JACK OF ALL TRADES, MASTER OF NONE

The multifaceted nature of H3K36 methylation

Henrik Lindehell

Dissertation for PhD
ISBN: 978-91-7855-952-7 (print)
ISBN: 978-91-7855-953-4 (pdf)
Cover art generated by AI with input from the author
Electronic version available at: http://umu.diva-portal.org/
Printed by: CityPrint i Norr AB
Umeå, Sweden 2022

The full quote:

“A jack of all trades is a master of none, but oftentimes better than a master of one.” – Robert Green describing William Shakespeare (1592)

Table of Contents

LIST OF PUBLICATIONS

This thesis is based on the following papers, which in the text will be referred to by their Roman numerals:

I. **Lindehell H***, Kim M*, Larsson J (2015) Proximity ligation assays of protein and RNA interactions in the male-specific lethal complex on *Drosophila melanogaster* polytene chromosomes. Chromosoma 124: 385-395

II. **Lindehell H**, Glotov A, Dorafshan E, Schwartz YB, Larsson J (2021) The role of H3K36 methylation and associated methyltransferases in chromosome-specific gene regulation. Sci. Adv. 7: eabh4390

III. **Lindehell H**, Schwartz YB, Larsson J (2022) Methylation of lysine 36 on histone H3 is required to control transposon activities in somatic cells. *(manuscript)*

* These authors contributed equally to this work

The following paper is not included in this thesis:

I. Mendoza-Garcia P, Basu S, Sukumar SK, Arefin B, Wolfstetter G, Anthonydhason V, Molander L, Uçkun E, **Lindehell H**, Lebrero-Fernandez C et al (2021) DamID transcriptional profiling identifies the Snail/Scratch transcription factor Kahuli as an Alk target in the *Drosophila* visceral mesoderm. Development 148: dev199465

ABSTRACT

Post-translational modifications of histones enable differential transcriptional control of the genome between cell types and developmental stages, and in response to environmental factors. The methylation of Histone 3 Lysine 36 (H3K36) is one the most complex and well-studied histone modifications and is known to be involved in a wide range of molecular processes. Commonly associated with active genes and transcriptional elongation, H3K36 methylation also plays a key role in DNA repair, repression of cryptic transcription, and guiding additional post-translational modifications to histones, genomic DNA, and RNA. In *Drosophila melanogaster*, trimethylated H3K36 has also been linked to dosage compensation of the single male X chromosome as a binding substrate for the Male-Specific Lethal (MSL) complex. However, this model has been challenged by structural and biochemical studies demonstrating higher MSL complex affinity for other methylated lysines. There is an additional system of chromosome-specific gene regulation in *D. melanogaster* where transcription from the small heterochromatic fourth chromosome is increased by Painting of fourth (POF), a protein specifically binding nascent RNA on the fourth chromosome. The fourth chromosome is thought to have been an ancestral X chromosome that reverted into an autosome. POF mediating high transcription levels from an autosome is believed to be a remnant of an ancient sex-chromosome dosage compensation mechanism.

Proximity ligation assays revealed no interaction between MSL complex components and methylated H3K36. This finding was corroborated by RNA sequencing of H3K36 methylation impaired mutants: the transcriptional output of the male X chromosome was unaffected in mutants where Lysine 36 on Histone 3 was replaced by an Arginine, abolishing methylation of this site. However, we found that knocking out *Set2*, which encodes the methyltransferase responsible for H3K36 trimethylation, significantly reduced X-linked transcription relative to autosomal transcription. This strongly suggests the existence of previously unrecognized alternate Set2 substrates. Interestingly, we also found that Ash1- and NSD-mediated methylation of H3K36 was required to maintain high expression from chromosome four.

Recent studies have also implicated H3K36 methylation in the silencing of transposon activity in somatic cells. By analyzing the transcription of transposable elements and Piwi-interacting RNAs (piRNAs), we identified dimethylation of H3K36 by Set2 as the main methylation mark involved in this process and showed that dual-stranded piRNA clusters are preferentially activated upon disturbing the H3K36 methylation machinery. These findings extends the long list of processes dependent on functional H3K36 methylation.

ABBREVIATIONS

ash1	*absent, small, or homeotic discs 1*
ChIP	Chromatin immunoprecipitation
CLAMP	*Chromatin-linked adaptor for MSL proteins*
HAS	High-affinity sites
LINE	*Long interspersed nuclear elements*
LTR	*Long terminal repeat*
mle	*maleless*
mof	*males absent on the first*
mRNA	Messenger ribonucleic acid
msl1	*male-specific lethal 1*
msl2	*male-specific lethal 2*
msl3	*male-specific lethal 3*
NSD	*Nuclear receptor binding SET domain protein*
PHD	Plant homeodomain
PionX	pioneering sites on the X
piRNA	*PIWI-interacting RNAs*
PLA	Proximity ligation assay
roX1	*long non-coding RNA on the X 1*
roX2	*long non-coding RNA on the X 2*
Set2	*SET domain containing 2*

INTRODUCTION

Drosophila melanogaster as a model organism

Molecular biology and genetics research requires suitable model organisms. Factors that should be considered when selecting a model organism include what you want to achieve, how much time and capital you have, and whether the questions you want to answer are biologically relevant outside the model. *Drosophila melanogaster* has been one of the most successful model organisms over the past century for several reasons including its small size, easy handling and short generation time. As such, it has been instrumental in the discovery of general genetic and biological mechanisms and the development of tools used to study many aspects of biology. *D. melanogaster* is well suited as a model in medical research because of its gene composition, which to a high degree resembles that of the mammalian genome and contain orthologs for approximately 75 percent of human disease-causing genes (Reiter *et al.*, 2001). Furthermore, *D. melanogaster* has become a powerful model for cancer research in recent years because most pathways relating to cancer growth and spread are conserved between flies and humans and many tissues relevant to cancer research are also present in flies (Mirzoyan *et al.*, 2019).

Arguably one of the most useful aspects of using *Drosophila* in genetic research is the presence of polytene chromosomes. Polytene chromosomes are chromosome complexes that have undergone multiple rounds of replication without cell division, forming huge structures of homologous chromatids that remain synapsed and can easily be examined under a regular light microscope. The centromeres of the replicating chromosomes clump together to form a central heterochromatic structure referred to as the chromocenter, to which all replicating chromosomes remain tethered. A distinctive feature of polytene chromosomes is that they have banded patterns reflecting their degree of chromatin condensation: darker bands correspond to denser chromatin regions with high concentration of DNA and chromatin-associated proteins, while the lighter interband regions consist of open and transcriptionally active chromatin. The banded patterns are chromosome-specific and were used in the early days of *Drosophila* genetics research (Bridges, 1935) to make remarkably precise genetic maps that have since proven invaluable for mapping genes, chromosomal rearrangements and chromatin interacting factors. More recent work has shown that the chromatin landscape of non-

polytene chromosomes can be used to accurately predict the locations of banded regions in polytene chromosomes (Vatolina *et al.*, 2011) and that polytene chromosome bands represent topologically associating domains (TADs) that are identical to TADs on non-polytene chromosomes (Eagen *et al.*, 2015). These revelations ensure that polytene chromosomes will remain powerful genetic tools that complement sequencing for years to come.

Chromosome wide gene regulation

The process of gene dose compensation occurs in organisms with sex chromosomes. Unlike autosomal pairs of chromosomes, which contain equal numbers of genes, sex chromosomes differ in both content and size. Since sex chromosomes contain many non-sex specific genes for which the optimal gene dose is often identical between the sexes (Ohno, 2013), there is a strong evolutionary incentive to equalize the gene dose between the heterogametic and homogametic sexes. Accordingly, several mechanisms that promote such equalization have evolved independently.

Different evolutionary strategies

In mammalian systems the imbalance between male and female is solved by the process of random X inactivation, i.e., the compaction and silencing of one X chromatid in females. This is orchestrated and transmitted through cell divisions by the combined action of multiple epigenetic regulators including DNA methylation, Polycomb repression, non-coding RNAs, silencing histone modifications, and variant histones (Lee, 2011). In the nematode *Caenorhabditis elegans*, both X chromosomes of the hermaphrodite sex are kept active but the transcriptional output of each one is reduced by 50% to match that of the single male X chromosome (Meyer, 2022). In both cases there is evidence suggesting upregulation of the X chromosomal gene dose in both sexes to equalize expression between autosomes and sex chromosomes (Deng *et al.*, 2011; Faucillion & Larsson, 2015; Lentini *et al.*, 2022).

Dosage compensation in *Drosophila*

Dosage compensation in *D. melanogaster* occurs through hyperactivation of the single male X chromosome and equalizes gene doses between males and females as well as between sex chromosomes and autosomes. The dosage compensation of the male X is facilitated by the Male-Specific Lethal (MSL) complex, which assembles exclusively in males and consists of at least five core proteins (MSL1, MSL2, MSL3, MOF, and MLE) and one of two long non-coding RNAs (*roX1* or *roX2*).

The main structural component of the MSL complex is the MSL1 protein, which after dimerization serves as a scaffold for interaction with MSL2, MSL3, and MOF (Hallacli *et al.*, 2012; Kadlec *et al.*, 2011; Kuroda *et al.*, 2016; Morales *et al.*, 2004).

MSL2 is the only male-specific component of the MSL complex and is blocked at the translational level in females (Kelley *et al.*, 1997). It is exclusively recruited to the male X-chromosome via its cysteine rich CXC-domain, which binds a sequence motif within a small number of sites known as pioneering sites on the X (PionX) (Villa *et al.*, 2016). After localizing to the PionX sites, the MSL complex populates an expanded set of sites called high affinity sites (HASs) (Straub *et al.*, 2008) or chromatin entry sites (CESs) (Alekseyenko *et al.*, 2008) before spreading to lower affinity sites and genes throughout the male X-chromosome. MSL2 is also an E3 ubiquitin ligase targeting Histone 2B Lysine 34, which in its ubiquitinylated state promotes methylation of H3K4 and H3K79, both of which have roles in increasing transcriptional activity (Wu *et al.*, 2011).

MSL3 is a chromodomain-containing protein involved in the spreading of the MSL complex from HASs throughout the male X chromosome (Sural *et al.*, 2008). However, there is some controversy regarding the interaction partner of the MSL3 chromodomain. It was first thought that MSL3 binds to the histone modification H3K36me3, which accumulates towards the 3′ end of active genes (Bannister *et al.*, 2005). However, structural and biochemical studies have shown that MSL3 cannot bind to H3K36me3 and instead shows higher affinity for mono-methylated H4K20 (Kim *et al.*, 2010; Moore *et al.*, 2010), a mark recently shown to promote transcription and chromatin openness (Shoaib *et al.*, 2021). The hypothesis that MSL3 merely colocalizes with H3K36me3 without interacting is supported by the findings presented in Papers I and II of this thesis. However, more research is needed to settle the issue.

The histone acetylase MOF is also part of the MSL complex and contributes to dosage compensation by locally increasing the abundance of H4K16 acetylated histones, which are associated with chromatin fiber decompaction and subsequent increased transcription (Akhtar & Becker, 2000). MOF also acetylates MSL3, which is necessary for the X chromosome localization of the latter protein (Buscaino *et al.*, 2003).

MLE is a DNA/RNA helicase and the primary factor responsible for incorporating the long non-coding RNAs *roX1* or *roX2* into the MSL complex (Meller *et al*, 2000). MLE binds *roX* RNAs via their stem-loop structures in an ATP-dependent manner (Maenner *et al.*, 2013) and its binding to the MSL complex is RNA-dependent (Richter *et al.*, 1996).

The MSL core complex also contain one of the aforementioned long non-coding *roX* RNAs. The two known *roX* RNAs, *roX1* and *roX2*, are functionally redundant for *D. melanogaster* dosage compensation and only one of them is present in a given MSL complex at any given time (Ilik *et al.*, 2013). Non-coding *roX* RNAs have also been shown to be essential for the MSL complex to spread in regions between HAS (Figueiredo *et al.*, 2014). The two *roX* RNAs differ greatly in size: *roX1* is nearly 4 kilobases long while a *roX2* transcript averages around 0.6 kilobases. Despite this considerable difference in size they share a homologous sequence motif known as the *roX* box (Franke & Baker, 1999; Park *et al.*, 2008) that is essential, together with stem-loop secondary structures present at the 3′ ends of both *roX* RNAs, for RNA stabilization and integration into the MSL complex (Park *et al.*, 2008).

There are additional factors contributing to MSL-driven dosage compensation. One is the protein Chromatin-Linked Adapter for MSL Proteins (CLAMP), which binds GA-rich motifs throughout the genome in both sexes of *D. melanogaster* and was specifically shown to be enriched within HAS found throughout the X chromosome (Larschan *et al.*, 2012; Soruco & Larschan, 2014). CLAMP interacts with MSL2 through an N-terminal Zinc-finger domain and is thought to act synergistically with the MSL2 CXC-domain to enhance the binding of the MSL complex to HASs (Tikhonova *et al.*, 2019; Tikhonova *et al.*, 2022). Interestingly, a recent high-throughput chromosome conformation capture study also implicates CLAMP in the formation of the three-dimensional structures of chromosomes and clustering of HAS on the X chromosome (Jordan & Larschan, 2021).

Another factor shown to aid X chromosome dosage compensation in *D. melanogaster* is the histone kinase JIL-1. JIL-1 phosphorylates Histone 3 at Serine 10 (H3S10), a mark that maintains open chromatin structures by counteracting heterochromatin formation (Jin *et al.*, 2000; Jin *et al.*, 1999; Regnard *et al.*, 2011; Wang *et al.*, 2001). Both JIL-1 and phosphorylated H3S10 were shown to be enriched on the male X chromosome (Wang *et al.*, 2001) and JIL-1 depletion reduces X-linked gene transcription (Regnard *et al.*, 2011). JIL-1 was also shown to physically interact with the MSL complex (Jin *et al.*, 2000), a relationship we examined in Paper I.

Even though multiple components of the MSL complex and its interactors have been identified, gene expression remains a highly complex process that can be regulated at multiple levels, and it is not yet fully clear what causes the MSL complex-mediated upregulation of dosage compensated genes on the male X. One theory gaining support is that the increase in transcription is coupled to enhanced transcriptional elongation rather than initiation. This theory is supported by the observations that H4K16ac accumulates preferentially towards the 3-prime end of genes rather than at promoters (Smith *et al.*, 2001), and that RNA polymerase density in gene bodies on the X chromosome is higher than on autosomes (Larschan *et al.*, 2011). However, it is unlikely an enhanced elongation rate by itself is sufficient to explain the observed upregulation of X-linked genes. Therefore, it probably happens in concert with other processes such as increased release of paused RNA polymerase (Ferrari *et al.*, 2013) and favorable spatial localization of the single male X chromosome within the nucleus (Ilyin *et al.*, 2022).

The *Drosophila melanogaster* fourth chromosome

D. melanogaster also has an additional chromosome that is subject to chromosome-wide regulation. The small heterochromatic fourth chromosome is believed to be an ancient sex chromosome that reverted to an autosome (Vicoso & Bachtrog, 2013, 2015). This is supported by the presence of another chromosome-specific compensatory mechanism centered around the protein Painting of fourth (POF) (Larsson *et al.*, 2001), which binds to nascent RNA of actively transcribed genes and increases their expression levels (Johansson *et al.*, 2012). The level of upregulation is comparable to that induced by MSL driven dosage compensation of the male X and is enough to ensure viability in haplo-4 flies (Johansson *et al.*, 2012; Johansson *et al.*, 2007; Stenberg *et al.*, 2009). No binding sites, such as HASs for the MSL complex, have been identified for the targeting of POF to the fourth chromosome. However,

POF colocalizes with Heterochromatin protein 1 (HP1), a key factor in heterochromatin formation, in coding regions of fourth chromosome genes (Johansson *et al.*, 2007). Additionally, POF was found to bind two loci on the X chromosome, *PoX1* and *PoX2* (POF-on-X), in close proximity to the *roX1* and *roX2* genes (Lundberg *et al.*, 2013). Further examination of these loci revealed that POF can bind to expressed genes if they have a downstream region with recurrent blocks of 1.688 satellite repeats (Kim *et al.*, 2018). The 1.688 satellite was also shown to be heavily enriched on the X chromosome compared to the autosomes (Gallach, 2014; Kim *et al.*, 2018; Kuhn *et al.*, 2011; Waring & Pollack, 1987). These observations together support the theory that POF was part of an ancient mechanism for sex chromosome targeted dosage compensation and provide an evolutionary link between the POF and MSL dosage compensation systems.

Post-translational histone modifications

Post-translational modifications of histones play a key role in chromatin organization and gene regulation. There is an ever-growing number of known modifications that can decorate histones and their amino acid tails, including acetylation, phosphorylation, ubiquitylation, and sumoylation. These enzyme-driven modifications serve two main roles in chromatin formation. The first is that a modified position will have different physiochemical properties to its unmodified precursor: modifications may change the charge, hydrophobicity, and/or steric bulk at the position, all of which may affect the chromatin's conformation. The second role of modified histones is that they serve as binding sites for chromatin-interacting proteins and thus create additional layers of control in chromatin-related processes. One of the most versatile histone modifications is the methylation of lysine residues. Unlike many other modifications, lysine methylation does not alter the histone's charge but does increase the local hydrophobicity and steric demand. Additional functional complexity arises from the fact that an individual lysine residue can be mono-, di- or trimethylated. Since each level of methylation has slightly different physicochemical effects and the different methylation levels have both shared and specific protein interactors, lysine methylation is associated with a long list of functions, some of which are in opposition to one-another.

Histone 3 Lysine 36 methylation

Methylation of histone 3 lysine 36 is one of the most extensively studied histone modifications and is highly conserved in eukaryotes. H3K36 can be mono-, di-, or trimethylated, adding to the mark's complexity and versatility. All three levels of H3K36 methylation have been implicated in preventing cryptic transcription, albeit with different and sometimes overlapping specificities (Difiore *et al.*, 2020; Li *et al.*, 2009a; Pfister *et al.*, 2014). A unique role for monomethylated H3K36 (H3K36me1) has yet to be identified and the mark shows only a slight preference for active promoters despite being ubiquitous within the genome (Barski *et al.*, 2007). Di- and trimethylated H3K36 (H3K36me2/H3K36me3), in contrast, have been implicated in many processes and have more specific genomic distributions. H3K36me2/me3 are both important factors in repair of double stranded DNA breaks (Fnu *et al.*, 2011; Pfister *et al.*, 2014) as well as the recruitment of DNA methyltransferases and the subsequent development of the DNA methylation landscape (Morselli *et al.*, 2015; Weinberg *et al.*, 2019; Xu *et al.*, 2020; Yano *et al.*, 2022).

However, the distributions of H3K36me2 and H3K36me3 differ: the former is mostly associated with intergenic regions and transcription start sites while the latter is enriched within exons and accumulates towards the 3′ end of coding genes. This exon-defining feature of H3K36me3 is linked to transcriptional elongation, which is the most thoroughly studied process associated with the mark.

H3K36 methyltransferases

There are three characterized histone 3 lysine 36 methyltransferases in *D. melanogaster*. SET domain–containing 2 (Set2) adds methyl groups to mono- and dimethylated H3K36 (Bell *et al.*, 2007), and is the only H3K36 methyltransferase known to trimethylate this position. Loss of Set2 function reduces the abundance of H3K36me3 tenfold and is lethal in the early pupal stage of development (Dorafshan *et al.*, 2019; Larschan *et al.*, 2007). Set2 binds and travels with RNA Polymerase II (RNAPol2) during elongation and performs its histone methyltransferase function co-transcriptionally (Kizer *et al.*, 2005). This process in turn influences numerous downstream pathways.

The *D. melanogaster* Nuclear receptor binding SET domain containing protein (NSD) is the least studied of the H3K36 methyltransferases. Research on its human homologs has suggested that it is capable of both mono- and dimethylation (Kudithipudi *et al.*, 2014; Li *et al.*, 2009b; Rayasam, 2003). However, unlike their mammalian counterparts, *D. melanogaster* NSD loss-of-function mutants are viable (Dorafshan *et al.*, 2019). The presence of H3K36me2 within the genome is thought to be largely due to the activity of NSD (Bell *et al.*, 2007), making this enzyme the main source of the substrate for Set2-mediated trimethylation.

Absent, small, or homeotic discs 1 (Ash1) is capable of mono- and di-methylating H3K36 (Tanaka *et al.*, 2007). It also counteracts erroneous PRC2 driven Polycomb repression (Yuan *et al.*, 2011) via a mechanism that was recently shown to be independent of its H3K36 methylation function (Dorafshan *et al.*, 2019). Unlike the other H3K36 methyltransferases, it has a very specific binding pattern on polytene chromosomes (Tripoulas *et al.*, 1996) and is therefore thought to only influence the overall H3K36 methylation landscape in specific locations.

H3K36 methylation and dosage compensation

Tri-methylated H3K36 has previously been postulated to be linked to dosage compensation of the male X chromosome in *D. melanogaster*. It has been suggested that the spreading of the MSL complex over the male

X chromosome following its initial binding to PionX and HASs is facilitated by the interaction between the MSL3 subunit's chromodomain and H3K36me3 (Larschan *et al.*, 2007; Sural *et al.*, 2008). This model is largely based on ChIP colocalization data and early structural studies but was not supported by proximity ligation assay between MSL3 and H3K36me3 (Paper I) or by differential expression analysis of H3K36 methylation-impaired mutants (Paper II). Moreover, structural and biochemical studies have suggested that the affinity of MSL3 for mono- and di-methylated histone 4 lysine 20 exceeds that for H3K36me3 (Kim *et al.*, 2010; Moore *et al.*, 2010), indicating a need to revise the H3K36 methylation-centered model of MSL complex spreading over the X chromosome.

H3K36 methylation in DNA repair

DNA damage is commonly caused by reactive oxygen species, radiation, and other environmental stressors. Multiple repair mechanisms have therefore evolved to repair different types of damage. Di- and trimethylated H3K36 is predicted to play a central role in homologous recombination (Aymard *et al.*, 2014; Daugaard *et al.*, 2012; Huen *et al.*, 2010), non-homologous end joining (Huen *et al.*, 2010; Musselman *et al*, 2012), and DNA mismatch repair (Li *et al.*, 2013). The common thread between these mechanistically distinct systems is that H3K36me2/me3 enables recruitment of various DNA repair effectors containing specific chromatin interacting structures such as Chromo-, Pro-Trp-Trp-Pro (PWWP-), and Tudor-domains.

H3K36 methylation and repression of cryptic transcripts

Cryptic transcription is the process of transcription from intragenic promoters, at which the transcription machinery cannot assemble under normal conditions (Cheung *et al.*, 2008; Kaplan *et al.*, 2003). These transcripts are stable and often polyadenylated and have in some cases been shown to translate into proteins (Cheung *et al.*, 2008). Increased cryptic transcription leads to accelerated cellular ageing and a shorter life span (Sen *et al*, 2015). Its initiation is prevented by deacetylation in the wake of RNAPol2. The Set2 methyltransferase travels together with RNAPol2 and methylates H3K36 towards the 3′ end of genes. The Rpd3S histone deacetylase complex recognizes trimethylated H3K36 via the combined action of the PHD- and Chromo-domains on the Rco1 and Eaf3 subunits (Li *et al.*, 2007). As more enzymes bearing domains recognizing H3K36 methylation have been discovered, more interactors associated with transcriptional fidelity have been identified. Two notable examples

are the short isoform of alternative splicing factor Psip1 (Pradeepa *et al.*, 2012) and the Isw1b chromatin remodeling complex (Maltby *et al.*, 2012).

H3K36 methylation as a guide for m6A RNA modification

Methylation of adenosine at N6 to form N6-methyladenosine (m6A) is the most common post-transcriptional modification in mRNA and is important in regulating mRNA stability and translation efficiency (Huang *et al.*, 2018; Wang *et al.*, 2014; Wang *et al.*, 2015). Intriguingly, recent studies have linked m6A modification of nascent RNAs to trimethylated H3K36. The interaction between the m6A methyltransferase complex and H3K36me3 is mediated by the Methyltransferase 14 (METTL14) subunit, which binds H3K36me3 (Huang *et al.*, 2019). The mechanism of this interaction is unclear because structural analyses of METTL14 revealed no known H3K36me-interacting domains (Wang *et al.*, 2016). However, it was shown that once this interaction forms, the m6A modification is installed on nascent RNA simultaneously with transcription (Huang *et al.*, 2019).

H3K36 methylation in cancer

Because of its many roles and interactors, the function and outcome of H3K36 methylation can be disturbed at multiple levels. Malfunctions of H3K36 histone methyltransferases have been associated with cancers as varied as acute myeloid leukemia (Wang *et al.*, 2007), multiple myeloma (Huang *et al.*, 2013; Swaroop *et al.*, 2019), head and neck squamous cell carcinoma (Papillon-Cavanagh *et al.*, 2017), lung squamous cell carcinoma (Brennan *et al.*, 2017), clear cell renal cell carcinoma (Dalgliesh *et al.*, 2010; Su *et al.*, 2017), neuroblastomas (Berdasco *et al.*, 2009), and breast cancer (Angrand *et al.*, 2001).

The two families of H3K36-related demethylases, KDM2 and KDM4, have also been implicated in cancer progression. The KDM demethylases are typically overexpressed in tumors, resulting in the removal of H3K36me2 and H3K36me3 from key locations such as promoters and tumor suppressor loci (He *et al.*, 2008; Mishra *et al.*, 2018; Wagner *et al.*, 2013). KDM demethylase interference is therefore among the most promising targets in pharmacological cancer research and multiple compounds blocking their activity are being tested as potential anticancer agents (Chu *et al.*, 2014; Suzuki *et al.*, 2013).

The erroneous substitution of Histone 3 Lysine 36 by methionine is also a common cancer marker. Although patients have been found with alterations to Histone 3.1 and Histone 3.2, the K-to-M mutation of

Histone 3.3 is the most prevalent variant in cancer (Lu *et al.*, 2016). Dramatic changes in the distributions of H3K36me2 and H3K36me3 along with reductions in their abundance have been observed in H3.3K36M cancer cells. These changes seem to result primarily from direct inhibition of the enzymatic activity of K36-related histone methyltransferases by H3.3K36M (Lu *et al.*, 2016). Moreover, a different substitution, H3K36R (Lysine to Arginine), which does not inhibit methyltransferase activity, was found in some cases of T-cell acute lymphoblastic leukaemia (Collord *et al.*, 2019). This finding underscores the importance of modifiable H3K36 for normal cell proliferation.

H3K36 methylation and neurodevelopmental disorders

H3K36 methylation is also associated with neurodevelopmental disorders. Mutations in the human H3K36 methyltransferases *SetD2* and *SETD5* have been linked to autism spectrum disorders (Lumish *et al.*, 2015), Luscan-Lumish syndrome (Luscan *et al.*, 2014), and 3p25.3 microdeletion syndrome (Grozeva *et al.*, 2014). H3K36 methylation was also shown to recruit DNA methyltransferases; the failure of this recruitment in the developing brain due to mutations in H3K36me writers and readers is associated with Tatton-Brown-Rahman syndrome (Tatton-Brown *et al.*, 2014) and Sotos syndrome (Kurotaki *et al.*, 2002).

Transposable elements

Transposable elements (TEs) are selfish genomic DNA segments capable of replication and translocation either on their own or with the help of other transposable elements. TEs are abundant in all eukaryotic branches and can comprise large portions of eukaryote genomes (e.g., 30% in *D. melanogaster*, 50% in humans, and >80% in Maize). There are an abundance of different types of transposable elements with differing means of mobilization, genetic contents, and sizes. The first rudimentary classification system for transposons (Finnegan, 1989) divides transposable element into two groups: class I elements, which all translocate via an RNA intermediate, and class II elements, which do not. These classifications are still used today but the system has been augmented to include lower levels of categorization including subclass, order, superfamily, and family (Wicker *et al.*, 2007). Two subclasses exist: the first comprises TEs capable of self-mobilization, while the second comprises those reliant on outside factors for movement. Since all class I elements are capable of self-mobilization, this level is only relevant for class II elements. The order-level subdivides transposons by their mode of insertion, which is largely dependent on their gene content. Categories at the two lowest levels, superfamily and family, are distinguished based on classic phylogeny in the form of protein structure and DNA sequence similarity data (Wicker *et al.*, 2007). However, due to the high DNA sequence variability between elements that are closely related in evolutionary terms, even the family level isn't easily defined. Consequently, there is heavy reliance on simplifications like the 80-80-80 rule. This rule states that for two elements to be part of the same family they must share at least 80 percent sequence identity in at least 80 percent of the internal domain and/or repeat regions and the full aligned transposon sequence must be at least 80 base pairs long to avoid misclassification (Wicker *et al.*, 2007).

The defining trait of class I elements is that their replication relies on reverse transcriptase and an RNA intermediate. Consequently, their original template remains intact. Class I elements are therefore generally referred to as copy-and-paste transposons. With only a few exceptions, class I elements can be further divided into two large subclasses: Long terminal repeats (LTRs) and non-LTRs.

LTRs are evolutionarily related to retroviruses and closely resemble them in structure and function (Eickbush & Malik, 2007). *Drosophila* LTRs are typically 5 to 7 kilobases long and flanked on either side by repeated

regions of 300-400 base pairs. Their gene content can differ slightly but they all contain *pol*, which encodes a fused protease, integrase, and reverse transcriptase. After translation, these components are cleaved by the protease. All LTRs also contain *gag*, which encodes a protein that forms a capsid around the LTR RNA template, the integrase and reverse transcriptase, and an endogenous tRNA that functions as a primer for reverse transcription of the LTR RNA into double-stranded DNA (Wilhelm & Wilhelm, 2001). The capsid is assembled in the cytoplasm and then shuttled into the nucleus, where it releases its contents. The integrase then inserts the LTR DNA into the host genome. Three major LTR superfamilies exist in *Drosophila*: Ty3/*Gypsy*, *BEL-Pao and* Ty1/*Copia*. Their relative abundance differs between fly species; *Gypsy* is the most abundant in *D. melanogaster*, comprising 8.2% of the genome (Smit *et al.*, 2015). Some *Gypsy* elements have been shown to carry *envelope*, a retrovirus glycoprotein that enables intercellular spreading (Keegan *et al.*, 2021). LTRs of this type are commonly referred to as endogenous retroviruses.

Class I non-LTRs differ from LTRs by their mode of mobilization. While they all contain a reverse transposase, their genetic makeup is more diverse than that of LTRs. The dominant non-LTRs in *D. melanogaster* are *Jockey* elements from the LINE superfamily. *Jockey* repeats have two open reading frames encoding reverse transcriptase, an endonuclease, and a protein with an RNA recognition motif and zinc finger domains (Metcalfe & Casane, 2014). After translation in the cytoplasm, the proteins bind their common template and the resulting particle is shuttled back into the nucleus. The mechanism of *Jockey* insertion is not fully understood but related LINE elements insert via a mechanism in which the endonuclease cuts one strand of the host DNA, exposing a 3′ end that acts as a primer for target-primed reverse transcription of the LINE template RNA (Cost, 2002). Finally, the other strand is cleaved and the complementary DNA is synthesized (Cost, 2002; Jiang *et al.*, 2014).

Class II elements are a large group of DNA templated transposons with a range of different insertion strategies. In *Drosophila* two groups of class II transposons dominate the genome: Terminal inverted repeats (TIRs) and Helitrons. TIRs translocate using a cut-and-paste mechanism enabled by a transposase protein that is translated from the TIR itself. The most abundant family of TIRs in *Drosophila melanogaster* is the P element, which has been extensively studied and comprises roughly 1% of the genome (Smit *et al.*, 2015). Helitrons are smaller than TIRs and other class I transposons and lack the genes to self-mobilize. Because of the lack of actively translocating Helitrons, little is known about their mobilization

and insertion; models resembling the copy-and-paste as well as cut-and-paste mechanisms have been proposed (Grabundzija *et al.*, 2018; Li & Dooner., 2009).

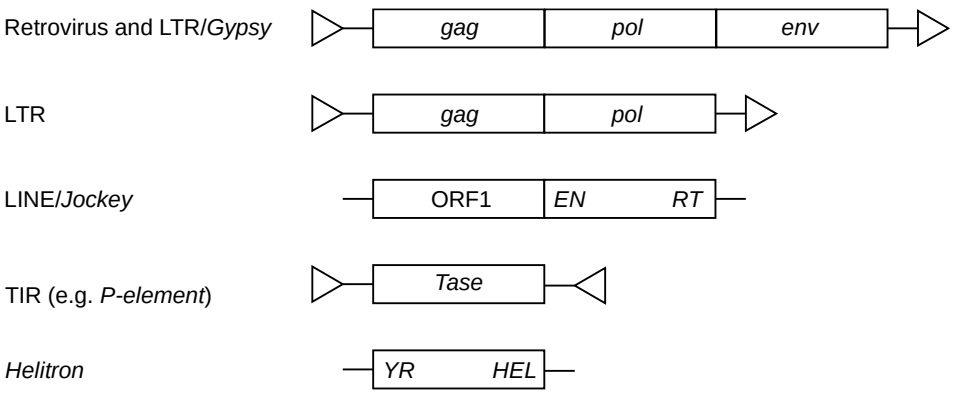

Figure 1: Schematic structures of different transposons
(gag=capsid protein, pol=fused protease integrase and reverse transcriptase, env=Envelope protein, EN=Endonuclease RT=Reverse transcriptase, Tase=Transposase, YR=Tyrosine recombinase, HEL=Helicase)

The PIWI-interacting RNA pathway

First discovered in *Drosophila* (Aravin *et al.*, 2001), PIWI-interacting RNAs (piRNAs) are short (23-30 bp) sequences that are almost exclusively associated with transposable elements (Aravin *et al.*, 2003). PiRNAs are tightly grouped into genomic clusters that can further be classified into uni-strand and dual-strand clusters depending on their mode of transcription (Brennecke *et al.*, 2007; Mohn *et al.*, 2014). Long transcripts of these piRNA clusters are exported to the cytoplasm, where they are processed into multiple individual piRNAs. The location of this pre-processing differs between germline and somatic cells. In the *Drosophila* germline, the piRNA pathway is hyperactivated and constitutes the cells' main defence against transposon activity that could alter the genome and thus potentially affect future generations. In germ cells, piRNA cluster transcripts are pre-processed in perinuclear compartments called 'nuage' granules (Lim & Kai, 2007). Conversely, in somatic cells, cleavage into individual piRNA precursors occurs in cytoplasmic Yb bodies (Olivieri *et al*, 2010). The precursors are then processed further by endoribonucleases and labelled with a 5′-monophosphate group (Ipsaro *et al.*, 2012; Nishimasu *et al.*, 2012). They may also undergo secondary structure remodeling (Vourekas *et al.*, 2015) before binding to an Argonaute protein belonging to the PIWI-clade for final processing and export as a mature PIWI-piRNA riboprotein complex (Czech & Hannon, 2016).

From this point forward the fate of the PIWI-piRNA complex depends on which PIWI protein was loaded during pre-processing. In *D. melanogaster* there are three Argonaute proteins in the PIWI-clade: P-element induced wimpy testis (Piwi), Aubergine (Aub), and Argonaute 3 (AGO3). If bound by Aub, the complex, guided by the loaded piRNA sequence, scans for transposable element transcripts in the cytoplasm and cleaves them. The cleavage product in turn produces a 5′ complementary piRNA that is recognized by AGO3 and guides the latter protein to unprocessed piRNA cluster transcripts. AGO3 then selectively cleaves these transcripts, amplifying specific piRNAs for free Aub to bind, leading to more efficient targeting of expressed transposons. This elegant sequence of molecular events was first described in *D. melanogaster* (Brennecke *et al.*, 2007; Gunawardane *et al.*, 2007) and has become known as the 'ping-pong cycle'. Similar piRNA amplification pathways have since been identified in many other species (Czech & Hannon, 2016).

If the piRNA precursor was instead bound by Piwi, the resulting complex is shuffled into the nucleus where, guided by the loaded piRNA, it scans

for transposons undergoing transcription (Post *et al.*, 2014). Then, via a mechanism that involves DNA methylation in mammals, it silences the transposon's activity (Aravin *et al.*, 2008). However, DNA methylation is either absent or exceedingly rare in *D. melanogaster*; the abundance of methylated DNA is far below the detection threshold of bisulphite sequencing (Capuano *et al.*, 2014). Also, no functional DNA methyltransferases have been identified in this species – the only possible DNA methyltransferase family member in the *D. melanogaster* genome is DNMT2, which has no effect on cytosine methylation (Takayama *et al.*, 2014) and was speculated to specifically methylate RNAs (Goll *et al.*, 2006; Schaefer *et al.*, 2010). Given the absence of DNA methylation in *D. melanogaster*, an alternative mechanism for the silencing of transposon loci exist. Transposon silencing by heterochromatin formation is often suggested as the mechanism responsible and is postulated to function through Piwi guided recruitment of heterochromatin related factors to transposon loci (Andreev *et al.*, 2022; Teo *et al.*, 2018). However, many of the described mechanisms of transposon silencing are germline specific. Mechanisms in somatic cells is not as well understood and currently a subject for intense study. Interestingly, a recent study proposesd that methylation of Histone 3 Lysine 36 plays an important role in transcriptional silencing of transposons in the developing eye of *D. melanogaster* (Chaouch *et al.*, 2021), a relationship closely examined in Paper III.

RESULTS AND DISCUSSION

Paper I

Proximity ligation assays of protein and RNA interactions in the male-specific lethal complex on *Drosophila melanogaster* polytene chromosomes

Summary

To visualize *in-situ* protein-protein and protein-RNA interactions within and around the MSL dosage compensation complex, we modified the Proximity Ligation Assay (PLA) technique to work on endoreplicated polytene chromosomes. PLA is an antibody-based technique used to detect interactions based on the physical proximity of different proteins and/or RNAs. Interactions can be detected with single molecule resolution directly on tissues and in cells. Primary antibodies raised in two different species are used to target the interaction partners of interest and are then bound by oligonucleotide-labeled secondary antibodies. Connector oligos are subsequently added, allowing a ligase to form a closed loop of DNA if the proteins are within interacting distance. One of the PLA probes then serves as a primer for DNA polymerase, enabling the formation of a long continuous DNA strand via rolling circle amplification. This strand remains tethered to the interaction site and is then hybridized with labeled oligos, allowing visualization and quantification of the interaction by brightfield or fluorescent microscopy.

Applying PLA to polytene chromosomes enables both the testing of protein interactions and the mapping of interactors to specific chromosomal locations. As a proof of principle, we tested the PLA system on the MSL dosage compensation complex, which specifically binds the male X chromosome. We anticipated that the results obtained would both validate our new method and clarify the precise makeup of the complex and its interacting partners, helping to resolve the ongoing debate about these matters.

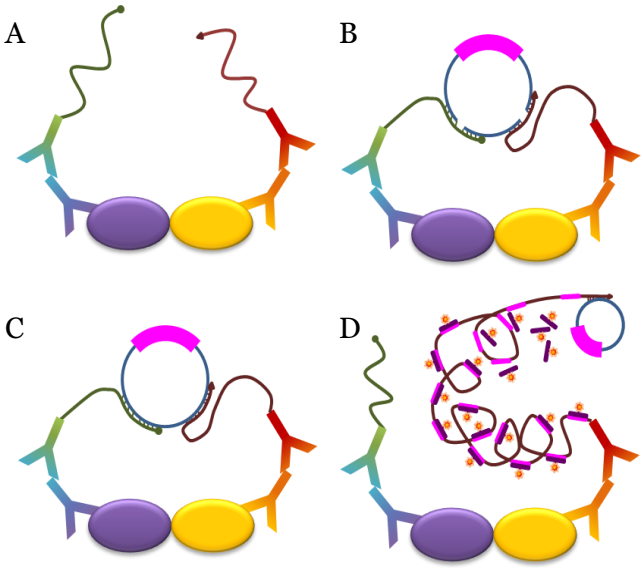

Figure 2: Proximity ligation assay
A) Two antibodies from different species, one for each interaction partner, are added followed by oligonucleotide-labeled secondary antibodies. B) Connector oligos are then added, allowing the antibody-bound oligos to form a loop if in sufficiently close proximity. C) The loop is closed by adding a ligase and one of the probes serves as a primer for rolling circle amplification upon adding DNA polymerase. D) Labeled oligos are hybridized to the tethered DNA strand, enabling microscope detection.

All core MSL components appear in proximity on the male X chromosome

By performing PLA on polytene chromosomes, we verified that all core protein components of the MSL complex (MSL1, MSL2, MSL3, MOF and MLE) colocalize within interaction range and generated a clear signal specific for the male X chromosome. We also confirmed the interaction between MLE and the *roX1* or *roX2* long non-coding RNAs using biotin-labeled antisense *roX1* and *roX2* RNA, demonstrating that the PLA system can be used to visualize both protein-protein and protein-RNA interactions on polytene chromosomes.

The MSL complex interacts with JIL-1, CLAMP, and Top2 but not H3K36me3

After validating the system, we moved on to study interactions between the MSL complex and other potential interactors. JIL-1 is a histone kinase enriched on the male X that is believed to counteract heterochromatin

formation through H3S10 phosphorylation, and has been suggested to interact with the MSL complex (Jin *et al.*, 2000; Jin *et al.*, 1999; Regnard *et al.*, 2011; Wang *et al.*, 2013). Our results using PLA support interaction with MSL1 and MSL2 but not MSL3. Given the length of the PLA probe oligos, this suggests that JIL-1 interacts with the MSL complex but the location of MSL3 within the complex is such that direct physical interaction between the two proteins is impossible.

CLAMP is another interactor of the MSL complex shown to be enriched at high affinity sites along the X chromosome (Larschan *et al.*, 2012; Soruco & Larschan, 2014). Our results show that CLAMP is in sufficiently close proximity the MSL complex components MSL2, MSL3 and MLE to physically interact. Topoisomerase 2 (Top2) is a DNA topology-altering enzyme enriched within X chromosomal genes in males and has been shown to be recruited there by the MSL complex component MLE (Cugusi *et al.*, 2013). Our results corroborate the proposed interaction between MLE and Top2 on the male X chromosome. We also show that MSL1 and Top2 colocalize in close proximity to each other. However, there was no detectable signal for interactions with MSL2 and MSL3, suggesting that their location in the MSL complex prevents direct physical interaction with Top2.

Interestingly, we were unable to corroborate the proposed interaction between MSL3 and the H3K36me3 histone mark (Larschan *et al.*, 2007). MSL3 was suggested to bind H3K36me3 via its chromo-domain to stabilize binding between active genes on the X chromosome and the MSL complex (Larschan *et al.*, 2007; Sural *et al.*, 2008). We cannot exclude the possibility that the lack of signal was due to transience of the interaction. However, since structural and biochemical studies have found that the MSL3 chromo-domain has a higher affinity for other histone lysine residues (Kim *et al.*, 2010; Moore *et al.*, 2010), it is conceivable that MSL3 and H3K36me3 merely colocalize while remaining too widely separated for direct physical interaction.

Discussion

PLA on polytene chromosomes is a fast and inexpensive tool that allows researchers to obtain valuable insights into potential protein interactions at specific genomic locations. PLA on polytene chromosomes offers a high-resolution method for distinguish simple protein colocalization (which can be detected using ChIP-seq and CUT&Tag experiments) from potential protein-protein interactions. Moreover, because it enables

visual detection of RNA-protein interactions, *in-situ* PLA on chromatin could be valuable for studying the interactions of non-coding RNAs.

In Paper I we showed that all components of the MSL complex interact throughout the male X chromosome. This is inconsistent with a 2013 model proposing partial and differential assembly of the MSL complex at specific locations on the X chromosome (Straub *et al.*, 2013). Moreover, it suggests that functional dosage compensation requires assembly of the entire MSL complex, in accordance with the currently prevailing model.

No signal was detected for the proposed interaction between the MSL3 chromo-domain and H3K36me3 (Sural *et al.*, 2008). This could be because the interaction is transient or the H3K36me3 antibody blocks the binding epitope for MSL3. However, more excitingly, it may indicate the existence of an alternative substrate for MSL3 binding. In accordance with this possibility, structural and biochemical studies have shown that MSL3 has a higher affinity for H4K20me1 than for H3K36me3, suggesting that H4K20me1 could be this alternative substrate (Kim *et al.*, 2010; Moore *et al.*, 2010). H3K20me1 was also shown to be involved in dosage compensation of the X chromosome in *C. elegans* (Brejc *et al.*, 2017) and to be a target for the human ortholog of the MSL complex, which catalyzes H4K16 acetylation like its *D. melanogaster* counterpart but does so genome-wide rather than exclusively on the male X chromosome (Kapoor-Vazirani & Vertino, 2014; Smith *et al.*, 2005). Whether H4K20me1 is the main substrate of MSL3 in *D. melanogaster* remains an open question.

Paper II

The role of H3K36 methylation and associated methyltransferases in chromosome-specific gene regulation

Summary

Dosage compensation is the process responsible for equalizing differences in X chromosome gene doses between males (XY) and females (XX). It often results from the coordinated effects of multiple factors, one of which is the targeting of specific modified histone residues by proteins involved in transcriptional regulation. In *D. melanogaster*, trimethylated H3K36 has been proposed to play a crucial role in the spreading of the MSL dosage compensation complex over the single male X chromosome (Larschan *et al.*, 2007; Sural *et al.*, 2008). This model has been challenged by structural studies showing that the binding affinity of the MSL complex is higher for methylated H4K20 than for trimethylated H3K36 (Kim *et al.*, 2010; Moore *et al.*, 2010). However, the loss of the H3K36 methyltransferase Set2 led to partial loss of MSL complex binding to dosage compensated genes (Larschan *et al.*, 2007). There is an additional system for chromosome wide gene regulation in *D. melanogaster* that acts on the small heterochromatic fourth chromosome. The fourth chromosome is an ancient sex chromosome that reverted to an autosome (Vicoso & Bachtrog, 2013, 2015), and it was hypothesized that an ancestral POF-mediated dosage compensation mechanism became trapped there in the process. Based on this knowledge, we wanted to investigate the link between H3K36 methylation and chromosome-wide transcriptional regulation in *D. melanogaster*.

H3K36me3 is enriched on the fourth chromosome but not on the male X

We initially performed immunostaining for H3K36me3 on salivary gland polytene chromosomes from third instar male larvae and were surprised to find that the X chromosome appeared to be less enriched in H3K36me3 than autosomes. We also observed a stronger H3K36me3 signal from the small forth chromosome than from other autosomes. Intrigued by these findings, we quantified H3K36me3 within exons on the different chromosomes using public ChIP data and verified the immunostaining results.

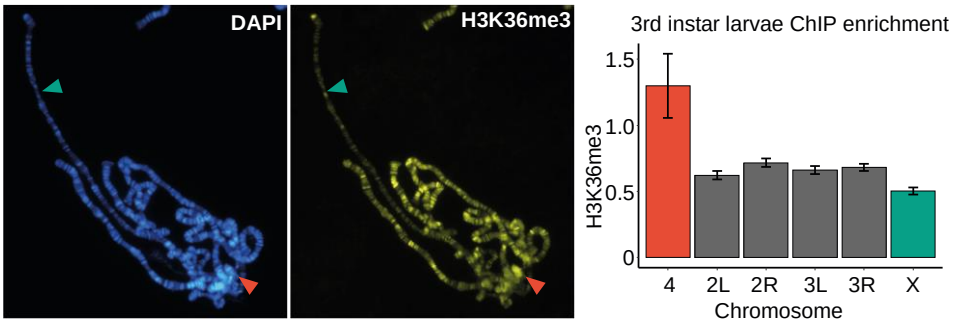

Figure 3: The male X and fourth chromosome differ in H3K36me3 enrichment compared to autosomes (Figure adapted from Paper 2)

Set2 is required for balanced transcription of the X chromosome independently of its function as a H3K36 histone methyltransferase

Because the interaction between the MSL complex and H3K36me3 was proposed to play a vital role in X-linked dosage compensation, we performed RNA-seq in mutants of the three H3K36-associated methyl transferases in *D. melanogaster* (Set2, NSD, and Ash1) and a histone replacement mutant where lysine 36 on the replication-dependent Histone 3 is substituted by an arginine (H3K36R). Due to the lethality of some mutants in the later developmental stages, we used brains from third instar male larvae in these experiments.

We found that there was significantly less X-linked transcription in *Set2* mutants than in both the wildtype and the autosomes of the mutants. The decrease in transcription originated from genes that are highly transcribed in the wildtype, extensively bound by MSL, and proximal to high-affinity sites. Intriguingly, the X-linked loss in transcription was not observed in the H3K36R mutant. Set2 does not methylate H3K36R, suggesting that the role of Set2 in dosage compensation is independent of its function as a H3K36 histone methyltransferase. The alternative mechanism by which Set2 enhances X-linked expression is currently unknown.

Ash1 and NSD together maintain balanced transcriptional output from the fourth chromosome

Mutants for the remaining H3K36 methyltransferases, *NSD* and *ash1*, exhibited no effect that distinguished the X chromosome from the autosomes. Interestingly however, both mutants individually showed small but significant reductions in transcriptional output from the fourth

chromosome. This effect was amplified in *ash1, NSD* double-knockout mutants to a degree similar to that seen for mutants lacking POF, the key protein linked to transcriptional upregulation of the fourth chromosome. To rule out the loss of POF from chromosome four in *ash1, NSD* mutants, we performed immunostaining experiments that revealed no difference in POF binding compared to the wildtype. The chromosome-specific decrease in transcription was thus recapitulated in H3K36R histone replacement mutants but not in *Set2* mutants, indicating that collaborative methylation by Ash1 and NSD is required for proper transcriptional output from the fourth chromosome.

Discussion

Two active systems of gene dose compensation exist in *D. melanogaster*. The first equalizes the transcriptional output of the single male X to that of the two X chromosomes in females and that of autosomes. The second is mediated by POF and specifically targets the fourth chromosome, where it binds nascent RNAs and increases their expression levels (Johansson *et al.*, 2012). In Paper II, we demonstrate differential enrichment of H3K36me3 between the fourth chromosome and the single male X – specifically, H3K36me3 is enriched on the fourth chromosome relative to the other autosomes, but its abundance on the male X is comparatively low. To quantify the impact of H3K36 methylation on chromosome-wide transcription levels we performed RNA sequencing in mutants lacking either the ability or the correct substrate to produce methylated H3K36.

Our finding that Set2 is required to maintain high transcriptional output independent of its histone methyltransferase activity was very intriguing and also consistent with previous reports suggesting that other methyltransferases may have alternative substrates (Dorafshan *et al.*, 2019). It was initially proposed that H3K36me3 plays a key role in the spreading of the MSL complex over the X chromosome (Sural *et al.*, 2008), but this model has been challenged by structural and biochemical studies suggesting that this role may actually be played by methylated H4K20 (Kim *et al.*, 2010; Moore *et al.*, 2010). This fits with our observation that X-linked transcription was not reduced in H3K36R histone replacement mutants and H3K36me3 was less abundant on the male X than on autosomes. In contrast we observed increased levels of H3K36me3 on the fourth chromosome and showed that this increase depends on uncompromised Ash1 and NSD methylation of H3K36 to maintain high transcriptional output. However, the fact that transcription from the male X was reduced in *Set2* mutants prompted us to speculate

about an alternative function of Set2. One potential pathway that could reduce X-linked transcription in males independently of H3K36me3 would involve reduced H4K16 acetylation, a mark associated with chromatin accessibility that is specifically enriched on the male X chromosome (Bone *et al.*, 1994; Turner *et al.*, 1992). Previous studies have demonstrated an X chromosome-specific reduction in H4K16ac abundance upon reducing Set2 activity (Bell *et al.*, 2008). Acetylation of H4K16 on the X chromosome is mediated by the MOF subunit of the MSL complex, so the reduced acetylation was attributed to reduced binding of the MSL complex to the male X driven by a reduction in Set2-mediated H3K36 trimethylation. However, this model doesn't fit if the MSL complex is actually directed to the male X by an alternative histone modification such as H4K20me1 (Kim *et al.*, 2010; Moore *et al.*, 2010). A recent study identified a list of non-histone substrates for SetD2, the human ortholog of Set2 (Schuhmacher *et al*, 2020). Intriguingly, one of these substrates is Histone deacetylase 9 (HDAC9), which has been shown to deacetylate H4K16 in mice (Cheng *et al.*, 2018). Although its substrates may vary between species, there are two homologs of HDAC9 in *D. melanogaster*: HDAC4 and HDAC6. Of these, HDAC6 appears to have a conserved lysine in a short peptide sequence similar to the one proposed to be modified by SetD2 in human HDAC9. HDAC proteins are primarily known to be substrates for acetyltransferases and kinases, but there is precedent for the methylation of deacetylases by Set-domain containing methyltransferases: human G9a methylates both HDAC1 (Rathert *et al.*, 2008) and Sirtuin 1 (Moore *et al.*, 2013). It is therefore conceivable that Set2 mediated methylation of a HDAC could inhibit its demethylase activity. If that were so, loss of *Set2* could lead to increased H4K16 deacetylation and thus to chromatin compaction that would reduce the efficiency of dosage compensation.

Paper III

Methylation of lysine 36 on histone H3 is required to control transposon activity in somatic cells

Summary

Transposable elements encompass a large group of genomic sequences capable of translocation throughout the genome. Processes regulating transposon activity have mostly been studied in the germline since only these cells can pass their modified genomes to offspring. However, accumulating evidence linking transposon activity in somatic cells to various developmental and neurological diseases has transformed the perception of somatic transposition and made it a growing field of research. Various mechanisms for transposon repression have evolved; one of the most efficient and well-studied is the PIWI/piRNA pathway, which both degrades already transcribed transposons and silences their future transcription. A recent study proposed a link between H3K36 methylation and transposon repression in *D. melanogaster* somatic cells (Chaouch *et al.*, 2021). We wanted to further investigate this relationship and determine the division of labor between the H3K36-specific methyltransferases in this process. It was not possible to do this in the earlier study because the H3.3K36M histone substitution mutant studied by Chaouch et al. is known to inhibit the enzymatic activity of H3K36 methyltransferases (Lu *et al.*, 2016). We therefore instead utilized a K36R histone substitution mutant without this toxicity (Lu *et al.*, 2016). This was crucial because two of the three H3K36 methyltransferases have been proposed to have additional functions independent of their lysine 36 methylation capability (Paper II and Dorafshan *et al.* (2019)).

Loss of H3K36 methylation increases transposon activity

RNA-seq data revealed that the lysine-to-arginine substitution mutant *H3K36R* had markedly higher transposon transcript levels than controls, indicating a significant loss of transposon repression. Moreover, a comparison to reanalyzed data for the H3.3K36M mutant (Chaouch *et al.*, 2021) revealed that the effect on transposon repression was more dramatic in *H3K36R*. Additionally, no significant increase in transposon activity was observed upon knocking out replication independent Histone 3.3. This suggests that the main mechanism of transposable element repression instead involves methylation of replication dependent Histone 3.

Set2-driven dimethylation of H3K36 is required for proper transposon repression

To clarify the division of labor between the H3K36 histone methyltransferases Set2, Ash1 and NSD in transposon repression, we quantified RNA-seq data on transposon expression in four mutants: the *H3K36R* histone H3 substitution mutant discussed above and *Set2*, *NSD*, and *ash1* knockout mutants. This revealed that most transposons expressed in *H3K36R* were also expressed in *Set2* while *NSD* and *ash1* exhibited reduced transposon activity. Grouping transposons that were exclusively activated in only one methyltransferase mutant revealed no obvious differences in genomic location between *Set2*, *NSD*, and *ash1*. However, upon grouping the transposons by order and superfamily, we found significant differences between the methyltransferases. When compared to the genomic abundance of transposon families, all three mutants expressed a disproportionately high proportion of LTR elements: the proportions of transposon activity mapped to LTRs in *Set2*, *NSD*, and *ash1* were 65%, 30%, and 23%, respectively. In *NSD* we observed significant enrichment of LINE repeats, with 19% of all activated transposons belonging to the *Jockey* superfamily. Aside from a slight but significant increase in the relative abundance of *Copia* LTRs, the *ash1*-exclusive transposon distribution resembled the genome-wide transposon family distribution more closely than those for the other methyltransferase mutants. ChIP-seq data for mono-, di- and trimethylated H3K36 in adult heads revealed H3K36me2 enrichment in LINE and LTR elements, which stands in contrast to the H3K36me3 enrichment of exons and the H3K36me1 enrichment of simple repeats. These results suggest that the H3K36me2 mark is required for proper transposon repression and that *Set2* is the main methyltransferase responsible for its application.

piRNA production from dual-strand piRNA clusters is preferentially triggered by the loss of H3K36 methylation

The PIWI/piRNA pathway is the main line of defense against transposon activity in the germline. It functions by degrading transcribed elements and silencing expression from transposon loci. This activity is targeted by short non-coding piRNAs containing complementary sequences of transposable elements. The piRNAs are loaded into PIWI-clade Argonaute proteins that mediate transposon silencing. PiRNAs are continuously produced in the germline but there is little data on their transcriptional activation in somatic cells. Interestingly, disturbed H3K36 methylation was recently shown to trigger piRNA production alongside

increased transposon activity (Chaouch *et al.*, 2021). However, since these results were obtained in cells overexpressing H3.3K36M, no conclusions could be drawn regarding the contributions of specific methyltransferases.

By analyzing RNA-seq data for the *H3K36R*, *Set2*, *NSD*, and *ash1* mutants, we were able to show that loss of the methylatable lysine 36 residue of Histone 3 significantly increased piRNA expression. The effect on piRNA transcriptional activation was also visible in the methyltransferase mutants and correlated positively with their levels of transposon activity (*Set2>NSD>ash1*). This suggests that the level of piRNA production in somatic cells depends on the level of transposon activation. Intriguingly, we observed higher levels of piRNAs originating from dual-strand piRNA clusters than from uni-strand clusters. This result was puzzling because transcription of dual-strand clusters requires the assembly of a germline-specific protein complex (Andersen *et al.*, 2017; Brennecke *et al.*, 2007; Mohn *et al.*, 2014). To investigate this, we performed gene-level differential expression analysis and confirmed that the four genes known to be required for dual-stand cluster expression (*rhino, deadlock, moonshiner* and *cutoff*) were transcriptionally activated in the brains of *D. melanogaster H3K36R* histone replacement mutants.

Discussion

H3K36 methylation already has a long list of functions and interactors. The addition of transposon repression further reinforces its position as the Swiss army knife of histone modifications. By taking apart the methylation machinery, we clarified the division of labor and transposon family specificity between the methyltransferases, the crucial role of Set2 dimethylation of H3K36 and the transposon dose dependent activation of dual-strand piRNA clusters in somatic cells. These findings are valuable to the community because there is a growing body of evidence linking transposon activity in somatic cells to multiple human diseases. For example, the expression of transposable elements in the human brain has been linked to several neurodevelopmental disorders (Jacob-Hirsch *et al.*, 2018; Krug *et al.*, 2017; Perrat *et al.*, 2013). In mammals, suppression of transposable elements is largely maintained by DNA methylation and the de novo methylation of DNA is established by DNA methyl transferase 3A (DNMT3A) and DNA methyl transferase 3B (DNMT3B) which are both known to be recruited to di- and trimethylated H3K36 (Morselli *et al.*, 2015; Neri *et al.*, 2017; Weinberg et al., 2019). Impairment of H3K36 methylation in the mammalian brain could thus induce transcriptional

activation of transposable elements, potentially causing neurodevelopmental disorders.

Although it was satisfying to demonstrate that methylated H3K36 has transposon-repressing functions, I felt that one major issue remained to be addressed after finalizing the manuscript for Paper III. Specifically, transposon silencing in mammalian systems depends on DNA methylation. However, although *D. melanogaster* lacks DNA methylation, H3K36 methylation still appears to mediate transposon suppression in this species. Transcription silencing in organisms lacking DNA methylation is managed by other mechanisms and an important mechanism of this type in eukaryotes is heterochromatin formation mediated by Heterochromatin Protein 1 (HP1). In the traditional model, HP1 binds methylated H3K9 and its spread across the chromosome is then facilitated by recruitment of Suppressor of variegation 3-9 (Su(var)3-9), a H3K9 methyltransferase that acts on adjacent histones to create additional HP1 binding sites. Silencing is then mediated by polymerization of closely bound HP1 monomers, leading to chromatin compaction. HP1 was also shown to bind and silence specific regions independently of initial H3K9 methylation (Figueiredo *et al.*, 2012). This prompted me to attempt to develop a plausible model to explain H3K36 methylation-mediated transposon silencing in *D. melanogaster* somatic cells.

In mammals, DNA methyltransferase 3A is recruited to H3K36me2/me3 via its PWWP domain (Weinberg *et al.*, 2019). The *D. melanogaster* genome has only 9 proteins that contain or are predicted to have a functional PWWP domain. These include the NSD H3K36 methyltransferase and JIL-1 Anchoring and Stabilizing Protein (JASper) (Albig *et al.*, 2019), both of which have been shown to bind methylated H3K36. The currently unnamed protein CG1815 also has a PWWP domain and was shown to interact with Su(var)205, the *Drosophila* ortholog of HP1 (Alekseyenko *et al.*, 2014). The closest human ortholog of CG1815, ZMYND8 (DIOPT score: 13/15), similarly has a PWWP-domain and was shown to interact with H3K36me2 (Adhikary *et al.*, 2016). Furthermore, loss of functional ZMYND8 is a high confidence marker for autism spectrum disorder (Iossifov *et al.*, 2014), which has been linked to both increased transposon activity in the brain (Misiak *et al.*, 2019) and mutations in H3K36 histone methyltransferases (Lumish *et al.*, 2015; O'Roak *et al.*, 2012a; O'Roak *et al.*, 2012b). Based on this knowledge, I proposed the existence of a DNA methylation-independent transposon silencing mechanism in which the key players are dimethylated H3K36 and its binding partner CG1815. CG1815 in turn acts as a target for the

delivery of Su(var)205, which initiates heterochromatin formation and thus represses transposon activity in somatic cells. This model, which is depicted schematically in Figure 4, remains to be tested.

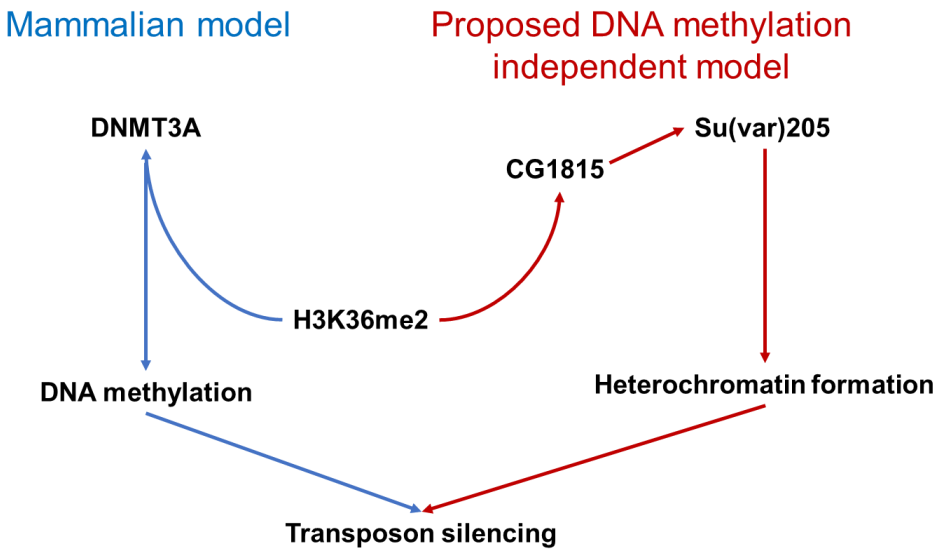

Figure 4: Hypothesized novel mechanism of DNA methylation-independent transposon silencing

KEY FINDINGS

Paper I

- *In-situ* PLA is a valuable tool enabling direct visualization and localization of potential protein/RNA interactions on polytene chromosomes.
- The interactions between the MSL complex and Jil-1, CLAMP, and Top2 were verified.
- The interaction between MSL3 and H3K36me3 could not be verified, suggesting that MSL3 has an alternative binding partner.

Paper II

- H3K36me3 is less enriched on the male X chromosome than on autosomes.
- Set2 promotes dosage compensation on the single male X chromosome independently of its H3K36 histone methyltransferase activity.
- NSD and Ash1 together maintain proper expression of genes on the small fourth chromosome.

Paper III

- Uncompromised lysine methylation of Lysine 36 on replication-dependent Histone 3 is required for transcriptional repression of transposons in somatic cells.
- Set2 is the main H3K36 methyl transferase involved in suppression of transposable elements.
- The PIWI/piRNA pathway is activated in somatic cells in a dosage-dependent manner as a response to transposon activity.
- Dual-strand piRNA clusters are preferentially expressed in H3K36me deficient somatic cells.

ACKNOWLEDGMENTS

First of all, I want to thank my wife Anna for encouraging me to embark on this scientific expedition into the unknown. Without you I would most likely have gone to medical school or decided to be a police officer. Since I am the worst at dealing with people in distress and danger, I guess it's not just me who should be thankful to you, but also all those other people in parallel universes who would have been forced to interact with these misplaced versions of me. You were also the one who had to deal with the emotional fallout of failed experiments and stress-induced tantrums before big presentations. Thank you for being my rock.

While talking about family, I also want to express gratitude to my kids. Doris, your enquiring mind and endless questions have helped me to formulate my work in an understandable way. You are probably the one at home with the best understanding of this thesis. Betsy, your relentless joy of life has taught me not to take everything so seriously. I always get stuck in my head trying to find the perfect formulation or finding the most minute errors in a text, but your "Fudge it, we'll do it live"-attitude has taught me relax and oftentimes see the bigger picture. Lennox, your unparalleled excitement about me leaving to go feed the flies has resulted in me doing this rather uninspiring task with a big stupid smile on my face. Thank you all three for being my F1 generation.

Now when that is done, let's move on to co-workers. I want to thank my supervisor Jan Larsson for the freedom to explore and go down rabbit holes sometimes leading nowhere. Your patience and guidance have been invaluable over the years and the time working in your group has been some of the best in my life. My kids also thank you for having candy in the office, which has sparked their interest in science because the slight chance of sweets was the main reason they wanted to tag along to the lab in the beginning.

Finally, I must thank all of the group members past and present who made every day a good day. Masha, Anna-Mia, Marie-Line, Karolis, Denise, Jacob, Margarida, and Lina; you've all been there for silly questions, coffee-breaks, and emotional support. Without you all I would have had more space in the office but way less fun. I'm also touched by your warm reception of me coming out as transgender; being able to be myself at work during my early days of transitioning has relieved me of so much anxiety.

REFERENCES

Adhikary S, Sanyal S, Basu M, Sengupta I, Sen S, Srivastava DK, Roy S, Das C (2016) Selective recognition of H3.1K36 dimethylation/H4K16 acetylation facilitates the regulation of all-trans-retinoic acid (ATRA)-responsive genes by putative chromatin reader ZMYND8. *J. Biol. Chem*. 291: 2664-2681

Akhtar A, Becker PB (2000) Activation of transcription through Histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* 5: 367-375

Albig C, Wang C, Dann GP, Wojcik F, Schauer T, Krause S, Maenner S, Cai W, Li Y, Girton J *et al* (2019) JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*. *Nat. Commun*. 10: 5343

Alekseyenko AA, Gorchakov AA, Zee BM, Fuchs SM, Kharchenko PV, Kuroda MI (2014) Heterochromatin-associated interactions of *Drosophila* HP1a with dADD1, HIPP1, and repetitive RNAs. *Genes Dev*. 28: 1445-1460

Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee O-K, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ *et al* (2008) A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. *Cell* 134: 599-609

Andersen PR, Tirian L, Vunjak M, Brennecke J (2017) A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* 549: 54-59

Andreev VI, Yu C, Wang J, Schnabl J, Tirian L, Gehre M, Handler D, Duchek P, Novatchkova M, Baumgartner L *et al* (2022) Panoramix SUMOylation on chromatin connects the piRNA pathway to the cellular heterochromatin machinery. *Nat. Struct. Mol. Biol*. 29: 130-142

Angrand P-O, Apiou F, Stewart AF, Dutrillaux B, Losson R, Chambon P (2001) NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics* 74: 79-88

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, Gaasterland T, Meyer J, Tuschl T (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* 5: 337-350

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Current Biology* 11: 1017-1027

Aravin AA, Sachidanandam R, Bourc'His D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* 31: 785-799

Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.* 21: 366-374

Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T (2005) Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J. Biol. Chem.* 280: 17732-17736

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837

Bell O, Conrad T, Kind J, Wirbelauer C, Akhtar A, Schübeler D (2008) Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in *Drosophila melanogaster*. *Mol. Cell. Biol.* 28: 3401-3409

Bell O, Wirbelauer C, Hild M, Scharf AND, Schwaiger M, Macalpine DM, Zilbermann F, Van Leeuwen F, Bell SP, Imhof A *et al* (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in *Drosophila*. *EMBO J.* 26: 4974-4984

Berdasco M, Ropero S, Setien F, Fraga MF, Lapunzina P, Losson R, Alaminos M, Cheung N-K, Rahman N, Esteller M (2009) Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *Proc. Natl. Acad. Sci. U.S.A.* 106: 21830-21835

Bone JR, Lavender J, Richman R, Palmer MJ, Turner BM, Kuroda MI (1994) Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev.* 8: 96-104

Brejc K, Bian Q, Uzawa S, Wheeler BS, Anderson EC, King DS, Kranzusch PJ, Preston CG, Meyer BJ (2017) Dynamic control of X chromosome conformation and repression by a histone H4K20 demethylase. *Cell* 171: 85-102.e123

Brennan K, Shin JH, Tay JK, Prunello M, Gentles AJ, Sunwoo JB, Gevaert O (2017) NSD1 inactivation defines an immune cold, DNA hypomethylated subtype in squamous cell carcinoma. *Sci. Rep.* 7: 17064

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-renerating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128: 1089-1103

Bridges CB (1935) Salivary chromosome maps: With a key to the banding of the chromosomes of *Drosophila melanogaster*. *J. Hered.* 26: 60-64

Buscaino A, Köcher T, Kind JH, Holz H, Taipale M, Wagner K, Wilm M, Akhtar A (2003) MOF-regulated acetylation of MSL3 in the *Drosophila* dosage compensation complex. *Mol. Cell* 11: 1265-1277

Capuano F, Mülleder M, Kok R, Blom HJ, Ralser M (2014) Cytosine DNA methylation is found in *Drosophila melanogaster* but absent in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and other yeast species. *Anal Chem* 86: 3697-3702

Chaouch A, Berlandi J, Chen CCL, Frey F, Badini S, Harutyunyan AS, Chen X, Krug B, Hébert S, Jeibmann A *et al* (2021) Histone H3.3 K27M and K36M mutations de-repress transposable elements through perturbation of antagonistic chromatin marks. *Mol. Cell* 81: 4876-4890.e4877

Cheng H-P, Gong D, Zhao Z-W, He P-P, Yu X-H, Ye Q, Huang C, Zhang X, Chen L-Y, Xie W *et al* (2018) microRNA-182 promotes lipoprotein lipase expression and atherogenesisby targeting Histone Deacetylase 9 in Apolipoprotein E-knockout mice. *Circulation Journal* 82: 28-38

Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. *PLoS Biol.* 6: e277

Chu C-H, Wang L-Y, Hsu K-C, Chen C-C, Cheng H-H, Wang S-M, Wu C-M, Chen T-J, Li L-T, Liu R *et al* (2014) KDM4B as a target for prostate cancer: Structural analysis and selective inhibition by a novel inhibitor. *J. Med. Chem.* 57: 5975-5985

Collord G, Martincorena I, Young MD, Foroni L, Bolli N, Stratton MR, Vassiliou GS, Campbell PJ, Behjati S (2019) Recurrent histone mutations in T-cell acute lymphoblastic leukaemia. Br. J. Haematol. 184: 676-679

Cost GJ (2002) Human L1 element target-primed reverse transcription *in vitro*. *EMBO J.* 21: 5899-5910

Cugusi S, Ramos E, Ling H, Yokoyama R, Luk KM, Lucchesi JC (2013) Topoisomerase II plays a role in dosage compensation in *Drosophila*. *Transcription* 4: 238-250

Czech B, Hannon GJ (2016) One loop to rule them all: The Ping-Pong cycle and piRNA-guided silencing. *Trends Biochem. Sci.* 41: 324-337

Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, Davies H, Edkins S, Hardy C, Latimer C *et al* (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* 463: 360-363

Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sørensen CS, Petersen NHT, Sorensen PHB, Lukas C, Bartek J *et al* (2012) LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nat. Struct. Mol. Biol.* 19: 803-810

Deng X, Hiatt JB, Nguyen DK, Ercan S, Sturgill D, Hillier LW, Schlesinger F, Davis CA, Reinke VJ, Gingeras TR *et al* (2011) Evidence for compensatory upregulation of expressed X-linked genes in mammals, *Caenorhabditis elegans* and *Drosophila melanogaster*. *Nat. Genet.* 43: 1179-1185

Difiore JV, Ptacek TS, Wang Y, Li B, Simon JM, Strahl BD (2020) Unique and shared roles for histone H3K36 methylation states in transcription regulation functions. *Cell Rep.* 31: 107751

Dorafshan E, Kahn TG, Glotov A, Savitsky M, Walther M, Reuter G, Schwartz YB (2019) Ash1 counteracts Polycomb repression independent of histone H3 lysine 36 methylation. *EMBO Rep.* 20: e46762

Eagen Kyle P, Hartl Tom A, Kornberg Roger D (2015) Stable chromosome condensation revealed by chromosome conformation capture. *Cell* 163: 934-946

Eickbush TH, Malik HS (2007) Origins and evolution of retrotransposons. In: *Mobile DNA II*, pp. 1111-1144.

Faucillion ML, Larsson J (2015) Increased expression of X-linked genes in mammals is associated with a higher stability of transcripts and an increased ribosome density. *Genome Biol. Evol.* 7: 1039-1052

Ferrari F, Plachetka A, Alekseyenko AA, Jung YL, Ozsolak F, Kharchenko PV, Park PJ, Kuroda MI (2013) "Jump start and gain" model for dosage compensation in *Drosophila* based on direct sequencing of nascent transcripts. *Cell Rep.* 5: 629-636

Figueiredo MLA, Kim M, Philip P, Allgardsson A, Stenberg P, Larsson J (2014) Non-coding *roX* RNAs prevent the binding of the MSL-complex to heterochromatic regions. *PLoS Genet.* 10: e1004865

Figueiredo MLA, Philip P, Stenberg P, Larsson J (2012) HP1a recruitment to promoters is independent of H3K9 methylation in *Drosophila melanogaster*. *PLoS Genet*. 8: e1003061

Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet*. 5: 103-107

Fnu S, Williamson EA, De Haro LP, Brenneman M, Wray J, Shaheen M, Radhakrishnan K, Lee S-H, Nickoloff JA, Hromas R (2011) Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc. Natl. Acad. Sci. U.S.A.* 108: 540-545

Franke A, Baker BS (1999) The *roX1* and *roX2* RNAs are essential components of the Compensasome, which mediates dosage compensation in *Drosophila*. *Mol. Cell* 4: 117-122

Gallach M (2014) Recurrent turnover of chromosome-specific satellites in *Drosophila*. *Genome Biol. Evol.* 6: 1279-1286

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh C-L, Zhang X, Golic KG, Jacobsen SE, Bestor TH (2006) Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2. *Science* 311: 395-398

Grabundzija I, Hickman AB, Dyda F (2018) *Helraiser* intermediates provide insight into the mechanism of eukaryotic replicative transposition. *Nat. Commun*. 9: 1278

Grozeva D, Carss K, Spasic-Boskovic O, Michael, Archer H, Helen, Park S-M, Canham N, Susan, Wilson M *et al* (2014) *De novo* loss-of-function mutations in SETD5, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. *Am. J. Hum. Genet*. 94: 618-624

Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* 315: 1587-1590

Hallacli E, Lipp M, Georgiev P, Spielman C, Cusack S, Akhtar A, Kadlec J (2012) Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in *Drosophila*. *Mol. Cell* 48: 587-600

He J, Kallin EM, Tsukada Y-I, Zhang Y (2008) The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15Ink4b. *Nat. Struct. Mol. Biol.* 15: 1169-1175

Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, Zhao BS, Mesquita A, Liu C, Yuan CL *et al* (2018) Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20: 285-295

Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, Chen Z, Deng X, Xiao G, Auer F *et al* (2019) Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nature* 567: 414-419

Huang Z, Wu H, Chuai S, Xu F, Yan F, Englund N, Wang Z, Zhang H, Fang M, Wang Y *et al* (2013) NSD2 is recruited through its PHD domain to oncogenic gene loci to drive multiple myeloma. *Cancer Res.* 73: 6277-6288

Huen MS, Huang J, Leung JW, Sy SM-H, Leung KM, Ching Y-P, Tsao SW, Chen J (2010) Regulation of chromatin architecture by the PWWP domain-containing DNA damage-responsive factor EXPAND1/MUM1. *Mol. Cell* 37: 854-864

Ilik Ibrahim A, Quinn Jeffrey J, Georgiev P, Tavares-Cadete F, Maticzka D, Toscano S, Wan Y, Spitale Robert C, Luscombe N, Backofen R *et al* (2013) Tandem stem-loops in *roX* RNAs act together to mediate X chromosome dosage compensation in *Drosophila*. *Mol. Cell* 51: 156-173

Ilyin AA, Kononkova AD, Golova AV, Shloma VV, Olenkina Oxana M, Nenasheva Valentina V, Abramov Yuri A, Kotov AA, Maksimov Daniil A, Laktionov Petr P *et al* (2022) Comparison of genome architecture at two stages of male germline cell differentiation in *Drosophila*. *Nucleic Acids Res.* 50: 3203-3225

Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE *et al* (2014) The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515: 216-221

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

Jacob-Hirsch J, Eyal E, Knisbacher BA, Roth J, Cesarkas K, Dor C, Farage-Barhom S, Kunik V, Simon AJ, Gal M *et al* (2018) Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. *Cell Res.* 28: 187-203

Jiang J, Zhao L, Yan L, Zhang L, Cao Y, Wang Y, Jiang Y, Yan T, Cao Y (2014) Structural features and mechanism of translocation of non-LTR retrotransposons in *Candida albicans*. *Virulence* 5: 245-252

Jin Y, Wang Y, Johansen J, Johansen KM (2000) JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. *J. Cell Biol.* 149: 1005-1010

Jin Y, Wang Y, Walker DL, Dong H, Conley C, Johansen J, Johansen KM (1999) JIL-1: A novel chromosomal tandem kinase implicated in transcriptional regulation in *Drosophila*. *Mol. Cell* 4: 129-135

Johansson AM, Stenberg P, Allgardsson A, Larsson J (2012) POF regulates the expression of genes on the fourth chromosome in *Drosophila melanogaster* by binding to nascent RNA. *Mol. Cell. Biol.* 32: 2121-2134

Johansson AM, Stenberg P, Bernhardsson C, Larsson J (2007) Painting of fourth and chromosome-wide regulation of the 4th chromosome in *Drosophila melanogaster*. *EMBO J.* 26: 2307-2316

Jordan W, Larschan E (2021) The zinc finger protein CLAMP promotes long-range chromatin interactions that mediate dosage compensation of the *Drosophila* male X-chromosome. *Epigenetics Chromatin* 14: 29

Kadlec J, Hallacli E, Lipp M, Holz H, Sanchez-Weatherby J, Cusack S, Akhtar A (2011) Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.* 18: 142-149

Kaplan CD, Laprade L, Winston F (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* 301: 1096-1099

Kapoor-Vazirani P, Vertino PM (2014) A dual role for the histone methyltransferase PR-SET7/SETD8 and histone H4 lysine 20 monomethylation in the local regulation of RNA polymerase II pausing. *J. Biol. Chem.* 289: 7425-7437

Keegan RM, Talbot LR, Chang Y-H, Metzger MJ, Dubnau J (2021) Intercellular viral spread and intracellular transposition of *Drosophila Gypsy*. *PLOS Genet.* 17: e1009535

Kelley RL, Wang J, Bell L, Kuroda MI (1997) Sex lethal controls dosage compensation in *Drosophila* by a non-splicing mechanism. *Nature* 387: 195-199

Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F, Khorasanizadeh S (2010) Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat. Struct. Mol. Biol.* 17: 1027-1029

Kim M, Ekhteraei-Tousi S, Lewerentz J, Larsson J (2018) The X-linked 1.688 satellite in *Drosophila melanogaster* promotes specific targeting by Painting of fourth. *Genetics* 208: 623-632

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell. Biol.* 25: 3305-3316

Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao W-W, Morrill K, Prazak L, Rozhkov N, Theodorou D, Hammell M *et al* (2017) Retrotransposon activation contributes to neurodegeneration in a *Drosophila* TDP-43 model of ALS. *PLOS Genet*. 13: e1006635

Kudithipudi S, Lungu C, Rathert P, Happel N, Jeltsch A (2014) Substrate specificity analysis and novel substrates of the protein lysine methyltransferase NSD1. *Chem. Biol.* 21: 226-237

Kuhn GCS, Küttler H, Moreira-Filho O, Heslop-Harrison JS (2011) The 1.688 repetitive DNA of *Drosophila*: Concerted evolution at different genomic scales and association with genes. *Mol. Biol. and Evol.* 29: 7-11

Kuroda MI, Hilfiker A, Lucchesi JC (2016) Dosage compensation in *Drosophila* - a model for the coordinate regulation of transcription. *Genetics* 204: 435-450

Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y *et al* (2002) Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet.* 30: 365-366

Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ, Kuroda MI (2007) MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell* 28: 121-133

Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, Park PJ, Kuroda MI (2011) X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature* 471: 115-118

Larschan E, Soruco MML, Lee O-K, Peng S, Bishop E, Chery J, Goebel K, Feng J, Park PJ, Kuroda MI (2012) Identification of chromatin-associated regulators of MSL complex targeting in *Drosophila* dosage compensation. *PLoS Genet*. 8: e1002830

Larsson J, Chen JD, Rasheva V, Rasmuson-Lestander Å, Pirrotta V (2001) Painting of fourth, a chromosome-specific protein in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 98: 6273-6278

Lee JT (2011) Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.* 12: 815-826

Lentini A, Cheng H, Noble JC, Papanicolaou N, Coucoravas C, Andrews N, Deng Q, Enge M, Reinius B (2022) Elastic dosage compensation by X-chromosome upregulation. *Nat. Commun.* 13: 1854

Li B, Gogol M, Carey M, Lee D, Seidel C, Workman JL (2007) Combined action of PHD and Chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* 316: 1050-1054

Li B, Jackson J, Simon MD, Fleharty B, Gogol M, Seidel C, Workman JL, Shilatifard A (2009a) Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *J. Biol. Chem.* 284: 7970-7976

Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li G-M (2013) The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSα. *Cell* 153: 590-600

Li Y, Dooner HK (2009) Excision of Helitron transposons in maize. *Genetics* 182: 399-402

Li Y, Trojer P, Xu C-F, Cheung P, Kuo A, Drury WJ, Qiao Q, Neubert TA, Xu R-M, Gozani O *et al* (2009b) The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem.* 284: 34283-34295

Lim AK, Kai T (2007) Unique germline organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 104: 6714-6719

Lu C, Jain SU, Hoelper D, Bechet D, Molden RC, Ran L, Murphy D, Venneti S, Hameed M, Pawel BR *et al* (2016) Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* 352: 844-849

Lumish HS, Wynn J, Devinsky O, Chung WK (2015) Brief Report: SETD2 mutation in a child with autism, intellectual disabilities and epilepsy. *J. Autism Dev. Disord.* 45: 3764-3770

Lundberg LE, Kim M, Johansson A-M, Faucillion M-L, Josupeit R, Larsson J (2013) Targeting of Painting of fourth to *roX1* and *roX2* proximal sites suggests evolutionary links between dosage compensation and the regulation of the fourth chromosome in *Drosophila melanogaster*. *G3: Genes Genomes Genet.* 3: 1325-1334

Luscan A, Laurendeau I, Malan V, Francannet C, Odent S, Giuliano F, Lacombe D, Touraine R, Vidaud M, Pasmant E *et al* (2014) Mutations in SETD2 cause a novel overgrowth condition. *J. Med. Genet.* 51: 512-517

Maenner S, Müller M, Fröhlich J, Langer D, Peter (2013) ATP-dependent *roX* RNA remodeling by the helicase maleless enables specific association of MSL proteins. *Mol. Cell* 51: 174-184

Maltby VE, Martin BJ, Schulze JM, Johnson I, Hentrich T, Sharma A, Kobor MS, Howe L (2012) Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. *Mol. Cell. Biol.* 32: 3479-3485

Meller VH, Gordadze PR, Park Y, Chu X, Stuckenholz C, Kelley RL, Kuroda MI (2000) Ordered assembly of *roX* RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila*. *Curr. Biol.* 10: 136-143

Metcalfe CJ, Casane D (2014) Modular organization and reticulate evolution of the ORF1 of Jockey superfamily transposable elements. *Mob. DNA* 5: 19

Meyer BJ (2022) The X chromosome in C. elegans sex determination and dosage compensation. *Curr. Opin. Genet. Dev.* 74: 101912

Mirzoyan Z, Sollazzo M, Allocca M, Valenza AM, Grifoni D, Bellosta P (2019) *Drosophila melanogaster*: A model organism to study cancer. *Front. Genet.* 10: 51

Mishra S, Van Rechem C, Pal S, Clarke TL, Chakraborty D, Mahan SD, Black JC, Murphy SE, Lawrence MS, Daniels DL *et al* (2018) Cross-talk between lysine-modifying enzymes controls site-specific DNA amplifications. *Cell* 174: 803-817.e816

Misiak B, Ricceri L, Sąsiadek MM (2019) Transposable elements and their epigenetic regulation in mental disorders: Current evidence in the field. *Front. Genet.* 10: 580

Mohn F, Sienski G, Handler D, Brennecke J (2014) The Rhino-Deadlock-Cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. *Cell* 157: 1364-1379

Moore Kaitlyn E, Carlson Scott M, Camp Nathan D, Cheung P, James Richard G, Chua Katrin F, Wolf-Yadlin A, Gozani O (2013) A general molecular affinity strategy for global Ddetection and proteomic analysis of lysine methylation. *Mol. Cell* 50: 444-456

Moore SA, Ferhatoglu Y, Jia Y, Al-Jiab RA, Scott MJ (2010) Structural and biochemical studies on the Chromo-barrel domain of Male Specific Lethal 3

(MSL3) reveal a binding preference for mono- or dimethyllysine 20 on histone H4. *J. Biol. Chem.* 285: 40879-40890

Morales V, Straub T, Neumann MF, Mengus G, Akhtar A, Becker PB (2004) Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *EMBO J.* 23: 2258-2268

Morselli M, Pastor WA, Montanini B, Nee K, Ferrari R, Fu K, Bonora G, Rubbi L, Clark AT, Ottonello S *et al* (2015) In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *eLife* 4: e06205

Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde M-E, Hong Z, Allen C, Roy S, Nuñez JK, Nickoloff J *et al* (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat. Struct. Mol. Biol.* 19: 1266-1272

Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, Oliviero S (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* 543: 72-77

Nishimasu H, Ishizu H, Saito K, Fukuhara S, Kamatani MK, Bonnefond L, Matsumoto N, Nishizawa T, Nakanaga K, Aoki J *et al* (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* 491: 284-287

O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K *et al* (2012a) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338: 1619-1622

O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD *et al* (2012b) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485: 246-250

Ohno S (2013) *Sex chromosomes and sex-linked genes*. Springer Science & Business Media

Olivieri D, Sykora MM, Sachidanandam R, Mechtler K, Brennecke J (2010) An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. *EMBO J.* 29: 3301-3317

Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, Ailles L, Karamchandani J, Marchione DM, Garcia BA *et al* (2017) Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nat. Genet.* 49: 180-185

Park S-W, Kuroda MI, Park Y (2008) Regulation of histone H4 Lys16 acetylation by predicted alternative secondary structures in *roX* noncoding RNAs. *Mol. Cell. Biol.* 28: 4952-4962

Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S (2013) Transposition-Driven Genomic Heterogeneity in the *Drosophila* Brain. *Science* 340: 91-95

Pfister Sophia X, Ahrabi S, Zalmas L-P, Sarkar S, Aymard F, Bachrati Csanád Z, Helleday T, Legube G, La Thangue Nicholas B, Porter Andrew CG *et al* (2014) SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.* 7: 2006-2018

Post C, Clark JP, Sytnikova YA, Chirn G-W, Lau NC (2014) The capacity of target silencing by *Drosophila* PIWI and piRNAs. *RNA* 20: 1977-1986

Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA (2012) Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLOS Genet.* 8: e1002717

Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, Komatsu Y, Shinkai Y, Cheng X, Jeltsch A (2008) Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* 4: 344-346

Rayasam GV (2003) NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J.* 22: 3153-3163

Regnard C, Straub T, Mitterweger A, Dahlsveen IK, Fabian V, Becker PB (2011) Global analysis of the relationship between JIL-1 kinase and transcription. *PLOS Genet.* 7: e1001327

Reiter LT, Potocki L, Chien S, Gribskov M, Bier E (2001) A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11: 1114-1125

Richter L, Bone JR, Kuroda MI (1996) RNA-dependent association of the *Drosophila* maleless protein with the male X chromosome. *Genes Cells* 1: 325-336

Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, Lyko F (2010) RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev.* 24: 1590-1595

Schuhmacher MK, Beldar S, Khella MS, Bröhm A, Ludwig J, Tempel W, Weirich S, Min J, Jeltsch A (2020) Sequence specificity analysis of the SETD2 protein

lysine methyltransferase and discovery of a SETD2 super-substrate. *Communications Biology* 3: 511

Sen P, Dang W, Donahue G, Dai J, Dorsey J, Cao X, Liu W, Cao K, Perry R, Lee JY *et al* (2015) H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev.* 29: 1362-1376

Shoaib M, Chen Q, Shi X, Nair N, Prasanna C, Yang R, Walter D, Frederiksen KS, Einarsson H, Svensson JP *et al* (2021) Histone H4 lysine 20 mono-methylation directly facilitates chromatin openness and promotes transcription of housekeeping genes. *Nat. Commun.* 12: 4800

Smit A, Hubley R, Green P, 2015. RepeatMasker Open-4.0. 2013–2015.

Smith ER, Allis CD, Lucchesi JC (2001) Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males. *J. Biol. Chem.* 276: 31483-31486

Smith ER, Cayrou C, Huang R, Lane WS, Côté J, Lucchesi JC (2005) A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol. Cell. Biol.* 25: 9175-9188

Soruco MML, Larschan E (2014) A new player in X identification: the CLAMP protein is a key factor in *Drosophila* dosage compensation. *Chromosome Res.* 22: 505-515

Stenberg P, Lundberg LE, Johansson A-M, Rydén P, Svensson MJ, Larsson J (2009) Buffering of segmental and chromosomal aneuploidies in *Drosophila melanogaster*. *PLoS Genet.* 5: e1000465

Straub T, Grimaud C, Gilfillan GD, Mitterweger A, Becker PB (2008) The chromosomal high-affinity binding sites for the *Drosophila* dosage compensation complex. *PLoS Genet.* 4: e1000302

Straub T, Zabel A, Gilfillan GD, Feller C, Becker PB (2013) Different chromatin interfaces of the *Drosophila* dosage compensation complex revealed by high-shear ChIP-seq. *Genome Res.* 23: 473-485

Su X, Zhang J, Mouawad R, Compérat E, Rouprêt M, Allanic F, Parra J, Bitker M-O, Thompson EJ, Gowrishankar B *et al* (2017) NSD1 inactivation and SETD2 mutation drive a convergence toward loss of function of H3K36 writers in clear cell renal cell carcinomas. *Cancer Res.* 77: 4835-4845

Sural TH, Peng S, Li B, Workman JL, Park PJ, Kuroda MI (2008) The MSL3 chromodomain directs a key targeting step for dosage compensation of the *Drosophila melanogaster* X chromosome. *Nat. Struct. Mol. Biol.* 15: 1318-1325

Suzuki T, Ozasa H, Itoh Y, Zhan P, Sawada H, Mino K, Walport L, Ohkubo R, Kawamura A, Yonezawa M *et al* (2013) Identification of the KDM2/7 histone lysine demethylase subfamily inhibitor and its antiproliferative activity. *J. Med. Chem.* 56: 7222-7231

Swaroop A, Oyer JA, Will CM, Huang X, Yu W, Troche C, Bulic M, Durham BH, Wen QJ, Crispino JD *et al* (2019) An activating mutation of the NSD2 histone methyltransferase drives oncogenic reprogramming in acute lymphocytic leukemia. *Oncogene* 38: 671-686

Takayama S, Dhahbi J, Roberts A, Mao G, Heo S-J, Pachter L, Martin DIK, Boffelli D (2014) Genome methylation in *D. melanogaster* is found at specific short motifs and is independent of DNMT2 activity. *Genome Res.* 24: 821-830

Tanaka Y, Katagiri Z-i, Kawahashi K, Kioussis D, Kitajima S (2007) Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* 397: 161-168

Tatton-Brown K, Seal S, Ruark E, Harmer J, Ramsay E, del Vecchio Duarte S, Zachariou A, Hanks S, O'Brien E, Aksglaede L (2014) Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat. Genet.* 46: 385-388

Teo RYW, Anand A, Sridhar V, Okamura K, Kai T (2018) Heterochromatin protein 1a functions for piRNA biogenesis predominantly from pericentric and telomeric regions in *Drosophila*. *Nat. Commun.* 9: 1735

Tikhonova E, Fedotova A, Bonchuk A, Mogila V, Larschan EN, Georgiev P, Maksimenko O (2019) The simultaneous interaction of MSL2 with CLAMP and DNA provides redundancy in the initiation of dosage compensation in *Drosophila* males. *Development* 146: dev179663

Tikhonova E, Mariasina S, Efimov S, Polshakov V, Maksimenko O, Georgiev P, Bonchuk A (2022) Structural basis for interaction between CLAMP and MSL2 proteins involved in the specific recruitment of the dosage compensation complex in *Drosophila*. *Nucleic Acids Res.* 50: 6521-6531

Tripoulas N, LaJeunesse D, Gildea J, Shearn A (1996) The *Drosophila ash1* gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. *Genetics* 143: 913-928

Turner BM, Birley AJ, Lavender J (1992) Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* 69: 375-384

Vatolina TY, Boldyreva LV, Demakova OV, Demakov SA, Kokoza EB, Semeshin VF, Babenko VN, Goncharov FP, Belyaeva ES, Zhimulev IF (2011) Identical functional organization of nonpolytene and polytene chromosomes in *Drosophila melanogaster*. *PLOS ONE* 6: e25960

Vicoso B, Bachtrog D (2013) Reversal of an ancient sex chromosome to an autosome in Drosophila. *Nature* 499: 332-335

Vicoso B, Bachtrog D (2015) Numerous transitions of sex chromosomes in Diptera. *PLOS Biol.* 13: e1002078

Villa R, Schauer T, Smialowski P, Straub T, Becker PB (2016) PionX sites mark the X chromosome for dosage compensation. *Nature* 537: 244-248

Vourekas A, Zheng K, Fu Q, Maragkakis M, Alexiou P, Ma J, Pillai RS, Mourelatos Z, Wang PJ (2015) The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. *Genes Dev.* 29: 617-629

Wagner KW, Alam H, Dhar SS, Giri U, Li N, Wei Y, Giri D, Cascone T, Kim J-H, Ye Y *et al* (2013) KDM2A promotes lung tumorigenesis by epigenetically enhancing ERK1/2 signaling. *J. Clin. Investig.* 123: 5231-5246

Wang CI, Alekseyenko AA, LeRoy G, Elia AEH, Gorchakov AA, Britton L-MP, Elledge SJ, Kharchenko PV, Garcia BA, Kuroda MI (2013) Chromatin proteins captured by ChIP–mass spectrometry are linked to dosage compensation in *Drosophila*. *Nat. Struct. Mol. Biol.* 20: 202-209

Wang GG, Cai L, Pasillas MP, Kamps MP (2007) NUP98–NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat. Cell Biol.* 9: 804-812

Wang X, Feng J, Xue Y, Guan Z, Zhang D, Liu Z, Gong Z, Wang Q, Huang J, Tang C *et al* (2016) Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex. *Nature* 534: 575-578

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G (2014) N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505: 117-120

Wang X, Zhao Boxuan S, Roundtree Ian A, Lu Z, Han D, Ma H, Weng X, Chen K, Shi H, He C (2015) N6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* 161: 1388-1399

Wang Y, Zhang W, Jin Y, Johansen J, Johansen KM (2001) The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* 105: 433-443

Waring GL, Pollack JC (1987) Cloning and characterization of a dispersed, multicopy, X chromosome sequence in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 84: 2843-2847

Weinberg DN, Papillon-Cavanagh S, Chen H, Yue Y, Chen X, Rajagopalan KN, Horth C, McGuire JT, Xu X, Nikbakht H *et al* (2019) The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature* 573: 281-286

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O *et al* (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8: 973-982

Wilhelm M, Wilhelm FX (2001) Reverse transcription of retroviruses and LTR retrotransposons. *Cell. Mol. Life Sci.* 58: 1246-1262

Wu L, Zee BM, Wang Y, Garcia BA, Dou Y (2011) The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation. *Mol. Cell.* 43: 132-144

Xu W, Li J, Rong B, Zhao B, Wang M, Dai R, Chen Q, Liu H, Gu Z, Liu S *et al* (2020) DNMT3A reads and connects histone H3K36me2 to DNA methylation. *Protein Cell* 11: 150-154

Yano S, Ishiuchi T, Abe S, Namekawa SH, Huang G, Ogawa Y, Sasaki H (2022) Histone H3K36me2 and H3K36me3 form a chromatin platform essential for DNMT3A-dependent DNA methylation in mouse oocytes. *Nat. Commun.* 13: 4440

Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B (2011) H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* 286: 7983-7989