.1807918115>.

Pherson M, Misulovin Z, Gause M, Mihindikulasuriya K, Swain A, Dorsett D. Polycomb repressive complex 1 modifies transcription of active genes. *Sci Adv* 2017;3:e1700944. <https://doi.org/10.1126/sciadv.1700944>.

Phillips-Cremins JE, Sauria MEG, Sanyal A, Gerasimova TI, Lajoie BR, Bell JSK, et al. Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell* 2013;153:1281–95. <https://doi.org/10.1016/j.cell.2013.04.053>.

Plath K. Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science* 2003;300:131–5. <https://doi.org/10.1126/science.1084274>.

Prudêncio P, Guilgur LG, Sobral J, Becker JD, Martinho RG, Navarro-Costa P. The Trithorax group protein dMLL3/4 instructs the assembly of the zygotic genome at fertilization. *EMBO Rep* 2018;19. <https://doi.org/10.15252/embr.201845728>.

Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, et al. Human tRNA genes function as chromatin insulators: Human tRNA genes function as chromatin insulators. *The EMBO Journal* 2012;31:330–50. <https://doi.org/10.1038/emboj.2011.406>.

Rach EA, Winter DR, Benjamin AM, Corcoran DL, Ni T, Zhu J, et al. Transcription Initiation Patterns Indicate Divergent Strategies for Gene Regulation at the Chromatin Level. *PLoS Genet* 2011;7:e1001274. <https://doi.org/10.1371/journal.pgen.1001274>.

Racko D, Benedetti F, Goundaroulis D, Stasiak A. Chromatin Loop Extrusion and Chromatin Unknotting. *Polymers* 2018;10:1126. <https://doi.org/10.3390/polym10101126>.

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biol* 2006;4:e309. <https://doi.org/10.1371/journal.pbio.0040309>.

- Ramakrishnan S, Pokhrel S, Palani S, Pflueger C, Parnell TJ, Cairns BR, et al. Counteracting H3K4 methylation modulators Set1 and Jhd2 co-regulate chromatin dynamics and gene transcription. *Nat Commun* 2016;7:11949. <https://doi.org/10.1038/ncomms11949>.
- Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun* 2018;9:189. <https://doi.org/10.1038/s41467-017-02525-w>.
- Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, et al. piRNA Production Requires Heterochromatin Formation in *Drosophila*. *Current Biology* 2011;21:1373–9. <https://doi.org/10.1016/j.cub.2011.06.057>.
- Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, et al. Cohesin Loss Eliminates All Loop Domains. *Cell* 2017;171:305-320.e24. <https://doi.org/10.1016/j.cell.2017.09.026>.
- Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 2014;159:1665–80. <https://doi.org/10.1016/j.cell.2014.11.021>.
- Ray P, De S, Mitra A, Bezstarosti K, Demmers JAA, Pfeifer K, et al. Combgap contributes to recruitment of Polycomb group proteins in *Drosophila*. *Proc Natl Acad Sci USA* 2016;113:3826–31. <https://doi.org/10.1073/pnas.1520926113>.
- Reddington JP, Perricone SM, Nestor CE, Reichmann J, Youngson NA, Suzuki M, et al. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol* 2013;14:R25. <https://doi.org/10.1186/gb-2013-14-3-r25>.
- Rennie S, Dalby M, van Duin L, Andersson R. Transcriptional decomposition reveals active chromatin architectures and cell specific regulatory interactions. *Nat Commun* 2018;9:487. <https://doi.org/10.1038/s41467-017-02798-1>.
- Ringrose L, Paro R. Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins. *Annu Rev Genet* 2004;38:413–43. <https://doi.org/10.1146/annurev.genet.38.072902.091907>.
- Rodriguez J, Munoz M, Vives L, Frangou CG, Groudine M, Peinado MA. Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells. *Proceedings of the National Academy of Sciences* 2008;105:19809–14. <https://doi.org/10.1073/pnas.0810133105>.
- Roelens B, Clémot M, Leroux-Coyau M, Klapholz B, Dostatni N. Maintenance of Heterochromatin by the Large Subunit of the CAF-1 Replication-Coupled Histone Chaperone

Requires Its Interaction with HP1a Through a Conserved Motif. *Genetics* 2017;205:125–37. <https://doi.org/10.1534/genetics.116.190785>.

Royzman I, Orr-Weaver TL. S phase and differential DNA replication during *Drosophila* oogenesis: Endo cell cycle in *Drosophila* oogenesis. *Genes to Cells* 1998;3:767–76. <https://doi.org/10.1046/j.1365-2443.1998.00232.x>.

Sabari BR, Dall’Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 2018;361:eaar3958. <https://doi.org/10.1126/science.aar3958>.

Saha N, Liu M, Gajan A, Pile LA. Genome-wide studies reveal novel and distinct biological pathways regulated by SIN3 isoforms. *BMC Genomics* 2016;17:111. <https://doi.org/10.1186/s12864-016-2428-5>.

Sahu SC, Swanson KA, Kang RS, Huang K, Brubaker K, Ratcliff K, et al. Conserved Themes in Target Recognition by the PAH1 and PAH2 Domains of the Sin3 Transcriptional Corepressor. *Journal of Molecular Biology* 2008;375:1444–56. <https://doi.org/10.1016/j.jmb.2007.11.079>.

Sakai A, Schwartz BE, Goldstein S, Ahmad K. Transcriptional and Developmental Functions of the H3.3 Histone Variant in *Drosophila*. *Current Biology* 2009;19:1816–20. <https://doi.org/10.1016/j.cub.2009.09.021>.

Salz HK, Flickinger TW, Mittendorf E, Pellicena-Palle A, Petschek JP, Albrecht EB. The *Drosophila* maternal effect locus deadhead encodes a thioredoxin homolog required for female meiosis and early embryonic development. *Genetics* 1994;136:1075–86. <https://doi.org/10.1093/genetics/136.3.1075>.

Santaguida S, Musacchio A. The life and miracles of kinetochores. *EMBO J* 2009;28:2511–31. <https://doi.org/10.1038/emboj.2009.173>.

Sanz AB, García R, Rodríguez-Peña JM, Nombela C, Arroyo J. Cooperation between SAGA and SWI/SNF complexes is required for efficient transcriptional responses regulated by the yeast MAPK Slt2. *Nucleic Acids Res* 2016:gkw324. <https://doi.org/10.1093/nar/gkw324>.

Savitsky M, Kim M, Kravchuk O, Schwartz YB. Distinct Roles of Chromatin Insulator Proteins in Control of the *Drosophila* Bithorax Complex. *Genetics* 2016;202:601–17. <https://doi.org/10.1534/genetics.115.179309>.

Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, et al. Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. *Molecular Cell* 2011;42:330–41. <https://doi.org/10.1016/j.molcel.2011.03.025>.

Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* 2017;171:34–57. <https://doi.org/10.1016/j.cell.2017.08.002>.

Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, Tolhuis B, et al. Functional Anatomy of Polycomb and Trithorax Chromatin Landscapes in *Drosophila* Embryos. *PLoS Biol* 2009;7:e1000013. <https://doi.org/10.1371/journal.pbio.1000013>.

Schultz DC. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes & Development* 2002;16:919–32. <https://doi.org/10.1101/gad.973302>.

Schwabish MA, Struhl K. The Swi/Snf Complex Is Important for Histone Eviction during Transcriptional Activation and RNA Polymerase II Elongation In Vivo. *Mol Cell Biol* 2007;27:6987–95. <https://doi.org/10.1128/MCB.00717-07>.

Schwabish MA, Struhl K. Evidence for Eviction and Rapid Deposition of Histones upon Transcriptional Elongation by RNA Polymerase II. *Mol Cell Biol* 2004;24:10111–7. <https://doi.org/10.1128/MCB.24.23.10111-10117.2004>.

Schwartz YB, Cavalli G. Three-Dimensional Genome Organization and Function in *Drosophila*. *Genetics* 2017;205:5–24. <https://doi.org/10.1534/genetics.115.185132>.

Schwartz YB, Kahn TG, Nix DA, Li X-Y, Bourgon R, Biggin M, et al. Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nat Genet* 2006;38:700–5. <https://doi.org/10.1038/ng1817>.

Schwartz YB, Kahn TG, Stenberg P, Ohno K, Bourgon R, Pirrotta V. Alternative Epigenetic Chromatin States of Polycomb Target Genes. *PLoS Genet* 2010;6:e1000805. <https://doi.org/10.1371/journal.pgen.1000805>.

Schwartz YB, Linder-Basso D, Kharchenko PV, Tolstorukov MY, Kim M, Li H-B, et al. Nature and function of insulator protein binding sites in the *Drosophila* genome. *Genome Res* 2012;22:2188–98. <https://doi.org/10.1101/gr.138156.112>.

Scott KC, Merrett SL, Willard HF. A Heterochromatin Barrier Partitions the Fission Yeast Centromere into Discrete Chromatin Domains. *Current Biology* 2006;16:119–29. <https://doi.org/10.1016/j.cub.2005.11.065>.

Secombe J, Li L, Carlos L, Eisenman RN. The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Genes & Development* 2007;21:537–51. <https://doi.org/10.1101/gad.1523007>.

Segal E, Widom J. What controls nucleosome positions? *Trends in Genetics* 2009;25:335–43. <https://doi.org/10.1016/j.tig.2009.06.002>.

Sentenac A. Eukaryotic RNA Polymerase. *Critical Reviews in Biochemistry* 1985;18:31–90. <https://doi.org/10.3109/10409238509082539>.

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al. Three-Dimensional Folding and Functional Organization Principles of the *Drosophila* Genome. *Cell* 2012;148:458–72. <https://doi.org/10.1016/j.cell.2012.01.010>.

Shaul O. How introns enhance gene expression. *The International Journal of Biochemistry & Cell Biology* 2017;91:145–55. <https://doi.org/10.1016/j.biocel.2017.06.016>.

Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet* 2014;15:272–86. <https://doi.org/10.1038/nrg3682>.

Silverstein RA, Ekwall K. Sin3: a flexible regulator of global gene expression and genome stability. *Curr Genet* 2005;47:1–17. <https://doi.org/10.1007/s00294-004-0541-5>.

Simon JA, Kingston RE. Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol Cell* 2013;49:808–24. <https://doi.org/10.1016/j.molcel.2013.02.013>.

Smolko AE, Shapiro-Kulnane L, Salz HK. The H3K9 methyltransferase SETDB1 maintains female identity in *Drosophila* germ cells. *Nat Commun* 2018;9:4155. <https://doi.org/10.1038/s41467-018-06697-x>.

Spellman PT, Rubin GM. Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J Biol* 2002;1:5. <https://doi.org/10.1186/1475-4924-1-5>.

Spradling A. *The development of Drosophila melanogaster*. 1993.

Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 2017;544:59–64. <https://doi.org/10.1038/nature21429>.

Stitzel ML, Seydoux G. Regulation of the Oocyte-to-Zygote Transition. *Science* 2007;316:407–8. <https://doi.org/10.1126/science.1138236>.

Strahl BD, Ohba R, Cook RG, Allis CD. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proceedings of the National Academy of Sciences* 1999;96:14967–72. <https://doi.org/10.1073/pnas.96.26.14967>.

Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. *Nature* 2017;547:241–5. <https://doi.org/10.1038/nature22989>.

Svensson MJ, Chen JD, Pirrotta V, Larsson J. The ThioredoxinT and deadhead gene pair encode testis- and ovary-specific thioredoxins in *Drosophila melanogaster*. *Chromosoma* 2003;112:133–43. <https://doi.org/10.1007/s00412-003-0253-5>.

Svensson MJ, Stenberg P, Larsson J. Organization and regulation of sex-specific thioredoxin encoding genes in the genus *Drosophila*. *Dev Genes Evol* 2007;217:639–50. <https://doi.org/10.1007/s00427-007-0175-y>.

Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. *Sci Adv* 2019;5:eaaw1668. <https://doi.org/10.1126/sciadv.aaw1668>.

Szenker E, Lacoste N, Almouzni G. A Developmental Requirement for HIRA-Dependent H3.3 Deposition Revealed at Gastrulation in *Xenopus*. *Cell Reports* 2012;1:730–40. <https://doi.org/10.1016/j.celrep.2012.05.006>.

Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, Orr-Weaver T, et al. SMAUG Is a Major Regulator of Maternal mRNA Destabilization in *Drosophila* and Its Translation Is Activated by the PAN GU Kinase. *Developmental Cell* 2007;12:143–55. <https://doi.org/10.1016/j.devcel.2006.10.005>.

Talbert PB, Henikoff S. Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* 2017;18:115–26. <https://doi.org/10.1038/nrm.2016.148>.

Tie F, Banerjee R, Fu C, Stratton CA, Fang M, Harte PJ. Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *Proc Natl Acad Sci USA* 2016;113:E744–53. <https://doi.org/10.1073/pnas.1515465113>.

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* 2009;136:3131–41. <https://doi.org/10.1242/dev.037127>.

Tirmarche S. Fonctions des thiorédoxines sexuelles et contrôle de l'état rédox des protamines chez la drosophile 2016:157.

Tirmarche S, Kimura S, Dubruille R, Horard B, Loppin B. Unlocking sperm chromatin at fertilization requires a dedicated egg thioredoxin in *Drosophila*. *Nat Commun* 2016;7:13539. <https://doi.org/10.1038/ncomms13539>.

Torné J, Ray-Gallet D, Boyarchuk E, Garnier M, Le Baccon P, Coulon A, et al. Two HIRA-dependent pathways mediate H3.3 de novo deposition and recycling during transcription. *Nat Struct Mol Biol* 2020;27:1057–68. <https://doi.org/10.1038/s41594-020-0492-7>.

Torres-Campana D, Kimura S, Orsi GA, Horard B, Benoit G, Loppin B. The Lid/KDM5 histone demethylase complex activates a critical effector of the oocyte-to-zygote transition. *PLoS Genet* 2020;16:e1008543. <https://doi.org/10.1371/journal.pgen.1008543>.

Ulianov SV, Khrameeva EE, Gavrilov AA, Flyamer IM, Kos P, Mikhaleva EA, et al. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res* 2016;26:70–84. <https://doi.org/10.1101/gr.196006.115>.

Urrutia AO. The Signature of Selection Mediated by Expression on Human Genes. *Genome Research* 2003;13:2260–4. <https://doi.org/10.1101/gr.641103>.

Van Bortle K, Nichols MH, Li L, Ong C-T, Takenaka N, Qin ZS, et al. Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol* 2014;15:R82. <https://doi.org/10.1186/gb-2014-15-5-r82>.

Van Bortle K, Ramos E, Takenaka N, Yang J, Wahi JE, Corces VG. Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. *Genome Res* 2012;22:2176–87. <https://doi.org/10.1101/gr.136788.111>.

Verdel A. RNAi-Mediated Targeting of Heterochromatin by the RITS Complex. *Science* 2004;303:672–6. <https://doi.org/10.1126/science.1093686>.

Vo ngoc L, Wang Y-L, Kassavetis GA, Kadonaga JT. The punctilious RNA polymerase II core promoter. *Genes Dev* 2017;31:1289–301. <https://doi.org/10.1101/gad.303149.117>.

Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, et al. Asymmetrically Modified Nucleosomes. *Cell* 2012;151:181–93. <https://doi.org/10.1016/j.cell.2012.09.002>.

Voigt P, Tee W-W, Reinberg D. A double take on bivalent promoters. *Genes & Development* 2013;27:1318–38. <https://doi.org/10.1101/gad.219626.113>.

Volpe TA. Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* 2002;297:1833–7. <https://doi.org/10.1126/science.1074973>.

Wakimoto BT, Hearn MG. The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of *Drosophila melanogaster*. *Genetics* 1990;125:141–54.

Wang C, Lin H. Roles of piRNAs in transposon and pseudogene regulation of germline mRNAs and lncRNAs. *Genome Biol* 2021;22:27. <https://doi.org/10.1186/s13059-020-02221-x>.

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 2004;431:873–8. <https://doi.org/10.1038/nature02985>.

Wang Q, Sun Q, Czajkowsky DM, Shao Z. Sub-kb Hi-C in *D. melanogaster* reveals conserved characteristics of TADs between insect and mammalian cells. *Nat Commun* 2018;9:188. <https://doi.org/10.1038/s41467-017-02526-9>.

Weber CM, Hafner A, Kirkland JG, Braun SMG, Stanton BZ, Boettiger AN, et al. mSWI/SNF promotes Polycomb repression both directly and through genome-wide redistribution. *Nat Struct Mol Biol* 2021;28:501–11. <https://doi.org/10.1038/s41594-021-00604-7>.

Weinberg DE, Shah P, Eichhorn SW, Hussmann JA, Plotkin JB, Bartel DP. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Reports* 2016;14:1787–99. <https://doi.org/10.1016/j.celrep.2016.01.043>.

Weiner A, Chen HV, Liu CL, Rahat A, Klien A, Soares L, et al. Systematic Dissection of Roles for Chromatin Regulators in a Yeast Stress Response. *PLoS Biol* 2012;10:e1001369. <https://doi.org/10.1371/journal.pbio.1001369>.

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 2011;473:343–8. <https://doi.org/10.1038/nature10066>.

Wolf G, Yang P, Füchtbauer AC, Füchtbauer E-M, Silva AM, Park C, et al. The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. *Genes Dev* 2015;29:538–54. <https://doi.org/10.1101/gad.252767.114>.

Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J* 2017;36:3573–99. <https://doi.org/10.15252/embj.201798004>.

Xi Y, Yao J, Chen R, Li W, He X. Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. *Genome Research* 2011;21:718–24. <https://doi.org/10.1101/gr.117101.110>.

Xie G, Chen H, Jia D, Shu Z, Palmer WH, Huang Y-C, et al. The SWI/SNF Complex Protein Snr1 Is a Tumor Suppressor in *Drosophila* Imaginal Tissues. *Cancer Res* 2017;77:862–73. <https://doi.org/10.1158/0008-5472.CAN-16-0963>.

Yamada T, Fischle W, Sugiyama T, Allis CD, Grewal SIS. The Nucleation and Maintenance of Heterochromatin by a Histone Deacetylase in Fission Yeast. *Molecular Cell* 2005;20:173–85. <https://doi.org/10.1016/j.molcel.2005.10.002>.

Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, et al. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Molecular Cell* 2015;59:744–54. <https://doi.org/10.1016/j.molcel.2015.07.018>.

Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, et al. Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 2015;518:556–9. <https://doi.org/10.1038/nature13994>.

Zamurrad S, Hatch HAM, Drelon C, Belalcazar HM, Secombe J. A *Drosophila* Model of Intellectual Disability Caused by Mutations in the Histone Demethylase KDM5. *Cell Rep* 2018;22:2359–69. <https://doi.org/10.1016/j.celrep.2018.02.018>.

Zhang H, Roberts DN, Cairns BR. Genome-Wide Dynamics of Htz1, a Histone H2A Variant that Poises Repressed/Basal Promoters for Activation through Histone Loss. *Cell* 2005;123:219–31. <https://doi.org/10.1016/j.cell.2005.08.036>.

Zhang K, Mosch K, Fischle W, Grewal SIS. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 2008;15:381–8. <https://doi.org/10.1038/nsmb.1406>.

Zhao F, Yu C, Liu Y. Codon usage regulates protein structure and function by affecting translation elongation speed in *Drosophila* cells. *Nucleic Acids Research* 2017;45:8484–92. <https://doi.org/10.1093/nar/gkx501>.

Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, et al. Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. *Cell Stem Cell* 2007;1:286–98. <https://doi.org/10.1016/j.stem.2007.08.004>.

Zhaunova L, Ohkura H, Breuer M. Kdm5/Lid Regulates Chromosome Architecture in Meiotic Prophase I Independently of Its Histone Demethylase Activity. *PLoS Genet* 2016;12:e1006241. <https://doi.org/10.1371/journal.pgen.1006241>.

Zhou Z, Dang Y, Zhou M, Li L, Yu C, Fu J, et al. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc Natl Acad Sci USA* 2016;113:E6117–25. <https://doi.org/10.1073/pnas.1606724113>.

Zrally CB, Dingwall AK. The chromatin remodeling and mRNA splicing functions of the Brahma (SWI/SNF) complex are mediated by the SNR1/SNF5 regulatory subunit. *Nucleic Acids Research* 2012;40:5975–87. <https://doi.org/10.1093/nar/gks288>.

Zrally CB, Marena DR, Nanchal R, Cavalli G, Muchardt C, Dingwall AK. SNR1 is an essential subunit in a subset of *Drosophila* brm complexes, targeting specific functions during development. *Dev Biol* 2003;253:291–308. [https://doi.org/10.1016/s0012-1606\(02\)00011-8](https://doi.org/10.1016/s0012-1606(02)00011-8).

Zrally CB, Middleton FA, Dingwall AK. Hormone-response genes are direct in vivo regulatory targets of Brahma (SWI/SNF) complex function. *J Biol Chem* 2006;281:35305–15. <https://doi.org/10.1074/jbc.M607806200>.