Targeting of proteins to chromatin in *Drosophila melanogaster*

von

Christian Albig

Jahr 2019

Aus dem Adolf-Butenandt-Institut im Biomedizinischen Centrum
der Ludwig-Maximilians-Universität München
Lehrstuhl für Molekularbiologie
Vorstand: Prof. Dr. rer. nat. Peter B. Becker

Targeting of proteins to chromatin in *Drosophila melanogaster*

Dissertation
zum Erwerb des Doktorgrades der Naturwissenschaften
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von
Christian Albig
aus Tübingen
Jahr 2019

Mit Genehmigung der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

Betreuer: Prof. Dr. rer. nat. Peter B. Becker
Zweitgutachter: Prof. Andreas Ladurner, PhD

Dekan: Prof. Dr. med. dent. Reinhard Hickel
Tag der mündlichen Prüfung: 21.01.2020

Eidesstattliche Versicherung

Albig, Christian

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

„Targeting of proteins to chromatin in *Drosophila melanogaster*“

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weitern, dass die hier vorliegende Dissertation nicht in der gleichen oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Gilching, 27.02.2020	Christian Albig
Ort, Datum	Unterschrift Doktorandin/Doktorand

für Pauline

Table of contents

Preamble 9

List of publications 10

Declaration of contributions 11

Summary 13

Zusammenfassung 14

1 Introduction 19

1.1 Chromatin 19

1.1.1 The nucleosome and higher-order chromatin structures 20

1.1.2 Histone variants and modifications 24

1.2 Chromatin binding factors 27

1.2.1 Cooperative DNA binding 28

1.2.2 DNA binding through CXC domains 29

1.2.3 DNA binding through C2H2 zinc finger domains 30

1.2.4 Nucleosome binding through PWWP domains 32

1.3 Sex chromosome dosage compensation 34

1.3.1 Dosage compensation in *Drosophila melanogaster* 36

1.3.2 The Dosage Compensation Complex 37

1.3.3 DCC targeting to the X chromosome 40

1.3.4 Transcription up-regulation by the DCC 42

1.4 H3K36me3 chromatin 42

1.4.1 The chromosomal kinase JIL-1 45

2 Aims and objectives 48

3 Results 50

3.1 Factor cooperation for chromosome discrimination in *Drosophila* 50

3.2 Chromosome topology guides the *Drosophila* Dosage Compensation Complex for target gene activation 97

3.3 JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila* 133

4 References 205
5 Table of figures 228
6 Abbreviations 230
7 Danksagung 232
8 Lebenslauf 234

Preamble

This cumulative PhD thesis presents two original research articles on the targeting mechanism of the dosage compensation complex to the male X chromosome and its transfer to target genes in *Drosophila melanogaster*. The supplementary files of each research article are included. Furthermore, one manuscript for an original research article submitted to a peer-reviewed journal and deposited on a public preprint server on the targeting mechanism and function of the chromosomal kinase JIL-1 in *D. melanogaster* is presented. The contributions of all co-authors to each research article are stated here and before each article presented in the chapter 'Results'. The lists of references and abbreviations for the chapter 'Introduction' can be found at the end of this thesis.

List of publications

Albig, C., Tikhonova, E., Krause, S., Maksimenko, O., Regnard, C., and Becker, P.B. (2019). Factor cooperation for chromosome discrimination in *Drosophila*. Nucleic acids research *47*, 1706-1724, DOI 10.1093/nar/gky1238.

Schauer, T., Ghavi-Helm, Y., Sexton, T., Albig, C., Regnard, C., Cavalli, G., Furlong, E.E., and Becker, P.B. (2017). Chromosome topology guides the *Drosophila* Dosage Compensation Complex for target gene activation. EMBO Reports *18*, 1854-1868, DOI 10.15252/embr.201744292.

Albig, C., Wang, C., Dann, G.P., Wojcik, F., Schauer, T., Krause, S., Maenner, S., Cai, W., Li, Y., Girton, J., Muir, T.W., Johansen, J., Johansen K.M., Becker, P.B., and Regnard, C. (2019). JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*. (under revision at Nature Communications) bioRxiv *624023*, DOI 10.1101/624023. [21.01.2020: Nat Communications *10*, 5343 (2019). DOI 10.1038/s41467-019-13174-6]

Declaration of contributions

Authors' contributions to 'Factor cooperation for chromosome discrimination in *Drosophila*':

C.A., C.R. and P.B.B. conceived the study; C.A. performed all experiments except for co-immunoprecipitation and Yeast Two-Hybrid experiments; S.K. expressed proteins and performed co-immunoprecipitation experiments; E.T. and O.M. conceived, performed and analyzed Yeast Two-Hybrid experiments; C.A. performed bioinformatics analyses; All authors analyzed data; C.R. and P.B.B. provided feedback and supervision; C.A., C.R. and P.B.B. wrote the manuscript; P.B.B. secured funding.

Authors' contributions to 'Chromosome topology guides the *Drosophila* Dosage Compensation Complex for target gene activation':

The 4C-seq, ChIP-seq and RNA-seq experiments with Kc cells were performed by T.Sc.; 4C-seq was designed and supervised by Y.G.-H. and E.E.M.F.; Hi-C was performed by T.Se. and supervised by G.C.; ChIP-seq on H3K36me3 was carried out by C.A.; RNA-seq in S2 cells was performed by C.R.; Datasets were analyzed and visualized by T.Sc.; The study was conceptualized and supervised by P.B.B.; T.Sc. and P.B.B. wrote the manuscript with contributions from all authors.

Authors' contributions to 'JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*':

C.R. conceived this study and performed experiments; C.A. performed MNase and sonication ChIP-seq experiments and all bioinformatics analysis also with support from T.S.; C.W. generated and characterized the cw2 mutant line with help from J.G., W.C. did the LacO-LacI targeting experiments with support from Y.L., and J.J. and K.M.J. supervised the work and secured funding; G.P.D. performed mononucleosome library experiments and F.W. generated the mononucleosomes and arrays for the kinase assays in T.W.M.'s lab; S.K. prepared recombinant proteins for all *in vitro* assays, RNA-seq experiments under the supervision of C.R. and spike-in ChIP-seq experiments under supervision of C.A.; All authors analyzed data; C.R. and C.A. wrote the manuscript with contributions from all authors; P.B.B. secured funding and established collaborations.

Summary

Dosage compensation of sex chromosomes in *Drosophila melanogaster* is an excellent model system to study various aspects of targeting of protein factors to chromatin. Dosage compensation prevents male lethality by up-regulating transcription from the single male X chromosome in the ~2-fold range to match the two active X chromosomes in females [reviewed in e.g. (Ferrari et al., 2014; Kuroda et al., 2016; Samata and Akhtar, 2018)]. This up-regulation is facilitated by the male-specific-lethal (MSL) dosage compensation complex (DCC). The DCC binds selectively to ~300 high affinity sites (HAS) on the X chromosome, containing a low-complexity GAGA-rich sequence motif, the MSL recognition element (MRE) (Alekseyenko et al., 2008; Straub et al., 2008). However, the DCC neglects thousands of other similar sequences in the genome outside of HAS. The DNA-binding subunit MSL2 alone can enrich X chromosomal MREs *in vitro*, although MSL2 misses most MREs within HAS (Villa et al., 2016). The Chromatin-Linked Adaptor for MSL Proteins (CLAMP) binds thousands of MREs genome-wide and contributes to DCC targeting to HAS (Kaye et al., 2018; Soruco et al., 2013). The role of CLAMP in facilitating MSL2 targeting to HAS was investigated by several approaches. Monitoring MSL2 chromatin binding *in vivo* by chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) showed the requirement of CLAMP for HAS targeting. Next, the interplay between CLAMP and MSL2 in genome-wide *in vitro* DNA binding was studied by DNA immunoprecipitation with high-throughput sequencing (DIP-seq) (Gossett and Lieb, 2008; Liu et al., 2005; Villa et al., 2016). The data revealed mutual recruitment of both factors to each other's binding sites and cooperative binding to novel sites. This DNA binding cooperativity extended each other's binding repertoire to facilitate robust binding of MREs located within HAS, although increased binding to other non-functional sites was observed. Both factors interacted directly with each other in co-IP experiments, providing an explanation for cooperative DNA binding. Whether CLAMP and MSL2 are required for keeping HAS nucleosome-free was studied by assay for transposase accessibly chromatin with high-throughput sequencing (ATAC-seq) (Buenrostro et al., 2013; Buenrostro et al., 2015). Both factors cooperate to stabilize each other's binding and to compete with nucleosome positioning at HAS.

After successful binding of the DCC to HAS, it interacts with neighboring target genes, which are marked by trimethylation of histone H3K36 (H3K36me3). There, the DCC catalyzes acetylation of H4K16 (H4K16ac) to boost transcription (Akhtar and Becker, 2000; Gelbart et al., 2009; Larschan et al., 2007; Prestel et al., 2010). The DCC employs the chromosome 3D organization, which seems to be invariant between males and females, to transfer from HAS to active genes (Ramirez et al., 2015; Ulianov et al., 2016). The contribution of HAS to the chromosome interaction network was studied by using different chromosome conformation

capture techniques. Hi-C analysis on sex-sorted embryos showed that, H4K16ac and H3K36me3 correlate well with the active compartments (Sexton et al., 2012). Interestingly, compartment switching on the X chromosome between males and females was correlated with H4K16ac and therefore attributed to dosage compensation. The involvement of the Pioneering sites on the X (PionX), a special sub-class of HAS, in chromosome architecture was studied by high-resolution 4C-seq in male and female cells. Chromosomal segments containing PionX made frequent contact with many loci within the active compartment and even looped over large domains of the inactive compartment (Ghavi-Helm et al., 2014). These long-range interactions between PionX with other PionX/HAS were more robust in males compared to females, indicating that the dosage compensation machinery reinforced them. Moreover, *de novo* induction of DCC assembly in female cells showed that the DCC uses long-range interaction within the active compartment to transfer from PionX to target genes marked by H3K36me3 for up-regulation of transcription.

The chromosomal kinase JIL-1, which catalyzes phosphorylation of histone H3S10, localizes also to actively transcribed genes marked by H3K36me3 and is two-fold enriched on the male X chromosome (Jin et al., 2000; Regnard et al., 2011; Wang et al., 2001). JIL-1 is implicated in maintaining overall chromosome organization and preventing the spreading of heterochromatin into the euchromatic part of the X chromosome in both sexes (Cai et al., 2014; Ebert et al., 2004; Jin et al., 1999). Furthermore, JIL-1 localizes to the non-LTR retrotransposon arrays of the telomeres to positively regulate their expression (Andreyeva et al., 2005; Silva-Sousa and Casacuberta, 2013; Silva-Sousa et al., 2012). The role of JIL-1 in regulating gene expression was studied using various methods. JIL-1 formed a stable complex with the novel PWWP domain-containing protein, JIL-1 Anchoring and Stabilizing Protein (JASPer). The JIL-1-JASPer (JJ)-complex specifically enriched H3K36me3 modified nucleosomes *in vitro* via JASPer's PWWP domain from a nucleosome library containing 115 different nucleosome types. Consistently, ChIP-seq experiments showed that the JJ-complex localizes to H3K36me3 chromatin at active gene bodies and at telomeric transposons *in vivo*. As previously described, the JJ-complex is also enriched on the male X chromosome relative to autosomes. Loss of JIL-1 resulted in loss of JASPer enrichment, a small increase in H3K9me2 and a decrease in H4K16ac on the X chromosome shown by spike-in ChIP-seq. Gene expression analysis by RNA-seq showed that the JJ-complex positively regulates expression of genes, in particular of genes from the male X chromosome, and of telomeric transposons. Furthermore, the JJ-complex associated with the Set1/COMPASS complex and with other remodelling complexes as shown by co-IP coupled to mass spectrometry analysis.

Zusammenfassung

Dosiskompensation der Geschlechtschromosomen in *Drosophila melanogaster* ist ein exzellentes Modelsystem um verschiedene Aspekte der gezielten Bindung von Proteinfaktoren an Chromatin zu studieren. Dosiskompensation verhindert die Letalität in Männchen, durch Hochregulierung der Transkription des einfachen männlichen X Chromosoms im ungefähr zweifachen Bereich um es den beiden aktiven X Chromosomen der Weibchens anzupassen [rezensiert in z.B. (Ferrari et al., 2014; Kuroda et al., 2016; Samata and Akhtar, 2018)]. Diese Hochregulation wird durch den männerspezifisch letalen (MSL) Dosiskompensationskomplex (DCC) bewerkstelligt. Der DCC bindet selektiv an ~300 hochaffine Stelle (HAS) auf dem X Chromosom, welche ein niederkomplexes, GAGA-reiches Sequenzmotiv beinhalten, das MSL Erkennungselement (MRE) (Alekseyenko et al., 2008; Straub et al., 2008). Jedoch missachtet der DCC tausende ähnliche Sequenzen im Genom, die außerhalb von HAS liegen. Die DNA bindende Untereinheit MSL2 kann *in vitro* alleine X chromosomale MREs anreichern, wenngleich MSL2 die meisten MREs in HAS auslässt (Villa et al., 2016). Der chromatingebundene Adapter für MSL Proteine (CLAMP) bindet tausende MREs genomweit und trägt zur gezielten Bindung des DCC an HAS bei (Kaye et al., 2018; Soruco et al., 2013). Die Beitrag von CLAMP zur MSL2 Bindung an HAS wurde durch mehrere Ansätze untersucht. Messung der MSL2 Chromatinbindung *in vivo* durch Chromatin-Immunopräzipitation mit Hochdurchsatzsequenzierung (ChIP-seq) zeigte, dass CLAMP für die HAS-Bindung benötigt wird. Danach wurde das Zusammenspiel von CLAMP und MSL2 bei der genomweiten DNA-Bindung *in vitro* durch DNA-Immunopräzipitation mit Hochdurchsatzsequenzierung (DIP-seq) sudiert (Gossett and Lieb, 2008; Liu et al., 2005; Villa et al., 2016). Die Daten zeigten eine gegenseitige Anreicherung beider Faktoren an den jeweils anderen Bindestellen und eine kooperative Binding von neuen Stellen. Diese DNA-Bindekooperativität vergrößerte das Bindungsrepertoire des jeweils Anderen um eine robuste Bindung der MREs in HAS zu ermöglichen, obgleich eine verstärkte Bindung von andere nicht-funktionelle Stellen beobachtet wurde. Beide Faktoren interagierten direkt mit einander in Ko-IP-Experimenten, was eine Erklärung für die kooperative DNA-Bindung liefert. Ob CLAMP und MSL2 benötigt werden um HAS nukleosomfrei zu halten wurde durch den Assay für tranposasezugängliches Chromatin mit Hochdurchsatzsequenzierung (ATAC-seq) studiert. (Buenrostro et al., 2013; Buenrostro et al., 2015). Beide Faktore kooperieren um ihre Bindung gegenseitig zu stabilisieren und mit Nukleosompositionierung an HAS zu konkurrieren.

Nach erfolgreicher Bindung des DCC an HAS geht dieser zu benachbarten Zielgenen über, welche durch Trimethylierung von Histon H3K36 (H3K36me3) markiert sind. Dort katalysiert der DCC die Acetylierung von H4K16 (H4K16ac) um die Transkription zu erhöhen (Akhtar and Becker, 2000; Gelbart et al., 2009; Larschan et al., 2007; Prestel et al., 2010). Der DCC bedient sich dabei dem Chromosomeninteraktionsnetzwerk um von HAS zu aktiven Genen

zu gelangen, welches invariant zwischen Männchen und Weibchen zu seien scheint (Ramirez et al., 2015; Ulianov et al., 2016). Der Beitrag der HAS zum Chromosomeninteraktionsnetzwerk wurde durch die Benutzung verschiedene Chromosomenkonformationserfassungstechniken studiert. Hi-C-Analysen von nach Geschlechtern sortierten Embryos zeigte, dass H4K16ac und H3K36me3 gut mit dem aktiven Kompartiment korrelieren (Sexton et al., 2012). Interessanterweise ist ein Kompartimentwechsel auf dem X Chromosom zwischen Männchen und Weibchen mit H4K16ac korreliert und wird deshalb der Dosiskompensation zu geschrieben. Die Beteiligung der Pionierstellen auf dem X (PionX), einer besonderen Subklasse der HAS, zur Chromosomenarchitektur wurde durch hochauflösendes 4C-seq in männlichen und weiblichen Zellen studiert. Dies zeigte, dass PionX-beinhaltende Chromosomensegmente häufigeren Kontakt zu zahlreichen anderen Loci innerhalb des aktiven Kompartiments haben und sogar über große Domänen des inaktiven Kompartiments hinweg Schleifen bilden. (Ghavi-Helm et al., 2014). Diese weit reichenden Interaktionen zwischen PionX mit anderen PionX/HAS sind robuster in Männchen verglichen zu Weibchen, dies deutet darauf hin, dass diese durch die Dosiskompensationsmaschinerie verstärkt werden. Darüberhinaus zeigte *de novo* Induktion der DCC-Assemblierung in weiblichen Zellen, dass der DCC weitreichende Interaktionen innerhalb des aktiven Kompartimentes nutzt um von PionX zu Zielgenen zu transferieren, welche durch H3K36me3 markiert sind, zur Hochregulierung der Transkription.

Die chromosomale Kinase JIL-1, welche die Phosphorylierung von Histone H3S10 katalysiert, lokalisiert auch an aktiv transkribierte Gene, die durch H3K36me3 markiert sind, und ist ungefähr zweifach auf dem männlichen X Chromosom angereichert (Jin et al., 2000; Regnard et al., 2011; Wang et al., 2001). Generell trägt JIL-1 zur Aufrechterhaltung der allgemeinen Chromosomenorganisation bei und verhindert die Ausbreitung von Heterochromatin in den euchromatischen Teil des X Chromosoms in beiden Geschlechtern (Cai et al., 2014; Ebert et al., 2004; Jin et al., 1999). Zudem lokalisiert JIL-1 an die nicht-LTR Retrotransposonansammlung der Telomere um deren Expression positive zu regulieren (Andreyeva et al., 2005; Silva-Sousa and Casacuberta, 2013; Silva-Sousa et al., 2012). Die Rolle von JIL-1 bei der Genexpressionsregulation wurde unter Benutzung verschiedener Methoden studiert. JIL-1 formt mit dem neuen Protein JIL-1 verankerndes und stabilisierendes Protein (JASPer), das eine PWWP-Domäne enthält, einen stabilen Komplex, der JIL-1-JASPer (JJ)-Komplex. Der JJ-Komplex erkennt spezifisch H3K36me3 modifizierte Nucleosome *in vitro* über JASPer's PWWP-Domäne, was durch die Benutzung einer Nucleosomenbibliothek gezeigt wurde, die 115 verschieden modifizierte Nucleosome enthält. Konsistenter weise lokalisiert der JJ-Komplex an H3K36me3-Chromatin an aktive Genkörper und telomerische Transposons *in vivo* und ist auf dem männlichen X Chromosom

angereichert, dies wurde mit ChIP-seq analysiert. Die Anreicherung hängt direkt von JIL-1 ab, was durch beigespickten ChIP-seq gezeigt wurde. Verlust von JIL-1 führt zu einer kleinen Zunahme von H3K9me2 und Abnahme von H4K16ac auf dem männlichen X Chromosom. Genexpressionsanalyse durch RNA-seq zeigte, dass der JJ-Komplex die Genexpression positiv reguliert, insbesondere vom X Chromosom, und von telomerischen Transposons. Außerdem assoziiert der JJ-Komplex mit dem Set1/COMPASS-Komplex und mit anderen Remodelulierungskomplexen, was durch Ko-IP gekoppelt an Massenspektrometrieanalyse gezeigt wurde.

1 Introduction

Reproduction within all kingdoms of life requires inheritance of genetic information encoded in the genomic deoxyribonucleic acid (DNA) between generations (Avery et al., 1944). In eukaryotic cells, a rather small nucleus accommodates long DNA molecules, the chromosomes, in the form of chromatin. While most cells within a higher multicellular organism harbor the identical genetic information, many functional different tissues and cell types must be formed. This cell type specific information is stored in the epigenome [reviewed in e.g. (Rivera and Ren, 2013; Yadav et al., 2018)]. Sequencing of the human genome in the 1990s and early 2000s was still heavily time and cost consuming (International Human Genome Sequencing Consortium, 2004; International Human Genome Sequencing Consortium et al., 2001; Venter et al., 2001). Recent advances in sequencing technologies gave rise to time and cost effective next generation sequencing (NGS) approaches allowing the sequencing of millions of DNA molecules in parallel in a high-throughput format [reviewed in e.g. (Goodwin et al., 2016; Levy and Myers, 2016)]. Approaches by big consortia like the '1000 Genomes Project', 'Encode', and 'modEncode' projects sequenced thousands of human genomes and localized numerous chromatin marks and proteins (Celniker et al., 2009; Genomes Project et al., 2015; Stamatoyannopoulos et al., 2012; Sudmant et al., 2015; The Encode Project Consortium et al., 2012). Most recently, the 'Human cell atlas' project set out to generate a comprehensive map of all human cells using the latest single-cell genomic approaches (Regev et al., 2017). NGS provides a powerful tool to study various genome- and chromatin-linked aspects on a genome-wide scale.

1.1 Chromatin

At the end of the 19th century, chromatin and its basic building blocks DNA and histones were first described [reviewed in (Cremer and Cremer, 1988; Doenecke and Karlson, 1984; Olins and Olins, 2003)]. In 1869, Friedrich Miescher isolated a phosphorus-rich substance named *Nuklein* from pus leukocytes in the laboratory of Felix Hoppe-Seyler in Tübingen (Miescher, 1871). A few years later, Alfred Kossel, another former student of Felix Hoppe-Seyler, described a basic compound isolated from avian erythrocyte nuclei and termed it *Histon* (Kossel, 1884). In 1889, Richard Altman discovered that the non-protein portion of the *Nuklein* has acidic properties and named it therefore *Nukleinsäure* (= nucleic acid) (Altmann, 1889). At the same time, Walther Flemming described a stainable substance in the nucleus, which he named *Chromatin* (Flemming, 1882). In 1888, Wilhelm Waldeyer used first the word *Chromosom* to describe the chromatin structures in the

nucleus (Waldeyer, 1888). Only a century later in the 1970s, the basic structural and functional units of chromatin, the nucleosomes called 'ν bodies' at that time were independently described by Olins and Olins (Olins and Olins, 1974) and Woodcock et al. (Woodcock et al., 1976) visualized by electron microscopy as 'beads on a string' [reviewed in (Olins and Olins, 2003)]. Around the same time, Hewish and Burgoyne could show by limited nuclease digestion of chromatin leading to distinctly sized DNA fragments that chromatin is organized in a regular repeating structure (Hewish and Burgoyne, 1973). In 1974, Roger Kornberg suggested that these repeating units consist of a ~200 base pairs (bp) DNA fragment and two of each of the four main histones (Kornberg, 1974; Kornberg and Thomas, 1974). Finally, in 1975, Oudet et al. confirmed the proposed model of Roger Kornberg and the fundamental chromatin unit was by then termed 'nucleosome' (Oudet et al., 1975).

By now, numerous factors contributing to chromatin organization and chromatin plasticity have been described [reviewed in e.g. (Rivera and Ren, 2013; Yadav et al., 2018)]. At a glance, besides the canonical histones many different histone variants exist, their exchange is facilitated by histone chaperones and nucleosome remodelers, which can also reposition nucleosome along the DNA. Histones can be post-transcriptionally modified, e.g. methylation or acetylation of lysine residues and phosphorylation of serine residues, by histone modifying enzymes, so called 'writers'. These post-transcriptional modifications (PTMs) can be recognized by PTM binding proteins, the 'readers'. The DNA itself can also be bound by transcription factors and modified by DNA-modifying enzymes, most prominent the methylation of cytosines. In addition, the chromatin fibre is organized in higher-order structures to reach further compaction involving chromatin looping and forming of topologically associating domains (TADs). Thereby, all chromatin-related processes, such as gene expression, transcription, DNA replication and DNA repair, are regulated at various levels.

1.1.1 The nucleosome and higher-order chromatin structures

The canonical nucleosome core particle builds up from a central histone octamer core, containing two of each of the canonical histones H2A, H2B, H3 and H4, entwined with 147 bp DNA in 1.65 left-handed superhelical turns (Figure 1) (Davey et al., 2002; Luger et al., 1997) [reviewed in e.g. (Andrews and Luger, 2011; Cutter and Hayes, 2015; Khorasanizadeh, 2004; Luger et al., 2012; Pepenella et al., 2014)]. Each histone possesses a globular domain, formed by three α-helices arranged into the so-called 'histone fold', and unstructured N- and C-terminal regions, the 'histone tails'. By forming a 'handshake' structure, the histones dimerize giving rise to H2A-H2B and H3-H4 heterodimers, while the

later ones further dimerize to form a H3-H4 tetramer. To form the histone octamer, a H3-H4 tetramer associates with two H2A-H2B dimers on opposite sides of the tetramer. The globular domains of the histones within the octamer form a disc-like structure, where 147 bp of DNA are wrapped around to form the canonical nucleosome core particle. Basic histone residues contact the DNA giving rise to 14 weak protein-DNA interactions, however, summed up these interactions stably position the nucleosome. The histone tails emerge between the DNA to form an accessible hub for PTMs. *In vivo*, various histone chaperones are implicated in nucleosome formation.

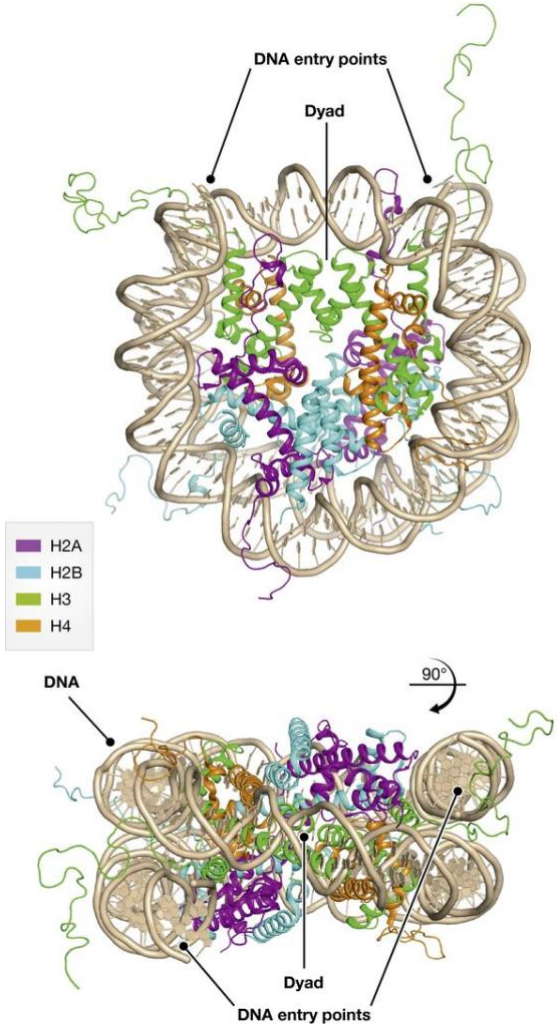


Figure 1. Structure of the canonical core nucleosome.

Representations of the X-ray crystal structure of the core nucleosome at 1.9 Å resolution (Davey et al., 2002). Cartoon of the nucleosome structure with H2A in purple, H2B in cyan, H3 in green, H4 in orange and DNA in wheat. Figure is reproduced from (Speranzini et al., 2016) with permission from John Wiley and Sons.

As *in vivo* DNA fragments have the size of chromosomes with tens to thousands of million base pairs, numerous nucleosomes are positioned on them forming the basis of chromatin [reviewed in e.g. (Bian and Belmont, 2012; Li and Reinberg, 2011; Luger et al., 2012;

Robinson and Rhodes, 2006; Tremethick, 2007; Woodcock et al., 2006)]. Therefore, neighboring nucleosome core particles are separated by short DNA stretches, the so-called 'linker DNA', to form nucleosome arrays, which have a diameter of 10 nm. The precise average linker length (mostly 20 – 40 bp) varies between organisms, but also between individual cell types and genomic locations. Binding of the linker histone H1 (or avian histone H5) at the nucleosome dyad and contacting both linker-DNAs lead to further compaction (Bednar et al., 2017). Histone H1 has an unstructured N-terminal region, a globular internal domain with a winged helix DNA-binding fold for binding the nucleosome dyad and a basic C-terminal domain. Further compaction of the primary nucleosomal array can be achieved by folding into a secondary structure, the 30 nm fiber. The extent to which the 30 nm fiber forms *in vivo* is still under debate. Further compaction is proposed to be alternatively achieved through interdigitated 10 nm fibers (Maeshima et al., 2016) [reviewed in e.g. (Ohno et al., 2018)]. However, based on various *in vitro* data two major models for the 30 nm fiber have been proposed (Figure 2). In the first model, the 'one-start' solenoid model, consecutive nucleosome form a super-helical path with 6 – 8 nucleosomes per turn involving bending of the linker DNA (Robinson et al., 2006). In the second model, the 'two-start' zigzag model, two nucleosomal rows fold into a two-start helix with straight linker DNA whereby very second next nucleosome interact with each other (Schalch et al., 2005). Recently, the special nucleosome positioning and orientation in chromosomes *in vivo* was studied by high-throughput chromosome conformation capture (Hi-C) resolving individual nucleosomes combined with simulated annealing-molecular dynamics in yeast (Ohno et al., 2019). The study showed that nucleosomes can fold into two secondary structures, either an α-tetrahedron or a β-rhombus fold, and the 3D nucleosome fold can be further altered through epigenetic marks at individual genomic loci. Generally, the secondary chromatin structure can be further compacted by folding into a tertiary structure. The highest compaction degree is achieved in metaphase chromosomes requiring condensin dependent loop formation (Gibcus et al., 2018). *In vivo*, higher-order chromatin structures are modulated and stabilized through binding of additional proteins to the chromatin fiber, including methyl-CpG-binding protein 2, heterochromatin protein 1 (HP1), high mobility group (HMG) proteins, Polycomb-group proteins. In the cell nucleus two general types of chromatin can be distinguished: euchromatin, which is more open and transcriptionally active, and heterochromatin, which is more compacted and transcriptionally inactive [reviewed in e.g. (Becker et al., 2016; Elgin and Reuter, 2013)].

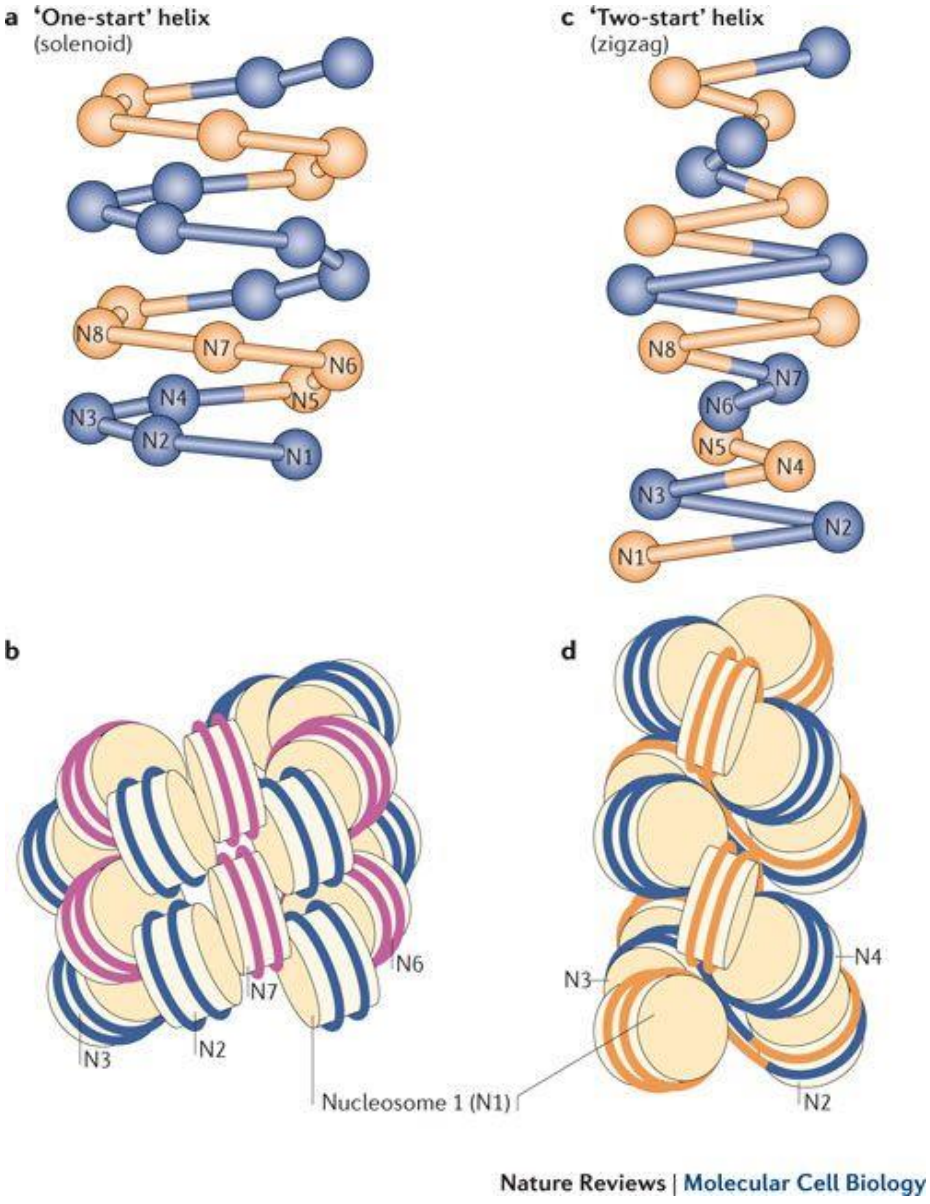


Figure 2. Models for 30 nm chromatin fibre.

Schematic representation of the two major structural models for chromatin secondary structure: left side, 'one-start' solenoid model, consecutive helical gyres are marked in blue and magenta/orange and right side, 'two-start' zigzag model, consecutive nucleosome pairs are marked in blue and orange. Figure is reproduced from (Luger et al., 2012) with permission of Elsevier.

Regulation apart from chromatin compaction is also achieved through positioning of individual nucleosomes along the DNA [reviewed in e.g. (Becker and Workman, 2013; Clapier and Cairns, 2009; Clapier et al., 2017; Luger et al., 2012; Radman-Livaja and Rando, 2010)]. Besides the intrinsic DNA sequence also adenosine triphosphate (ATP)-dependent nucleosome remodelers actively alter the nucleosome position and DNA-binding factors create local boundaries contributing to the precise positioning of nucleosomes. ATP-dependent nucleosome remodelers perform diverse functions to alter the chromatin structure such as nucleosome sliding, histone variant exchange and nucleosome eviction.

ATP-dependent nucleosome remodelers can be divided into four subfamilies differing in protein composition of the remodeling complexes. However, all remodeling complexes possess an ATPase 'motor' subunit that translocate DNA within the nucleosome.

1.1.2 Histone variants and modifications

The chromatin structure can be further modified through the exchange of canonical histones within the nucleosome with replacement histone variants [reviewed in e.g. (Gurard-Levin et al., 2014; Happel and Doenecke, 2009; Henikoff and Smith, 2015; Izzo and Schneider, 2016; Maze et al., 2014; Venkatesh and Workman, 2015)]. The genes of the canonical histones are present in multiple copies and clustered in the genome. They are only expressed during S-phase of the cell cycle and incorporated into chromatin in a replication-dependent manner. Their messenger ribonucleic acid (mRNA) lack introns and a polyA tail, instead they have a 3' stem loop structure crucial for regulating mRNA stability and translation. In contrast, histone variant genes are expressed throughout the whole cell cycle and incorporated in a replication independent manner involving specialized nucleosome remodelers and histone chaperones. Their mRNA can contain introns and they possess a 3' polyA tail. Except for histone H4, all canonical histones have replacement variants, which can differ at the protein level from the canonical histone by only few amino acids but my also contain additional non-histone domains. These differences in protein sequence can lead to changes in nucleosome stability, chromatin structure, PTM pattern and interaction with specific interaction partners.

The chromatin landscape can be altered at the level of histones not only by the exchange of histone variants but also by placing PTMs on histones (

Figure 3) [reviewed in e.g. (Andrews et al., 2016; Bannister and Kouzarides, 2011; Gelato and Fischle, 2008; Jenuwein and Allis, 2001; Kouzarides, 2007; Musselman et al., 2012; Rothbart and Strahl, 2014; Taverna et al., 2007)]. PTMs are often small chemical groups added to amino acid residues, e.g. covalent linkage of methyl-groups to lysine and arginine residues, acetyl-groups to lysine residues and phosphate-groups to serine and threonine residues. Probably methylation is the most wide—spread histone PTM, whereby lysine residues can be mono-, di- and trimethylated, while the arginine residues can be mono- and dimethylated. Larger groups like ubiquitin and Small Ubiquitin-related Modifier (SUMO) can also be post-transcriptionally added to lysine residues. In recent years, many novel histone modifications were identified, e.g. adenosine diphosphate (ADP) ribosylation, crotonylation or addition of β-N-acetylglucosamine.

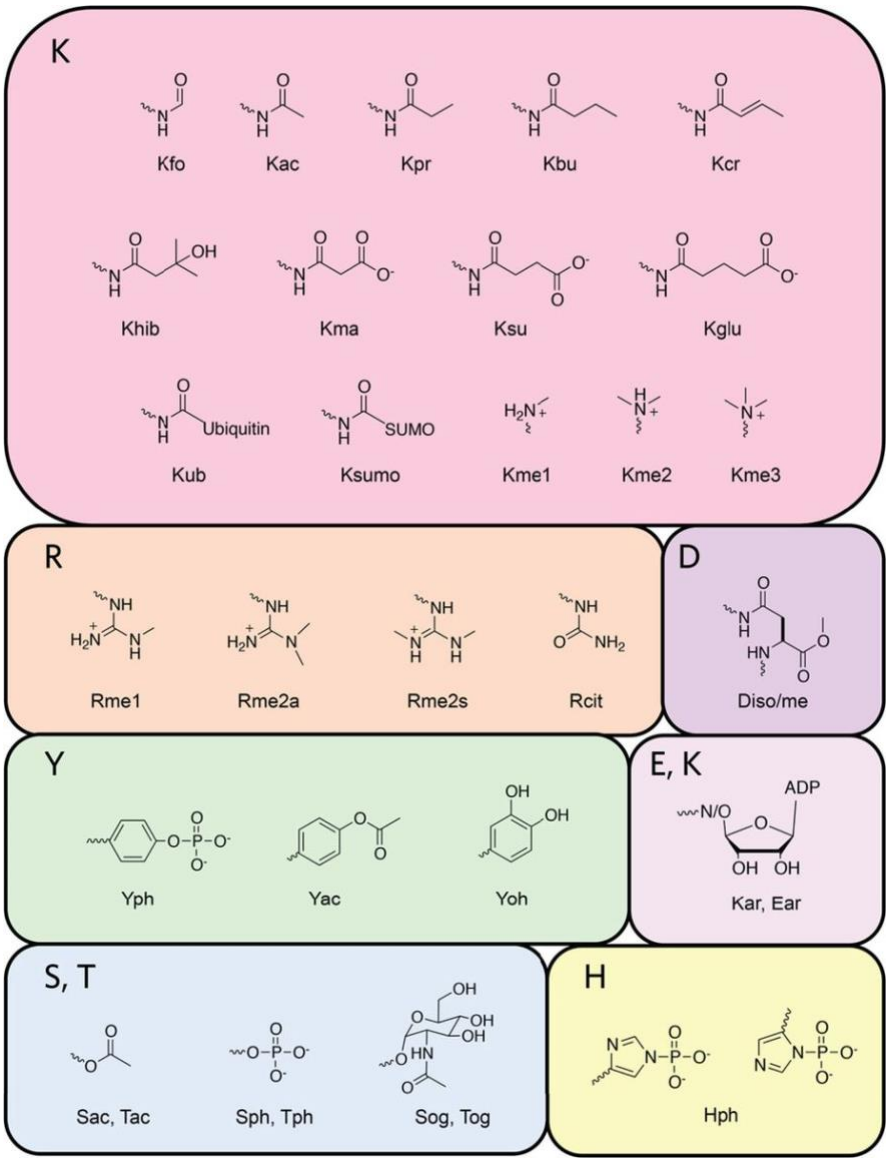


Figure 3. Identified post-transcriptional modifications in histones.

Modifications are grouped according to the modified amino acid, using the abbreviations: ar - ADP ribosylation; cit - citrullination; fo - formylation; glu - glutarylation; ma - malonylation; oh - hydroxylation; og - O-GlcNAcylation; su - succinylation; ub - ubiquitination. Figure is reproduced from (Andrews et al., 2016) with permission from Springer Nature.

Histone PTMs are often placed at specific amino acid residues in the histone tails, still some residues in the globular domain or at the C-terminus can be modified (Figure 4). PTMs are placed by site-specific histone modifying enzymes, the ‘writers’, and the presence of specific PTMs is a dynamic process as most can also be specifically removed by other histone modifying enzymes, the ‘erasers’. Histone modifications can shape the chromatin landscape in different ways. The addition of PTMs can cause direct biophysical changes to the chromatin fiber, as charges can be added to or removed from the histones. While methylation of lysine residues does not alter the positive net charge, acetylation neutralizes the charge, which weakens the interaction of the histone tail with the negatively charged

DNA and causes a more open chromatin state. Additionally, many chromatin binding proteins, the 'readers', can specifically recognize individual PTMs as they have modification-specific binding domains. In many cases, specific PTMs accumulate in distinct genome region where they are associated with specific chromatin processes, e.g. constitutive heterochromatin is marked by trimethylation of histone H3 at lysine 9 (H3K9me3) and active promoters with trimethylation of histone H3 at lysine 4 (H3K4me3). Therefore, the chromatin landscape can be separated into discrete chromatin states due to the prevalent histone marks and associated chromatin factors (Ernst and Kellis, 2010, 2012; Filion et al., 2010; Kharchenko et al., 2011).

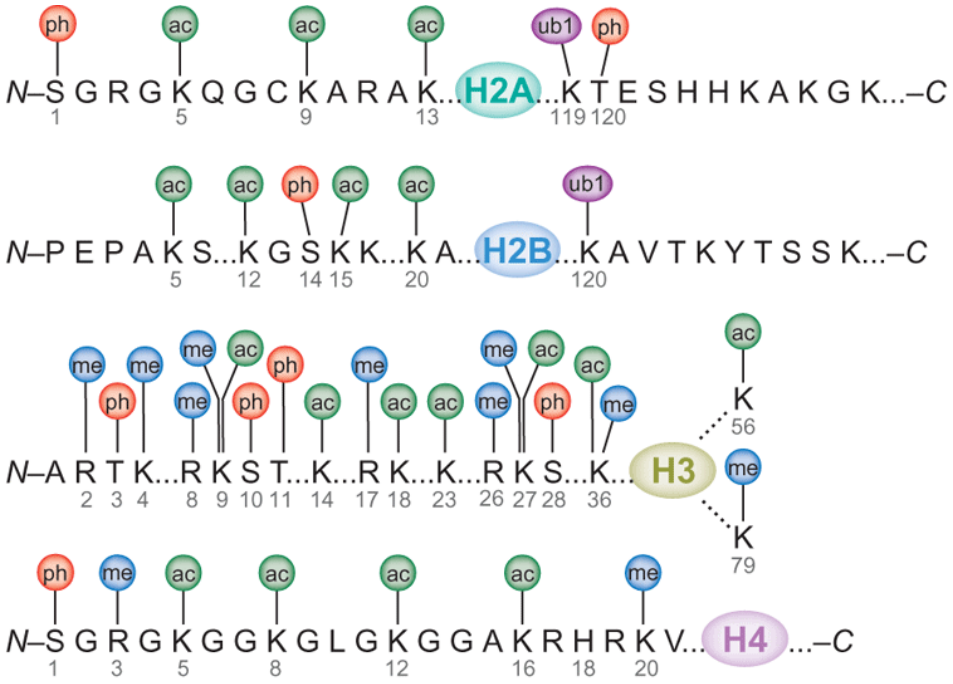


Figure 4. Post-transcriptional modifications of histone tails.

Modifications of specific histone tail residues. Positions are indicated by numbers. Colored ovals represent globular histone domains. Figure is reproduced from (Bhaumik et al., 2007) with permission from Springer Nature.

1.2 Chromatin binding factors

Chromatin binding proteins often recruit and regulate other chromatin modifying, chromatin remodeling, or transcription regulating complexes to target genome loci. Chromatin binders can either directly bind to DNA in a chromatin environment or interact with histone PTMs. Until now, many distinct protein domain families were discovered to recognize and bind specifically to certain histone PTMs [reviewed in e.g. (Musselman et al., 2012; Patel and Wang, 2013; Taverna et al., 2007; Weaver et al., 2018; Yun et al., 2011)]. Probably the most extensively studied group are the binders of methylated lysine residues, they include ATRX-DNMT3-DNMT3L (ADD) domain, ankyrin domain, bromo-adjacent homology domain, chromo-barrel domain, chromodomain, double chromodomain, malignant brain tumor (MBT) domain, plant homeodomain (PHD), PWWP (Pro-Trp-Trp-Pro) domain, tandem Tudor domain, Tudor domain, WD40 domain and zinc finger CW domain. The methyl-lysine binding domains contain an aromatic cage formed by two to four aromatic residues, which specifically recognize the methyl mark. The Tudor domain and WD40 domain are also capable in binding to methylated arginine residues. Substantially fewer protein domains are known to recognize acetylated lysine residues. So far the bromodomain, double PHD finger domain and double pleckstrin homology domain were described to bind acetyl-lysine. Similarly, only a few binding domains were identified to specifically interact with phosphorylated serine and threonine residues of histones, including tandem BRCT domain, BIR domain and the 14-3-3 protein family.

Some of the histone reader domains, including bromodomains, PHD fingers, PWWP domains, chromodomains and Tudor domains, are also described to have additional nucleic acid binding activity [reviewed in (Weaver et al., 2018)]. Binding of reader domains to DNA and/or RNA is largely sequence-independent. Accordingly, nucleic acid binding may be independent of histone binding, or binding to DNA/RNA increases histone binding and the other way around. The additional nucleic acid binding of reader domains can result in increased affinity for chromatin and retention at target sites. Furthermore, multivalent histone and nucleic acid binding can increase the affinity for nucleosomes or enable proper positioning on the nucleosome to increase specificity for certain histone residues. Another mechanism could be that binding of the reader domain to DNA releases DNA-interacting histone tails, to make the histone tail accessible for reader/writer domains.

Besides histone reader domains, numerous domains with sequence specific DNA binding activity are described (Fulton et al., 2009; Vaquerizas et al., 2009) [reviewed in e.g. (Lambert et al., 2018; Luscombe et al., 2000)]. Proteins with specific DNA binding activity are often transcription factors (TF) as they participate in transcription regulation. The major DNA binding domain families include C2H2 zinc finger (ZnF), Homeodomain, basic helix-loop-helix

domain, basic leucine zipper domain, and nuclear hormone receptor domain. DNA binding domains have two DNA recognition modes to readout the DNA sequence. One form of DNA interaction is referred to as ‘direct readout’, whereby amino acid side chains of the DNA binding domain directly contact the base pairs via direct or water mediated hydrogen bonds and hydrophobic interactions. The other form is known as ‘indirect readout’, whereby the binding domain senses the DNA shape, which in turn is a sequence-specific structural feature of the DNA. Sequence-specific DNA binding factors usually recognize a set of hundreds to thousands similar DNA sequence over a range of affinities within the genome [reviewed in e.g. (Inukai et al., 2017; Pan et al., 2010; Rohs et al., 2010; Siggers and Gordan, 2014; Slattery et al., 2014)]. The consensus binding site motif of bound DNA sequences is commonly modelled by position weight matrix (PWM). However, only a subset of DNA sequences which resemble to the binding site motif are bound by the DNA binding factor *in vivo*. Besides the DNA sequence, additional features contribute to binding site selection *in vivo*, including DNA shape, flanking DNA, variable spacing of bipartite binding motifs, DNA methylation/modification, multimeric binding, cofactor interaction, interaction/cooperation with other DNA binding factors, DNA accessibility, nucleosome occupancy and competition with nucleosomes.

1.2.1 Cooperative DNA binding

DNA binding factors are non-uniformly distributed throughout the genome, they rather form clusters or hotspots at distinct loci. Since binding motifs are only slightly enriched within the binding clusters, cooperativity between DNA binding factors through various mechanisms seems to contribute to the enrichment of factors at binding clusters [reviewed in e.g. (Morgunova and Taipale, 2017; Siggers and Gordan, 2014)]. Cooperation can be achieved through direct protein-protein interaction of two DNA binding factors in the absence of DNA. There, the protein complex poses a higher affinity for DNA as the individual factors. Similarly, if the interaction of two factors is too weak to form a stable soluble complex, DNA binding can favor this interaction by bringing both factors in close proximity or introducing conformational changes required for stable interaction, resulting in increased DNA binding affinity. Furthermore, solely DNA-mediated cooperativity between two factors that lack direct protein-protein interaction can still strengthen their DNA binding. Changes in DNA shape or dynamics induced by binding of one factor can result in increased affinity for DNA of another factor. Additionally, an indirect cooperativity mechanism is described that functions through competition for DNA binding between binding factors and nucleosomes allowing binding of several factors next to each other at the created nucleosome free region (NFR).

1.2.2 DNA binding through CXC domains

The Tesmin/TSO1-like CXC domain is a rather rare protein domain, which can be found in the Animalia and Plantae kingdom and is absent in the Fungi kingdom (Marin, 2003). The Tesmin/TSO1-like CXC domains are particularly conserved within the C-terminal part with the consensus formula [C-$(X)_4$-C-(X)-C-$(X)_6$-C-X_4-5-C-$(X)_2$-C], while the N-terminal part is more variable. Three subtypes of Tesmin/TSO1-like CXC domains can be distinguished (Marin, 2003; Mitchell et al., 2019): (a) “CRC domain” found in the human Tesmin protein and in the *Arabidopsis thaliana* TSO1 protein; Many proteins with “CRC domains” have two domains in tandem separated by a RNPXAFXPK linker (Andersen et al., 2007; Song et al., 2000; Sutou et al., 2003); To this class also the LIN-54 homolog family members belong, in *Drosophila melanogaster* this is mip120 (Beall et al., 2002; Marin, 2003; Schmit et al., 2007); (b) “CXC domain” found in the homologs of the polycomb group protein Enhancer of zeste [E(z)] (Abel et al., 1996; Chen et al., 1996; Goodrich et al., 1997; Jones and Gelbart, 1993); (c) “E3 ubiquitin-protein ligase Msl2, CXC domain” found in the homologs of the *D. melanogaster* male-specific-lethal (MSL) 2 protein (Figure 5) (Bashaw and Baker, 1995; Smith et al., 2005; Zhou et al., 1995); However, the “E3 ubiquitin-protein ligase Msl2, CXC domain” is specific to animals (Marin, 2003). The CXC domain seems to be a conserved DNA binding domain as some protein members are shown to bind DNA, including *D. melanogaster* MSL2 (Fauth et al., 2010; Zheng et al., 2014), human LIN54 (Marceau et al., 2016; Schmit et al., 2009) and soybean Cysteine-rich polycomb-like protein 1 belonging to the “CRC domain” family too (Cvitanich et al., 2000). Additionally, the DNA binding mode of *D. melanogaster* MSL2 and human LIN54 were studied using X-ray crystallography and nuclear magnetic resonance (NMR) (Figure 5) (Marceau et al., 2016; Zheng et al., 2014; Zheng et al., 2012). *D. melanogaster* MSL2 harbors a single CXC fold, which is formed by three zinc ions coordinated by nine cysteine residues. The Zn_3Cys_9 cluster is wrapped by flexible loops and a short α-helix in a two turn right-handed helix. When bound to DNA, two CXC domains bind next to each other on opposite strands making protein-protein contacts. Each CXC domain contacts the DNA phosphate backbone by hydrogen bonding and inserts two arginine residues, one into the minor and one into major groove, forming base-specific hydrogen bonds. The human LIN54 harbors two tandem “CRC domains” with independent CXC folds, which are connected by a flexible linker. To form the CXC fold, nine cysteine residues coordinate three zinc ions, which are connected by flexible loops and short α-helix making 1.5 right-handed helical turns. In addition, each fold contains a C-terminal α-helix, which is absent in the MSL2 CXC fold. The tandem “CRC domains” bind to adjacent, opposite sites at the minor grove to the target DNA and the domains make no protein-protein contacts. Each domain makes extensive hydrogen bonding to the DNA phosphate backbone

and base-specific contacts by inserting two tyrosine residues into the minor groove. The tyrosine residues form hydrogen bonds to two adjacent bases from opposite strands.

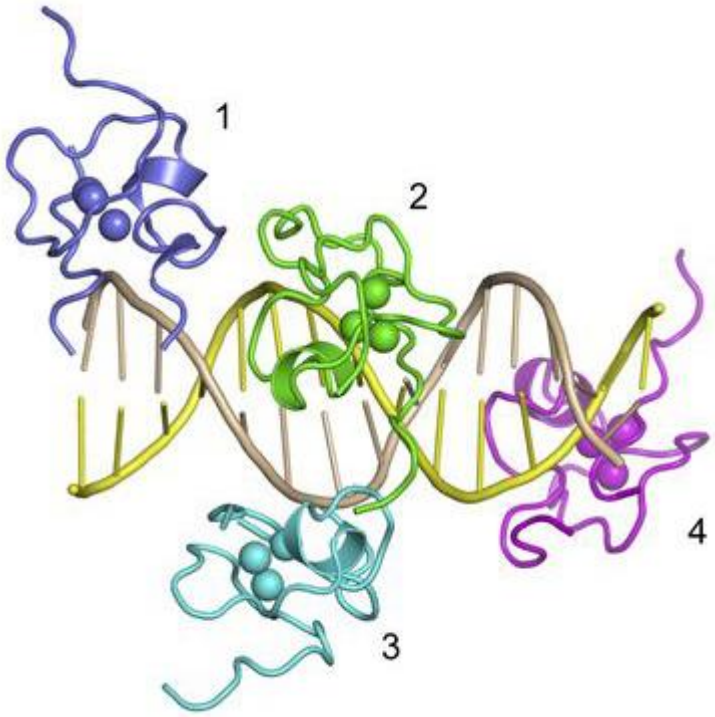


Figure 5. CXC domain binding to DNA.

Representation of NMR assembly of MSL2 CXC domain bound to DNA. Cartoon representations of four CXC domains of MSL2 binding to DNA. Figure is reproduced from (Zheng et al., 2014) with permission of Cold Spring Harbor Laboratory Press.

1.2.3 DNA binding through C2H2 zinc finger domains

Zinc finger domains are found throughout all kingdoms of life and are one of the most abundant protein domains in eukaryotes [reviewed in e.g. (Brayer and Segal, 2008; Brown, 2005; Iuchi, 2001; Razin et al., 2012; Wolfe et al., 2000)]. Various types of ZnF domains exist, which have in common that they use cysteine and/or histidine residues to complex a single zinc ion for stabilizing their fold. The founding member of the ZnF superfamily is the "classical" C2H2 ZnF domain. The C2H2 ZnF domain consists of ~30 amino acid residues with the consensus sequence (F/Y)-X-C-X_{2-5}-C-X_3-(F/Y)-X_5-ψ-X_2-H-X_{3-5}-H, ψ representing hydrophobic residues. The first structure of C2H2 ZnF domain was solved in 1989, showing that the single zinc ion is tetrahedrally coordinated by two cysteine and two histidine residues (Lee et al., 1989; Wolfe et al., 2000). Further, the C2H2 ZnF domain folds into two N-terminal, antiparallel β-sheets with the connecting loop harboring the two cysteine residues and one C-terminal α-helix with the two histidine residues at the C-terminal part of the α-helix (Figure 6). C2H2 ZnF domains are involved in DNA binding, RNA binding and protein-protein

interaction. Many proteins contain patches of multiple C2H2 ZnF domains, some proteins even > 30. Based on the number and distribution of C2H2 ZnF domains within a protein, C2H2 ZnF containing proteins can be categorized into three groups (Iuchi, 2001; Razin et al., 2012): (a) “triple-C2H2 ZnF proteins”, this group includes the Krüppel-like family and SP1-like family of TFs; (b) “multiple-adjacent-C2H2 ZnF proteins” is the largest group and the proteins contain clusters of > 4 closely spaced C2H2 ZnF domains; one well-studied member of this group is the CTCF protein containing a cluster of 11 C2H2 ZnF domains (c) “separated-paired-C2H2 ZnF proteins”, this group comprises the lowest number of TFs, including tramtrack, PRDII-BF1 and basonuclin; This group of ZnF proteins contain one or more pairs of C2H2 ZnF domains, where individual pairs are separated by many intervening amino acid residues. Alternatively, C2H2 ZnF proteins are classified according to additional domains present in the protein, such as the KRAB domain, SCAN domain and BTB/POZ domain.

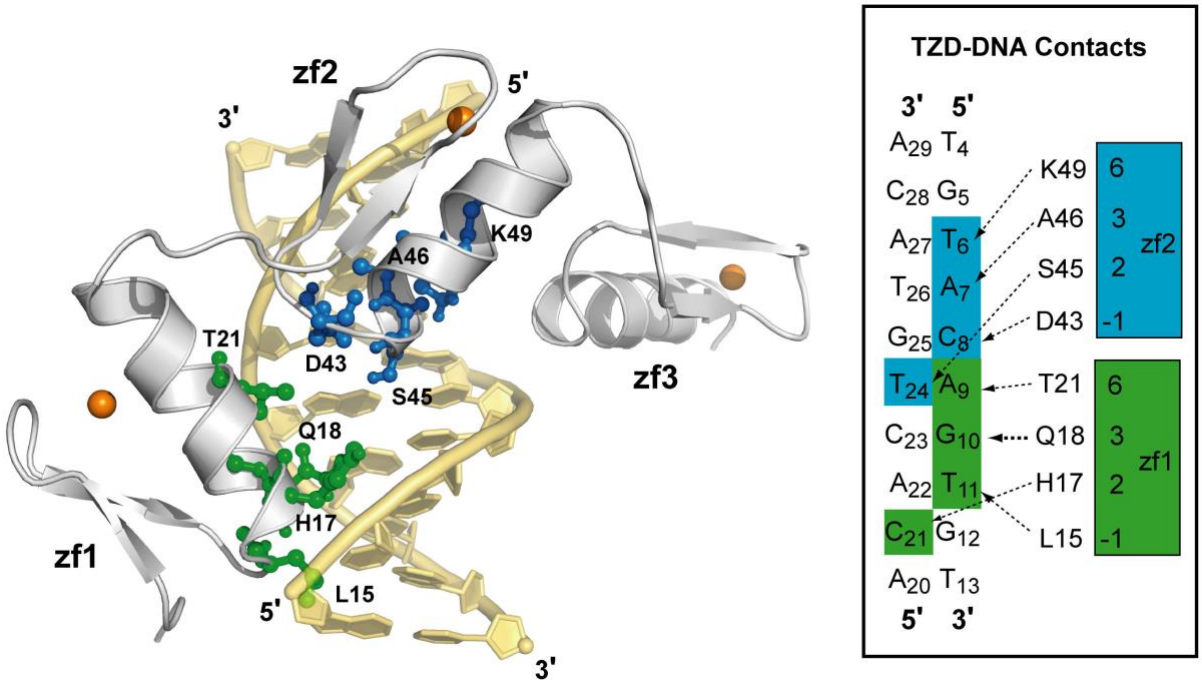


Figure 6. C2H2 zinc finger domain binding to DNA.

Representation of NMR assembly of TZD ZnF domain bound to DNA. Left side, cartoon representation of three C2H2 ZnF domains of TZD binding to DNA, with side chains at position -1, 2, 3 and 6 as sticks and balls. Right side, schematic representation of direct DNA base contact of ZnF domains. Figure is reproduced from (Chou et al., 2017) with permission of PLOS.

C2H2 ZnF domains are mostly described in DNA binding, as many TFs contain multiple C2H2 ZnF domains. Binding of C2H2 ZnF domains to DNA is studied extensively using structural approaches (Elrod-Erickson et al., 1996; Pavletich and Pabo, 1991; Wolfe et al., 2001; Wolfe et al., 2000; Wuttke et al., 1997). For DNA binding, the C2H2 ZnF domain makes numerous contacts to the phosphate back bone and direct base contacts by inserting the α-helix into the major groove. For direct base contact, the amino acid residues at position

-1, 3, and 6 form hydrogen bonds with one DNA strand and the residue at position 2 contacts the complementary strand (Figure 6). C2H2 ZnF domains occur often in clusters and neighboring C2H2 ZnF domains wrap around DNA when bound. Each C2H2 ZnF domain contacts four consecutive bases and adjacent C2H2 ZnF domains are shifted by three base pairs, resulting in the overlap by one base pair. Neighboring C2H2 ZnF domains are most commonly connected by a five amino acid-linker with the consensus sequence TGEKP. The linker contributes to correct spacing of adjacent C2H2 ZnF domains and responsible for capping the C-terminus of the preceding α-helix in a DNA binding-dependent manner (Laity et al., 2000). Capping of the α-helix C-terminus provides rigidity and locks the α-helix into the correct place when bound to DNA, contributing to high affinity binding. While scanning the DNA strand for appropriate binding sites, the linker does not fold onto the α-helix C-terminus providing flexibility between adjacent domains. However, structural studies of GLI and TFIIA as well as biochemical and genomic studies of CTCF showed that proteins with multiple adjacent C2H2 ZnF domains can just use a subset of their C2H2 ZnF domains to bind DNA, allowing certain flexibility and deviation in DNA binding site recognition (Filippova et al., 1996; Filippova et al., 2002; Nakahashi et al., 2013; Nolte et al., 1998; Pavletich and Pabo, 1993; Renda et al., 2007).

1.2.4 Nucleosome binding through PWWP domains

The PWWP domain belongs to the Royal superfamily, which also includes the Tudor domain, chromodomain, and MBT domain [reviewed in e.g. (Qin and Min, 2014; Rona et al., 2016; Weaver et al., 2018)]. Members of the Royal superfamily can bind specifically to methylated lysine and/or arginine residues in histone tails. They share a conserved aromatic cage structure, which is responsible for methyl-lysine/arginine binding, present in a β-barrel fold formed by three to five β-sheets. The PWWP domain is named after the characteristic Pro-Trp-Trp-Pro sequence motif present in the N-terminal part of the domain. However, especially the first and second position of the motif is variable. The PWWP domain was first described in WHSC1 and is conserved within all eukaryotes (Stec et al., 1998). Over the last years, much structural information on PWWP domains and their binding to nucleosomes was obtained (Figure 7) (Eidahl et al., 2013; Laguri et al., 2008; Qiu et al., 2002; Sue et al., 2004; van Nuland et al., 2013; Vezzoli et al., 2010; Wen et al., 2014; Wu et al., 2011a).

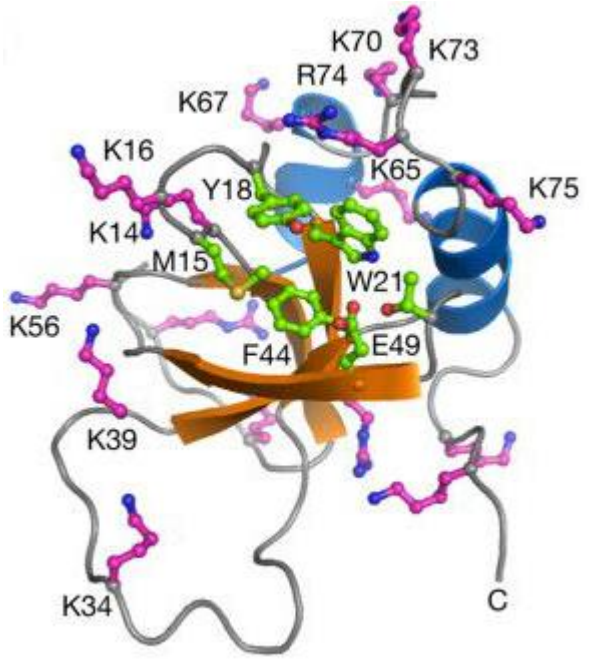


Figure 7. Protein structure of PSIP1 PWWP domain.

Representation of NMR assembly of PSIP1 PWWP domain. Cartoon representation of PSIP1 PWWP domain with aromatic side chains in green balls and sticks and basic side chains in purple. Figure is reproduced from (van Nuland et al., 2013) with permission of BioMed Central.

The PWWP domain is formed by a β-barrel fold of five antiparallel β-sheets and a helix bundle of one to six α-helices. Within the β-barrel, a loop region of variable length and secondary structure between β2 and β3 and often a 3_{10} helix between β4 and β5 are inserted. Furthermore, three aromatic residues located in the loop between β1 and β2, in the PWWP motif and in the β3-strand form an aromatic cage accommodating the methyl-lysine side chain. Most PWWP domains are suggested to bind trimethylated histone H3 at lysine 36 (H3K36me3), also binding to trimethylated histone H3 at lysine 79 (H3K79me3) and mono/trimethylated histone H4 at lysine 20 (H4K20me1/3) were described (Dhayalan et al., 2010; Eidahl et al., 2013; Li et al., 2013a; van Nuland et al., 2013; Vermeulen et al., 2010; Wen et al., 2014; Wu et al., 2011a). A second interaction surface is formed by solvent exposed basic residues mostly located at β1-β2 loop, PWWP motif and β2-strand next to the aromatic cage. This positively charged interaction surface mediates sequence-unspecific DNA binding through interactions with the phosphate backbone (Figure 8). As shown for PWWP domain binding to H3K36me3 nucleosome, the positively charged DNA binding surface and methyl-lysine binding aromatic cage synergize for binding with high affinity (Eidahl et al., 2013; van Nuland et al., 2013). The PWWP domains contacts the phosphate backbone of both DNA gyres next to the H3 tail exit site through its basic surface and the aromatic cage engages with the K36me3 residue.

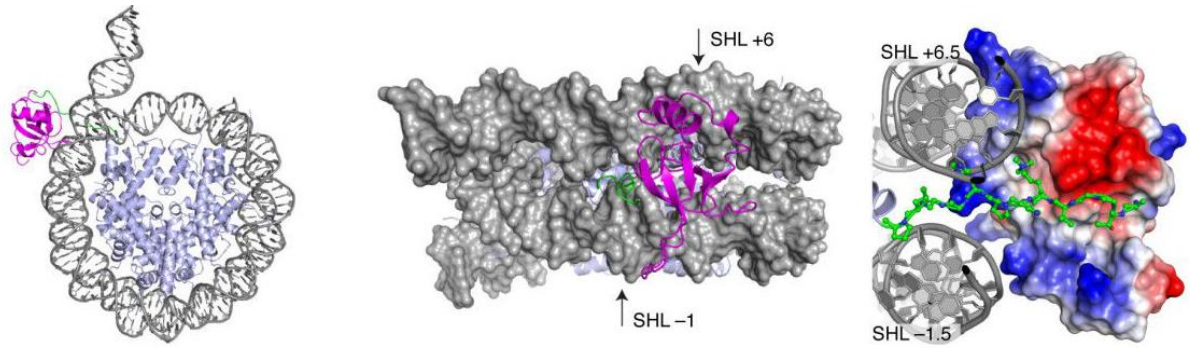


Figure 8. PWWP domain binding to H3K36me3 nucleosome.
Structural model of PSIP1 PWWP domain bound to H3K36me3 nucleosome. Left side and middle, cartoon representation of lowest energy structure of PSIP1 PWWP domain docked to H3K36me3 nucleosome. Right side, detailed view on PSIP1 PWWP domain bound to H3K36me3 nucleosome with residues 31 to 38 of the H3 N-terminal tail as balls and sticks in green, DNA a cartoon representation and electrostatic potential of molecular surface of PWWP domain. Figure is reproduced from (van Nuland et al., 2013) with permission of BioMed Central.

1.3 Sex chromosome dosage compensation

Most metazoans propagate by various means of sexual reproduction. Here, sex determination is genetically encoded by sex chromosomes, which often belong to the XY system. In the XY system, females have two X chromosomes in addition to a diploid set of autosomes, while males have a single X chromosome and an often degenerated, gene-poor Y chromosome [reviewed in e.g. (Chandler, 2017; Disteche, 2012, 2016; Gu and Walters, 2017; Lucchesi, 2018)]. The concept of sex chromosome evolution was already introduced one hundred years ago by Hermann J. Muller, who proposed that reduced recombination between the X and Y chromosome leads to degeneration of the Y chromosome and later he also presented the concept of dosage compensation (Muller, 1914). In 1967, Susumu Ohno extended that idea and proposed that the gene-rich X and degenerated Y chromosome evolved from a normal pair of autosomes that acquired a sex-determining locus (Ohno, 1967). The different number of X chromosomes creates an imbalance of X-linked genes between males and females and between the single male X chromosome and the diploid autosomes. This imbalance is adjusted by a process called dosage compensation, first proposed by Hermann J. Muller (Muller, 1932, 1950). During evolution, various mechanisms for sex chromosome dosage compensation evolved individually, the best studied systems are eutherian mammals (humans/mice), *Caenorhabditis elegans* and *Drosophila melanogaster* (Figure 9).

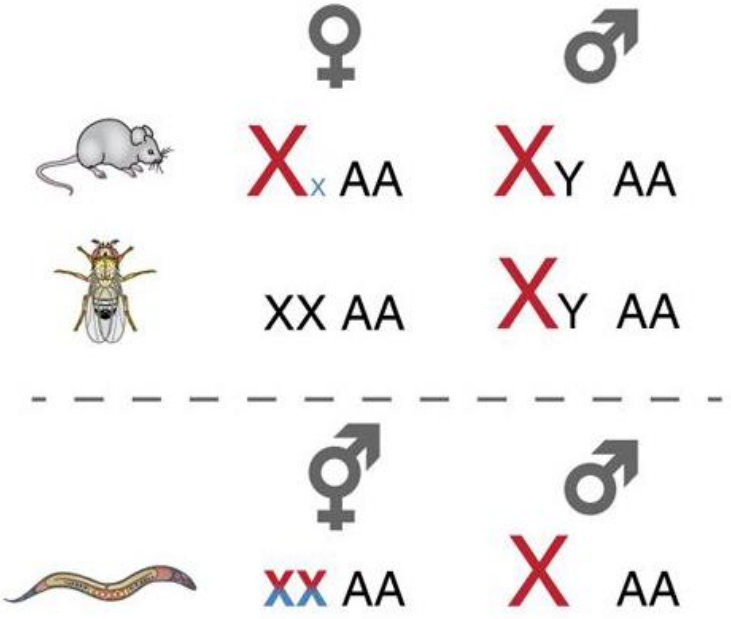


Figure 9. Dosage compensation of sex chromosomes.
Schematic representation of sex chromosome dosage compensation in *Mus musculus*, *Drosophila melanogaster* and *Caenorhabditis elegans*. Figure is reproduced from (Ferrari et al., 2014) with permission of Springer Nature.

In mammalian dosage compensation, one of the female X chromosomes is inactivated by the formation of the heterochromatic Barr body [reviewed in e.g. (Bonora and Disteche, 2017; Brockdorff, 2017; Brockdorff and Turner, 2015; Migeon, 2017; Pinheiro and Heard, 2017; Sahakyan et al., 2017)]. Inactivation of the X chromosome is triggered through the expression of the long non-coding RNA (lncRNA) Xist, which coats the whole chromosome. Xist expression leads to the recruitment of polycomb complexes which repress gene expression by placing heterochromatic histone marks including trimethylated histone H3 at lysine 27 (H3K27me3) and ubiquitylation of histone H2A at lysine 119 (H2AK119ub). Furthermore, X chromosome silencing includes incorporation of the histone variant macroH2A. Stochastic inactivation of one of the two female X chromosomes leaves females with only one active X chromosome like males. This creates an imbalance between X-linked genes and autosomal genes in both sexes. Therefore, the gene expression from the single active X chromosome is up-regulated in both sexes through a mechanism that involves acetylation of histone H4 at lysine 16 (H4K16ac) by males absent on the first (MOF) at promoters and enrichment of paused RNA polymerase II and increased RNA stability (Deng et al., 2013).

In *C. elegans*, males have a single X chromosome and lack a Y chromosome, while hermaphrodites have two X chromosomes. In hermaphrodites, the two X chromosomes are silenced to 50% during dosage compensation. Dosage compensation in worms is facilitated by the dosage compensation complex (DCC), which is a specialized condensin complex

[reviewed in e.g. (Albritton and Ercan, 2018; Strome et al., 2014)]. Binding of the DCC to the two X chromosomes in hermaphrodites leads to reduced expression of X-linked genes through chromosome compaction, reduced H4K16ac level and increased monomethylation of histone H4 at lysine 20, which represses RNA polymerase II binding to promoters. As in mammals, this silencing by half of both X chromosomes results in an imbalance between X-linked and autosomal genes, which is also present in males. Therefore, the expression of X-linked genes is up-regulated in both sexes.

1.3.1 Dosage compensation in *Drosophila melanogaster*

D. melanogaster dosage compensation differs from mammals and *C. elegans* as it lacks inactivation of the female X chromosomes and only the single male X chromosome is upregulated to match the two active female X chromosomes to prevent male lethality (Mukherjee and Beermann, 1965) [reviewed in e.g. (Birchler, 2016; Conrad and Akhtar, 2012; Ferrari et al., 2014; Keller and Akhtar, 2015; Kuroda et al., 2016; Lucchesi and Kuroda, 2015; Maenner et al., 2012; McElroy et al., 2014; Samata and Akhtar, 2018; Straub and Becker, 2007, 2011)]. That male-specific upregulation of X-linked gene expression in the two-fold range is facilitated by the so-called dosage compensation complex (DCC). The DCC is a ribonucleoprotein complex consisting of five male-specific-lethal (MSL) protein subunits and a long, non-coding (lnc) RNA component (Figure 10). The first DCC members, including MSL1, MSL2, MSL3 and maleless (MLE), were discovered by genetic screens or isolated in nature around the early 1980s (Belote and Lucchesi, 1980a, b; Fukunaga et al., 1975; Uchida et al., 1981). Later in the early 1990s, bulk acetylation of H4K16 was described to be restricted to the male X chromosome and depending on the presence of MSL proteins (Bone et al., 1994; Turner et al., 1992). The remaining DCC components were only discovered in the late 1990$_{th}$, including the fifth MSL protein males absent on the first (MOF), the histone acetyltransferase specific for H4K16ac, and the two lncRNA subunits, RNA-on-the-X (roX) 1 and 2 (Amrein and Axel, 1997; Franke and Baker, 1999; Hilfiker et al., 1997; Meller et al., 1997). As indicated by their naming, both RNAs roX1 and roX2 are encoded on the X chromosome and co-localize with the MSL proteins on the male X chromosome (Kelley et al., 1999; Meller et al., 2000). Co-immunoprecipitation (IP) experiments indicated that the MSL proteins form a stable complex together with the roX RNAs (Alekseyenko et al., 2014; Copps et al., 1998; Kelley et al., 1995; Meller et al., 2000; Smith et al., 2000; Wang et al., 2013). Furthermore, the DCC assembles exclusively in males, as translation of MSL2 protein is restricted to males (Gebauer et al., 1999; Gebauer et al., 2003; Grskovic et al., 2003; Hennig et al., 2014; Kelley et al., 1995; Kelley et al., 1997; McDowell et al., 1996; Szostak et al., 2018). The Sex lethal (SXL) protein, which regulates sex determination through sex-specific

splicing, is only functional in females due to female-specific expression and splicing during early embryogenesis. SXL prevents translation of msl2 mRNA by a splicing-independent mechanism through inhibiting ribosome assembly on msl2 mRNA. In males, a non-functional splicing isoform of SXL is expressed allowing msl2 mRNA translation and accumulation of MSL2 protein leading to DCC formation and finally dosage compensation. The mechanism of dosage compensation in the model organism *D. melanogaster* is widely conserved throughout various *Drosophilid* species (Alekseyenko et al., 2013; Quinn et al., 2016; Sural et al., 2008).

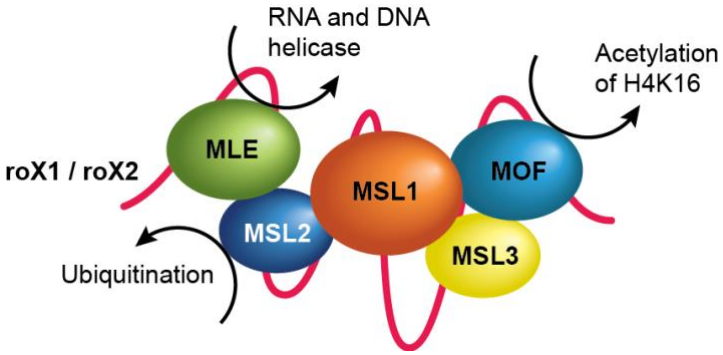


Figure 10. Dosage compensation complex.
Schematic representation of the DCC in *D. melanogaster*. Figure is adapted from (Straub and Becker, 2007) with permission of Springer Nature.

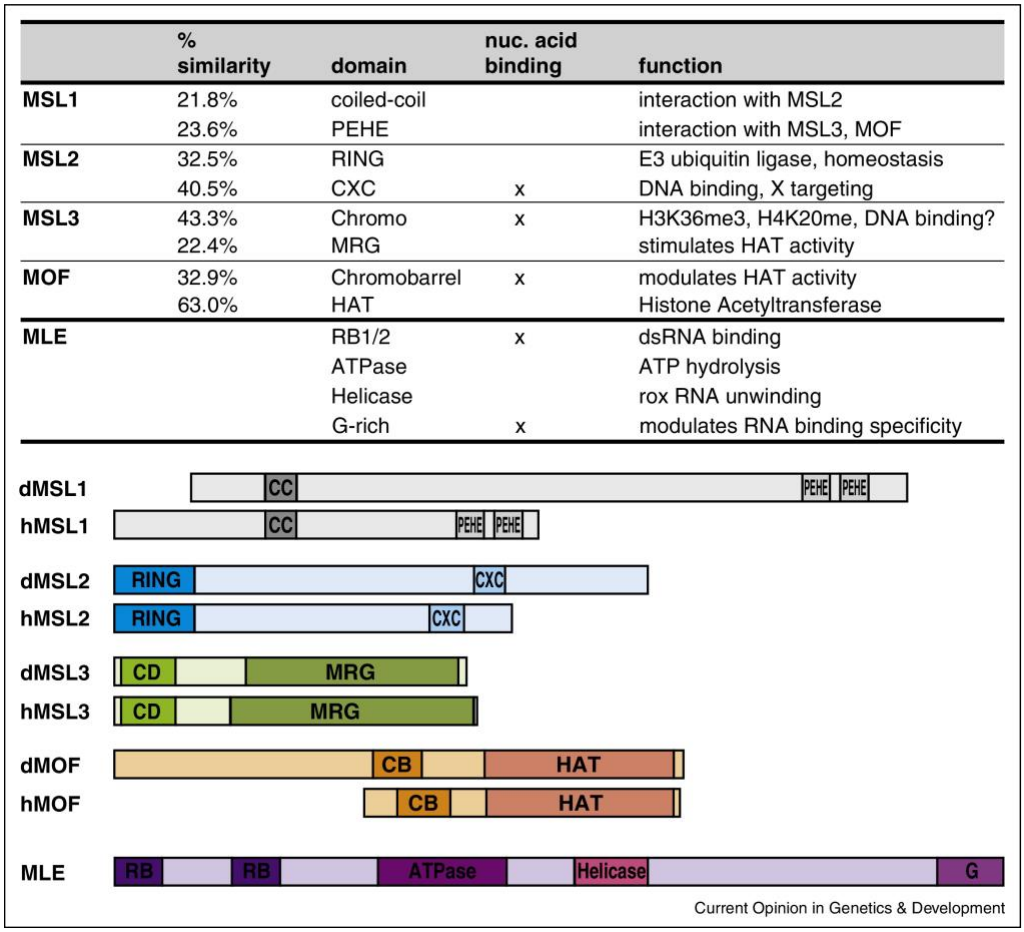
1.3.2 The Dosage Compensation Complex

The DCC forms most likely a dimeric complex through homodimerization of MSL1 via its N-terminal coiled-coil domain (Figure 11) (Hallacli et al., 2012). MSL1 seems to be a rather unstructured protein and in addition to the N-terminal coiled-coil domain it has only a C-terminal PEHE domain. MSL1 serves as a scaffold for the other DCC members, as it interacts via its coiled-coil domain with the really interesting new gene (RING) finger domain of MSL2, via the N-terminal half of the PEHE domain with the histone acetyltransferase (HAT) domain of MOF and via the C-terminal half of the PEHE domain with the MRG domain of MSL3 (Hallacli et al., 2012; Kadlec et al., 2011; Li et al., 2005; Morales et al., 2005; Morales et al., 2004; Scott et al., 2000; Wu et al., 2011b). MSL2 is the male-specific subunit of the complex and it harbors two domains, the N-terminal RING finger domain and a C-terminal CXC domain followed by a proline-rich region including a patch of basic residues. MSL2's RING finger domain possesses an E3 ubiquitin ligase activity besides being required for interaction with MSL1 (Kruse and Gu, 2009; Schunter et al., 2017; Villa et al., 2012; Wu et al., 2011b). MSL2 can ubiquitylate itself and other members of the DCC, including MSL1, MSL3 and MOF. It is thought that MSL2 contributes to DCC homeostasis by ubiquitylating surplus DCC and marking it for proteasome-dependent degradation as DCC levels need to

be tightly regulated (Demakova et al., 2003; Johansson et al., 2011; Kelley et al., 1995; Kelley et al., 1997; Villa et al., 2012). MSL2's CXC domain is a DNA binding domain required for DCC targeting to the X chromosome (Fauth et al., 2010; Villa et al., 2016; Zheng et al., 2014). MSL3 belongs to the MGR15/MSL3 protein family and has, besides the C-terminal MRG domain required for interaction with MSL1, an N-terminal chromodomain (Buscaino et al., 2006; Morales et al., 2005). The chromodomain of MSL3 specifically recognizes the active chromatin mark H3K36me3 and facilitates DCC binding to its target genes (Bell et al., 2008; Larschan et al., 2007; Sural et al., 2008). MSL3's chromodomain is also described to bind DNA and RNA as well as methylated H4K20 (Akhtar et al., 2000; Kim et al., 2010; Moore et al., 2010). Furthermore, MSL3 recruits MOF via the interaction with MSL1 to nucleosomes and triggers acetylation of H4 (Conrad et al., 2012; Morales et al., 2005; Morales et al., 2004). MOF belongs to MYST-family of HATs and specifically catalyzes H4K16ac (Akhtar and Becker, 2000; Feller et al., 2015; Smith et al., 2000). MOF contains a C-terminal HAT domain, which catalyzes H4K16ac and mediates the interaction with MSL1, and a C2HC ZnF domain just N-terminal to the HAT domain required for substrate binding (Akhtar and Becker, 2001). Additionally, MOF has a chromobarrel domain located in the central part of the protein, which binds nucleic acids and stimulates the HAT activity (Akhtar et al., 2000; Conrad et al., 2012; Nielsen et al., 2005).

So far, all DCC subunits are assembled into a protein complex by protein-protein interactions with MSL1. The fifth subunit MLE seems to be more transiently connected to the DCC. MLE is an ATP-dependent RNA/DNA helicase of the DExH box subfamily and remodels the roX RNA for incorporation into the DCC (Ilik et al., 2013; Lee et al., 1997; Maenner et al., 2013; Meller et al., 2000; Militti et al., 2014; Morra et al., 2008; Prabu et al., 2015; Quinn et al., 2014; Richter et al., 1996). MLE has a modular architecture with the central catalytic domain consisting of two RecA ATPase domains and a helicase-associated 2 (HA2) domain. On the N-terminal site, MLE harbors two double-stranded (ds) RNA binding domains, essential for roX binding, and an OB-fold and glycine-rich region at the C-terminus (Ankush Jagtap et al., 2019; Izzo et al., 2008; Lv et al., 2019; Prabu et al., 2015). MLE especially binds and remodels both roX RNAs trough recognition of U-rich roX boxes located in prominent stem loop structures at the 3' end (Ilik et al., 2017; Ilik et al., 2013; Maenner et al., 2013; Park et al., 2007; Quinn et al., 2014; Stuckenholz et al., 2003; Sural et al., 2008). Even though, both roX RNAs share common features, they are very different in sequence composition and size. RoX1 is much longer with ~3.7 kb and roX2 is only ~0.6 kb. However, they appear to be genetically redundant as the individual mutants are viable and only the double mutant is male-specific-lethal (Meller and Rattner, 2002). While the roX RNAs are required for DCC binding to target genes and their up-regulation, the localization of their genes on the X chromosome is dispensable for X chromosome recognition (Deng and Meller, 2006;

Figueiredo et al., 2014; Ilik et al., 2017; Kageyama et al., 2001; Kelley et al., 1999; Meller, 2003; Meller et al., 2000; Meller and Rattner, 2002). Until now, it remains elusive, whether only one roX RNA is part of one DCC particle at any given time or both roX RNAs associate simultaneously. Also their precise function is unknown.



	% similarity	domain	nuc. acid binding	function
MSL1	21.8%	coiled-coil		interaction with MSL2
	23.6%	PEHE		interaction with MSL3, MOF
MSL2	32.5%	RING		E3 ubiquitin ligase, homeostasis
	40.5%	CXC	x	DNA binding, X targeting
MSL3	43.3%	Chromo	x	H3K36me3, H4K20me, DNA binding?
	22.4%	MRG		stimulates HAT activity
MOF	32.9%	Chromobarrel	x	modulates HAT activity
	63.0%	HAT		Histone Acetyltransferase
MLE		RB1/2	x	dsRNA binding
		ATPase		ATP hydrolysis
		Helicase		rox RNA unwinding
		G-rich	x	modulates RNA binding specificity

Figure 11. Protein domain architecture of MSL protein.

Schematic representation and summary table of the D. melanogaster and human MSL proteins. Figure is reproduced from (Keller and Akhtar, 2015) with permission of Elsevier.

The DCC combines multiple enzymatic activities and interaction surfaces for X chromosome recognition. Binding of the DCC to the X chromosome was studied in more detail by genomic approaches using chromatin immunoprecipitation followed by microarray hybridisation (ChIP-chip) or high-throughput sequencing (ChIP-seq). The DNA-binding MSL2 is the key subunit for DCC targeting to the X chromosome and H3K36me3 binding MSL3 for targeting the active genes on the X chromosome for transcription up-regulation through acetylation of H4K16 by MOF (Alekseyenko et al., 2012; Alekseyenko et al., 2006; Alekseyenko et al., 2008; Gelbart et al., 2009; Gilfillan et al., 2006; Kind et al., 2008; Larschan et al., 2007; Legube et al., 2006; Straub et al., 2008; Straub et al., 2013; Sural et al., 2008).

1.3.3 DCC targeting to the X chromosome

The DCC must distinguish the X chromosome from autosomes for faithful dosage compensation. The DCC targets the X chromosome through binding to ~300 high affinity sites (HAS), also referred to as chromosomal entry sites. The terms 'high affinity sites' or 'chromosomal entry sites' were defined already by genetic studies analysing MSL2 binding in absence of other subunits using immunofluorescence microscopy of polytene chromosome spreads (Dahlsveen et al., 2006; Demakova et al., 2003; Gilfillan et al., 2007; Gu et al., 1998; Kelley et al., 1999; Lyman et al., 1997). More recent studies using ChIP-seq showed that MSL2 is the central DNA binding subunit that recruits the DCC to HAS (Alekseyenko et al., 2008; Straub et al., 2008; Straub et al., 2013; Villa et al., 2016). Already, the early genetic studies noted that HAS are GA-rich. Indeed, the HAS contain a ~21 bp low complexity, GAGA-rich consensus sequence motif termed MSL recognition element (MRE) (Figure 12) (Alekseyenko et al., 2008; Straub et al., 2008).

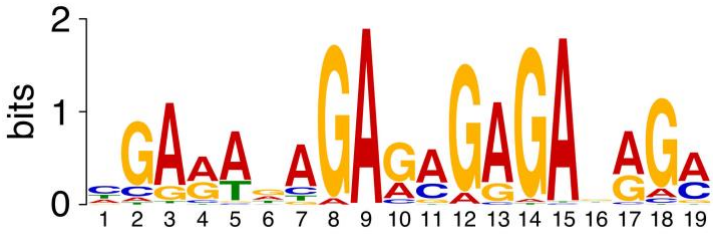


Figure 12. MSL recognition element.

Sequence motif of high affinity sites. Figure is adapted from (Villa et al., 2016) with permission of Springer Nature.

However, the MRE motif is not sufficient for X chromosome discrimination as the genome contains several thousand MRE motifs, which are only slightly enriched on the X chromosome. Furthermore, only ~2% of the MRE motifs are bound by the DCC. DCC recruitment to its genomic binding sites is modulated by epigenetic factors, as HAS are preferentially within the 3' end of active genes, in an active chromatin environment and depleted of nucleosomes (Alekseyenko et al., 2012; Alekseyenko et al., 2008; Ramirez et al., 2015; Straub et al., 2008; Straub et al., 2013). Furthermore, HAS tend to cluster as a function of chromosome domain interactions in the active X-chromosomal compartment, which is DCC-independent and invariant between males and females, as revealed by using the genome-wide chromosome conformation capture (Hi-C) and loci-specific (4C) approaches (Ramirez et al., 2015). However, earlier, more anecdotal DNA fluorescence *in situ* hybridization experiments suggested that the DCC brings some HAS in close proximity (Grimaud and Becker, 2009).

Most recently, using an *in vitro* genome-wide DNA immunoprecipitation (DIP) assay, the direct genomic binding sites of MSL2 were determined (Villa et al., 2016). Remarkably, *in vitro*, MSL2 selected sites containing the MRE motif and enriched X chromosomal sites, including ~20% of the HAS. More interestingly, binding of MSL2 to a subset of these sites required the presence of the CXC domain. These CXC-dependent sites were termed Pioneering-sites-on-the-X (PionX). The PionX harbour a consensus motif with a 5' extension to the MRE motif and a particular DNA shape with a high DNA roll between the first two base pairs. This PionX motif contributes to X chromosome discrimination as it is ~10-fold enriched on the X chromosome. However, the PionX signature, consisting of DNA motif and shape, is only present in a subset of HAS and therefore *per se* not sufficient to explain the binding of the DCC binding to all HAS. To address whether further proteins contribute to DCC targeting to HAS, Larschan and colleagues performed a gene-wide RNA interference (RNAi) screen couple to a reporter assay for DCC binding to HAS (Larschan et al., 2012). Besides general transcription regulators, they discovered some proteins with potential sequence-specific DNA binding activity contributing to DCC recruitment. One of the proteins with potential sequence-specific DNA binding activity is the gene product of CG1832. This protein was named Chromatin Linked Adaptor for MSL Proteins (CLAMP) as it is implicated in DCC recruitment to the X chromosome (Kaye et al., 2018; Larschan et al., 2012; Soruco et al., 2013). CLAMP is a seven C2H2 ZnF protein, with one N-terminal ZnF and a cluster of six ZnF at the C-terminus. The C-terminal six ZnF domains mediate sequence-specific DNA binding to the MRE motif (Kaye et al., 2018; Kuzu et al., 2016; Soruco et al., 2013). Furthermore, CLAMP is an essential gene in both sexes and binds to thousands of MRE sequences genome-wide (Soruco et al., 2013; Urban et al., 2017b). However, CLAMP binds to HAS mostly in males and in a DCC-dependent manner. There, CLAMP is suggested to interact with the DCC (Lindehell et al., 2015; Wang et al., 2013). Besides being implicated in DCC targeting, CLAMP is suggested to have further functions outside dosage compensation (Kaye et al., 2017; Rieder et al., 2017; Urban et al., 2017a; Urban et al., 2017c). CLAMP promotes long-range chromatin accessibility at its binding sites and increases global X chromosome accessibility in males. Furthermore, CLAMP is implicated in histone locus regulation and associates with boundary elements. Besides CLAMP, the small interfering RNA (siRNA) pathway, in particular siRNAs from the 1.688 g/cm$_3$ satellite repeats, may contribute to X chromosome recognition (Deshpande and Meller, 2018; Joshi and Meller, 2017; Menon et al., 2014; Menon and Meller, 2015).

Successful X chromosome discrimination and HAS targeting by the DCC is the first step in dosage compensation. In the second step, the DCC targets transcribed genes and boosts their transcription output.

1.3.4 Transcription up-regulation by the DCC

To prevent male lethality, the DCC must adjust the imbalance of X-linked gene expression between males and females. Therefore, the DCC reaches out from HAS to actively transcribed genes on the male X chromosome, making use of the chromosome interaction network (Ramirez et al., 2015). Genes actively transcribed by RNA polymerase II are marked by H3K36me3 at the gene body, a mark placed co-transcriptionally by the histone methyltransferase (HMT) Set2 (Bell et al., 2007; Carrozza et al., 2005; Larschan et al., 2007; Stabell et al., 2007). H3K36me3 is recognized by MSL3, which targets the DCC to actively transcribed genes for up-regulation (Bell et al., 2008; Larschan et al., 2007; Sural et al., 2008; Wang et al., 2013). There, MOF the histone acetyl transferase within the DCC acetylates H4K16 over gene bodies (Conrad et al., 2012; Gelbart et al., 2009; Kind et al., 2008). Outside of dosage compensation, MOF is also part of the NSL complex placing H4K16ac at promoters of housekeeping genes (Cai et al., 2010; Feller et al., 2012; Lam et al., 2012; Raja et al., 2010). However, H4K16ac at the gene body in the context of dosage compensation induces chromatin decondensation for transcription up-regulation in the two-fold range (Akhtar and Becker, 2000; Allahverdi et al., 2011; Bell et al., 2010; Corona et al., 2002; Liu et al., 2011; Prestel et al., 2010; Robinson et al., 2008; Shogren-Knaak et al., 2006; Smith et al., 2001). Transcription activation is achieved through increased transcription elongation of RNA polymerase II (Ferrari et al., 2013; Larschan et al., 2011). Further mechanisms could contribute to transcription up-regulation, for example reduced negative supercoiling mediated by topoisomerase II and supercoiling factor and the interaction of MSL1 with the elongation factor SPT5 and the general TF TFIIH via CDK7 (Chlamydas et al., 2016; Cugusi et al., 2013; Dunlap et al., 2012; Furuhashi et al., 2006; Prabhakaran and Kelley, 2012). Additional chromatin regulators are linked to dosage compensation and male X chromosome regulation, including the remodeler ISWI, the heterochromatin factors suppressor of variegation [Su(var)] 3-7 and HP1, the nuclear pore components Megator and Nup153, and the chromosomal kinase JIL-1, which is ~2-fold enriched on the male X chromosome (Badenhorst et al., 2002; Corona et al., 2002; de Wit et al., 2005; Jin et al., 2000; Mendjan et al., 2006; Regnard et al., 2011; Spierer et al., 2008; Spierer et al., 2005; Vaquerizas et al., 2010; Wang et al., 2001).

1.4 H3K36me3 chromatin

Methylation of histone lysine residues is associated with gene expression regulation. Up to three methyl-groups can be added to a lysine residue, which however retains its positive charge. Methylation of histone H3 at lysine 36 is a hallmark of active transcription and

increases gradually from mono- over di- to trimethylation from the 5’ to 3’ end of transcribed genes [reviewed in e.g. (Huang and Zhu, 2018; Li et al., 2016; McDaniel et al., 2017; Venkatesh and Workman, 2013; Wagner and Carpenter, 2012; Woo et al., 2017)]. Methylation of H3K36 is catalyzed by specific HMTs containing the common, catalytic SET domain. In *Saccharomyces cerevisiae*, a single HMT, Set2, catalyzes all three methylation states of H3K36, while higher eukaryotes have distinct enzymes for mono-/dimethylation and trimethylation. In *D. melanogaster*, Mes-4 mediates H3K36me1/me2, Ash1 mediates H3K36me2 and Set2 places H3K36me3 (Bell et al., 2007; Schmahling et al., 2018; Tanaka et al., 2007). The Set2 enzymes have a conserved domain architecture from yeast to mammals. In *S. cerevisiae*, the SET domain is located at the N-terminus while higher eukaryotes have an extended, low-complexity N-terminus preceding the SET domain. The Set2-Rpb1 interaction (SRI) domain is located at the C-terminus, mediating the interaction with RNA polymerase II. The SRI domain binds specifically to two consecutive heptapeptide repeats in the C-terminal domain (CTD) of RNA polymerase II containing serines at residue 2 and 5 in each repeat. The CTD of RNA polymerase II becomes differentially phosphorylated during the transcription cycle (Komarnitsky et al., 2000; Mayer et al., 2010). The S5 residues become phosphorylated during formation of the preinitiation complex by TFIIH, which leads to the recruitment of various elongation factors for transition of RNA polymerase II into active elongation. During the elongation phase, S5 becomes gradually dephosphorylated towards the 3’ end while S2 is gradually phosphorylated by P-TEFb and allows the binding of Set2 and transcription-coupled trimethylation of H3K36 (Carrozza et al., 2005; Gopalakrishnan et al., 2019; Kizer et al., 2005; Li et al., 2003; Stabell et al., 2007).

H3K36 trimethylation by Set2 and its biological roles are best studied in *S. cerevisiae*. There, other transcription elongation factors and histone chaperons are shown to contribute to the regulation of H3K36me3 through Set2. The H3/H4 histone chaperone Spt6, which also binds to RNA polymerase II CTD to facilitate nucleosome-reassembly in the wake of the transcribing polymerase, is required for binding of Interacts with Spt6-1 (IWS1), recruiting Set2 to the CTD (Gopalakrishnan et al., 2019; Yoh et al., 2008; Youdell et al., 2008). Furthermore, the elongation factors facilitates chromatin transcription (FACT) complex, Paf1 complex and anti-silencing factor 1 are implicated in Set2-mediated H3K36me3 (Chu et al., 2007; Chu et al., 2006; Lin et al., 2010). Heterogeneous nuclear ribonucleoprotein L associates with Set2 and stimulates H3K36me3, while the prolyl isomerase Fpr4 mediates H3P38 isomerization, which changes the conformation of H3K36 making trimethylation unfavorable (Nelson et al., 2006; Yuan et al., 2009). In addition, H3K36 methylation marks can be actively removed by Jumonji C domain containing lysine demethylases (KDM). The human KDM2A and KDM4A are shown to demethylate H3K36me2/3, in flies this is mediated

by Kdm2, Kdm4A and Kdm4B (Blackledge et al., 2010; Holowatyj et al., 2015; Whetstine et al., 2006; Zhou et al., 2012).

To translate the H3K36 methylation signal into a biological output, various readers recognize the different methylation states of H3K36 and trigger associated pathways. One function of Set2-mediated H3K36me3 in yeast is the prevention of aberrant transcription initiation within the coding region of transcribed genes (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005; Sun et al., 2008). The histone deacetylase complex reduced potassium dependency 3 small (Rpd3S) is recruited to H3K36me3 through the binding of the subunit ESA1-associated factor 3 (Esa3) [MORF-related gene 15 (MRG15) in humans and flies] by its chromodomain to maintain a hypoacetylated state at the gene body. The Isw1b ATP-dependent chromatin remodeling complex contributes to this pathway by properly spacing nucleosome providing a suitable substrate for deacetylation (Lee et al., 2013; Maltby et al., 2012; Smolle et al., 2012). In the Isw1b complex, Isw1 interacts with Ioc4 a PWWP domain containing protein, which recruits the complex to H3K36me3. In humans, a histone deacetylase complex is also recruited to H3K36me3 via the chromodomain-containing MRG15 subunit (Jelinic et al., 2011; Kumar et al., 2012). In addition, various PWWP domain-containing proteins are recruited to H3K36me3, including ZMYND11, which is implicated in transcription elongation and mRNA splicing, DNA (cytosine-5)-methyltransferase 3A and B to establish *de novo* DNA methylation, and GLYR1 interacting with a H3K4 demethylase and responsible for nucleosome-destabilization to facilitate transcription. (Baubec et al., 2015; Dhayalan et al., 2010; Fang et al., 2013; Fei et al., 2018; Guo et al., 2014; Morselli et al., 2015; Suzuki et al., 2006; Vermeulen et al., 2010; Wen et al., 2014). H3K36me3 readers include PSIP1, which is implicated in various processes from alternative splicing, to DNA double strand repair by homologous recombination, to DNA mismatch repair mediated by MSH6, to integration of the human immunodeficiency virus in active portions of the genome (Cherepanov et al., 2005; Ciuffi et al., 2005; Daugaard et al., 2012; Laguri et al., 2008; Li et al., 2013a; Pfister et al., 2014; Pradeepa et al., 2012; Sharma et al., 2018). Moreover, the activity of the Polycomb repressive complex 2 (PRC2) is modulated by H3K36me3, which is bound by the Tudor domain containing subunits PHF1 and PHF19 (Ballare et al., 2012; Brien et al., 2012; Cai et al., 2013; Schmitges et al., 2011; Yuan et al., 2011).

In *D. melanogaster*, MSL3 is the best studied protein shown to bind H3K36me3 via its chromodomain for dosage compensation (Larschan et al., 2007; Sural et al., 2008). Furthermore, the chromodomain-containing MRG15 is suggested to bind H3K36me3, as shown for its yeast homolog (Filion et al., 2010; Joshi and Struhl, 2005; Sun et al., 2008). MRG15 in flies is part of the Tip60 HAT complex/Domino remodeling complex and Ash1 HMT complex (Huang et al., 2017; Kusch et al., 2004; Schmahling et al., 2018). Furthermore,

the PWWP domain-containing NDF is suggested to bind H3K36me3, as shown for its human homolog GLYR1, to facilitate transcription by destabilizing nucleosomes (Fang et al., 2013; Fei et al., 2018; Vermeulen et al., 2010; Wang et al., 2013). Although the chromosomal JIL-1 kinase is not shown to directly bind H3K36me3, JIL-1 co-localizes genome-wide with the H3K36me3 mark (Regnard et al., 2011).

1.4.1 The chromosomal kinase JIL-1

In the late 1990s, the kinase JIL-1 was discovered in the lab of K. M. Johansen and J. Johansen, who most contributed to our understanding of JIL-1 function (Jin et al., 1999). JIL-1 harbors two central serine/threonine kinase domains arranged in tandem. The JIL-1 N-terminus contains an asparagine-rich and an alanine-rich stretch, and a bipartite nuclear localization signal. Furthermore, two PEST-related sequences are located in the N-terminus and one between the kinase domains, which are implicated in protein degradation (Rogers et al., 1986). JIL-1 is expressed ubiquitously at all developmental stages and associates to chromatin at all cell cycle phases (Jin et al., 1999; Zhang et al., 2003). Based on polytene chromosome squashes, JIL-1 localizes to an active chromatin environment at the decondensed interbands and is ~2-fold enriched on the dosage-compensated male X chromosome (Jin et al., 2000; Jin et al., 1999). Basal JIL-1 levels at gene bodies correlate with H3K36me3 levels and enrichment at male X-linked genes with additional H4K16ac levels (Regnard et al., 2011). Since JIL-1 is enriched on the male X chromosome in a DCC dependent way, it is linked to dosage compensation. Moreover, JIL-1 binds MSL1 and MSL3 via its kinase domains and is required for full transcription up-regulation of X-linked genes by the DCC (Jin et al., 2000; Regnard et al., 2011; Wang et al., 2013). At active chromatin, JIL-1 phosphorylates H3 at serine 10 (H3S10ph) during interphase, while H3S10ph in mitosis is catalyzed by aurora B kinase (Adams et al., 2001; Giet and Glover, 2001; Regnard et al., 2011; Wang et al., 2001). JIL-1 is an essential protein in both sexes since it is implicated in higher-order chromosome organization and ectopic recruitment causes chromatin reorganization (Bao et al., 2008; Deng et al., 2008; Deng et al., 2005; Li et al., 2013b; Wang et al., 2001). H3S10 phosphorylation by JIL-1 during interphase counteracts heterochromatin spreading by preventing H3K9 di-/trimethylation by Su(var)3-9 and subsequent binding of HP1 in euchromatic regions (Figure 13) (Cai et al., 2014; Deng et al., 2007; Wang et al., 2011a; Wang et al., 2012; Wang et al., 2014; Zhang et al., 2006). The function of JIL-1 in the maintainance of the hetero- and euchromatin boundary is best illustrated in the phenomenon of position-effect variegation (Bao et al., 2007; Deng et al., 2010; Girton et al., 2013; Lerach et al., 2006; Wang et al., 2011b). Indeed, truncations within JIL-1's C-terminus have a Su(var) phenotype, counteracting heterochromatin spreading into euchromatin (Ebert et al.,

2004). Additionally, JIL-1 is shown to phosphorylate the heterochromatin factor Su(var)3-9, the HMT specific for di- and trimethylation of H3 at lysine 9 (H3K9me2/3), at its N-terminus, which is implicated in protein-protein interaction and chromatin binding (Boeke et al., 2010; Melcher et al., 2000; Schotta et al., 2002). Besides localizing to active H3K36me3 chromatin at gene bodies, JIL-1 is also found at the arrays of retrotransposons at telomeres, consisting of the non-LTR retrotransposons HeT-A, TART and TAHRE (HTT arrays), to regulates their expression (Andreyeva et al., 2005; Silva-Sousa and Casacuberta, 2013; Silva-Sousa et al., 2012).

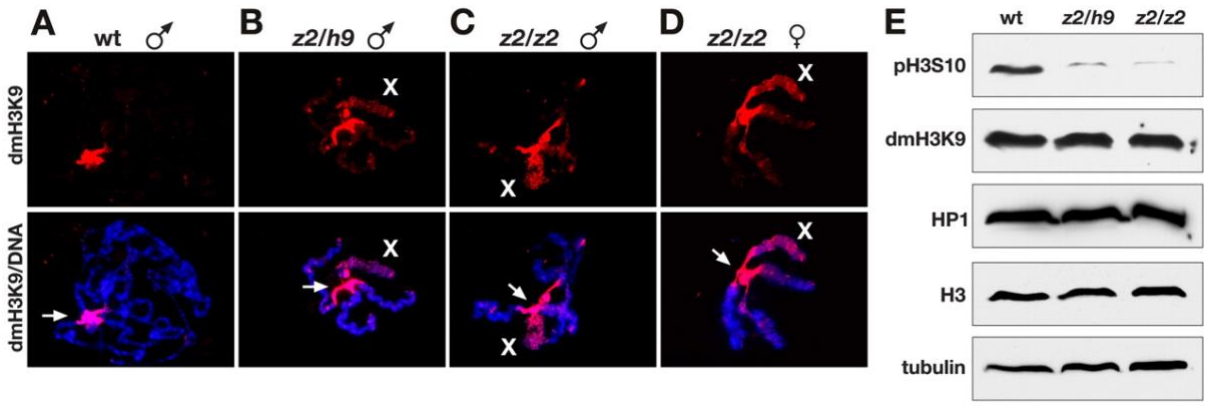


Figure 13. Heterochromatin spreading in JIL-1 mutants.

Left side, polytene squashs from wild type and JIL-1 mutant L3 larvae stained with α-H3K9me2 antibody in red and with Hoechst in blue. Right side, western blots analysis of wild type and JIL-1 mutant larval extracts for H3S10ph, H3K9me2 and HP1. Figure is reproduced from (Zhang et al., 2006) with permission of Company of Biologists.

2 Aims and objectives

The first aim was to mechanistically dissect the contribution of CLAMP to MSL2 targeting to HAS on the male X chromosome. To this end, MSL2 binding *in vivo* in absence of CLAMP was monitored by high-resolution ChIP-seq, as potentially MSL2 alone could be sufficient to bind PionX *in vivo*. To analyze cooperative effects between CLAMP and MSL2 for HAS targeting, *in vitro* genome-wide DNA immunoprecipitation with high-throughput sequencing (DIP-seq) was used (Gossett and Lieb, 2008; Liu et al., 2005). As direct protein-protein interactions are implicated in cooperative DNA binding, CLAMP interaction with MSL2 was assayed by co-IP experiments to map individual interaction domains. Further, chromatin organization could also contribute to binding site selection and cooperativity, genome-wide chromatin accessibility was studied in male and female cells after CLAMP and MSL2 depletion by assay for transposase accessibly chromatin with high-throughput sequencing (ATAC-seq) (Buenrostro et al., 2013; Buenrostro et al., 2015).

The second aim was to refine the role of chromosome architecture in DCC transfer from HAS to target genes. Therefore, chromosome architecture and compartmentalization in male and female embryos was analyzed by genome-wide chromosome conformation capture (Hi-C) (Sexton et al., 2012). To investigate potential sex-dependent differences, Hi-C data were connected to H3K36me3 and H4K16ac chromatin marks assayed by ChIP-seq as well as to gene expression measured by RNA sequencing (RNA-seq). To more specifically study the contact frequency of PionX with the active compartment, locus-specific 4C-seq after MSL2 depletion in male cells or *de novo* induction of DCC formation by *sxl* RNAi in female cells was performed and related to gene expression changes (Ghavi-Helm et al., 2014).

The third aim was to reveal the targeting mechanism of JIL-1 to H3K36me3 chromatin and its role in transcription regulation. The interaction of JIL-1 with the new PWWP domain containing protein, JIL-1 Anchoring and Stabilizing Protein (JASPer), was identified by co-IP experiments. To map the individual interaction domains, the JIL-1-JASPer complex (JJ-complex) was recombinantly expressed for subsequent co-IP assays. To analyze whether JASPer can recruit JIL-1 to H3K36me3 chromatin, *in vitro* nucleosome pull-down experiments using recombinant JASPer and JJ-complex were performed by employing a nucleosome library comprising 115 different nucleosome types (Dann et al., 2017). Furthermore, the *in vivo* localization of JASPer and JIL-1 relative to H3K36me3 was studied by high-resolution ChIP-seq in male and female cells. To gain mechanistic insight into JJ-complex enriched at the male X chromosome, JASPer, H4K16ac and MSL3 ChIP-seq after JIL-1 depletion were performed. Potential spreading of H3K9me2 into euchromatin after JIL-1 depletion was also studied by ChIP-seq. Moreover, the role of the JJ-complex in gene expression regulation and expression of transposable elements was studied by RNA-seq

after JASPer and JIL-1 depletion in male and female cells. The association of JJ-complex with other chromatin factors was investigated by co-IP coupled to mass spectrometry analysis.

3 Results

3.1 Factor cooperation for chromosome discrimination in *Drosophila*

Factor cooperation for chromosome discrimination in *Drosophila*

Christian Albig[1,2], Evgeniya Tikhonova[3], Silke Krause[1], Oksana Maksimenko[3], Catherine Regnard[1], and Peter B. Becker[1]

[1]Molecular Biology Division, Biomedical Center, Faculty of Medicine and Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-University, 82152 Martinsried, Germany;

[2]Graduate School for Quantitative Biosciences (QBM), Ludwig-Maximilians-University, 81377 Munich, Germany;

[3]Department of the Control of Genetic Processes, Institute of Gene Biology, Russian Academy of Sciences, 119334 Moscow, Russia;

Nucleic Acids Research, Volume 47, Issue 4, 28 February 2019, Pages 1706–1724, DOI 10.1093/nar/gky1238

Authors' contributions:

C.A., C.R. and P.B.B. conceived the study; C.A. performed all experiments except for co-immunoprecipitation and Yeast Two-Hybrid experiments; S.K. expressed proteins and performed co-immunoprecipitation experiments; E.T. and O.M. conceived, performed and analyzed Yeast Two-Hybrid experiments; C.A. performed bioinformatics analyses; All authors analyzed data; C.R. and P.B.B. provided feedback and supervision; C.A., C.R. and P.B.B. wrote the manuscript; P.B.B. secured funding.

3.2 Chromosome topology guides the *Drosophila* Dosage Compensation Complex for target gene activation

Chromosome topology guides the *Drosophila* Dosage Compensation Complex for target gene activation

Tamás Schauer[1], Yad Ghavi-Helm[2], Tom Sexton[3], Christian Albig[1], Catherine Regnard[1], Giacomo Cavalli[4,5], Eileen EM Furlong[2] and Peter B Becker[1]

[1]Molecular Biology Division, Biomedical Center and Center for Integrated Protein Science Ludwig-Maximilians-University, Munich, Germany

[2]European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg, Germany

[3]Institute of Genetics and Molecular and Cellular Biology, Illkirch, France

[4]Institute of Human Genetics, CNRS, Montpellier, France

[5]University of Montpellier, Montpellier, France

EMBO Reports, Volume 18, Issue 10, 1 October 2017, Pages 1854–1868, DOI 10.15252/embr.201744292

Authors' contributions:

The 4C-seq, ChIP-seq and RNA-seq experiments with Kc cells were performed by T.Sc.; 4C-seq was designed and supervised by Y.G.-H. and E.E.M.F.; Hi-C was performed by T.Se. and supervised by G.C.; ChIP-seq on H3K36me3 was carried out by C.A.; RNA-seq in S2 cells was performed by C.R.; Datasets were analyzed and visualized by T.Sc.; The study was conceptualized and supervised by P.B.B.; T.Sc. and P.B.B. wrote the manuscript with contributions from all authors.

3.3 JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*

JASPer controls interphase histone H3S10 phosphorylation by chromosomal kinase JIL-1 in *Drosophila*

Christian Albig[1,2], Chao Wang[3], Geoffrey P. Dann[4,5], Felix Wojcik[4], Tamás Schauer[6], Silke Krause[1], Sylvain Maenner[1,7], Weili Cai[3], Yeran Li[3], Jack Girton[3], Tom W. Muir[4], Jørgen Johansen[3], Kristen M. Johansen[3], Peter B. Becker[1,*] and Catherine Regnard[1,*]

[1]Molecular Biology Division, Biomedical Center, Faculty of Medicine and Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-University of Munich, 82152 Martinsried, Germany;

[2]Graduate School for Quantitative Biosciences (QBM), Ludwig-Maximilians-University of Munich, 81377 Munich, Germany;

[3]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa, 50011, USA;

[4]Department of Chemistry, Princeton University, Frick Laboratory, Princeton, New Jersey, 08544, USA;

[5]Current address: Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, USA;

[6]Bioinformatics Unit, Biomedical Center, Faculty of Medicine, Ludwig-Maximilians-University of Munich, 82152 Martinsried, Germany;

[7]Current address: UMR7365 CNRS-UL, IMoPA, University of Lorraine, 54505 Vandoeuvre-les Nancy, France;

Under revision at Nature Communications, bioRxiv, 30 April 2019, 624023, DOI 10.1101/624023. [21.01.2020: Nat Communications *10*, 5343 (2019). DOI 10.1038/s41467-019-13174-6]

Authors' contributions:

C.R. conceived this study and performed experiments; C.A. performed MNase and sonication ChIP-seq experiments and all bioinformatics analysis also with support from T.S.; C.W. generated and characterized the cw2 mutant line with help from J.G., W.C. did the LacO-LacI targeting experiments with support from Y.L., and J.J. and K.M.J. supervised the work and secured funding; G.P.D. performed mononucleosome library experiments and F.W. generated the mononucleosomes and arrays for the kinase assays in T.W.M.'s lab; S.K. prepared recombinant proteins for all *in vitro* assays, RNA-seq experiments under the supervision of C.R. and spike-in ChIP-seq experiments under supervision of C.A.; All authors analyzed data; C.R. and C.A. wrote the manuscript with contributions from all authors; P.B.B. secured funding and established collaborations.

4 References

Abel, K.J., Brody, L.C., Valdes, J.M., Erdos, M.R., McKinley, D.R., Castilla, L.H., Merajver, S.D., Couch, F.J., Friedman, L.S., Ostermeyer, E.A.*, et al.* (1996). Characterization of EZH1, a human homolog of Drosophila Enhancer of zeste near BRCA1. Genomics *37*, 161-171.

Adams, R.R., Maiato, H., Earnshaw, W.C., and Carmena, M. (2001). Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. The Journal of cell biology *153*, 865-880.

Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Molecular cell *5*, 367-375.

Akhtar, A., and Becker, P.B. (2001). The histone H4 acetyltransferase MOF uses a C2HC zinc finger for substrate recognition. EMBO Rep *2*, 113-118.

Akhtar, A., Zink, D., and Becker, P.B. (2000). Chromodomains are protein-RNA interaction modules. Nature *407*, 405-409.

Albritton, S.E., and Ercan, S. (2018). Caenorhabditis elegans Dosage Compensation: Insights into Condensin-Mediated Gene Regulation. Trends in genetics : TIG *34*, 41-53.

Alekseyenko, A.A., Ellison, C.E., Gorchakov, A.A., Zhou, Q., Kaiser, V.B., Toda, N., Walton, Z., Peng, S., Park, P.J., Bachtrog, D.*, et al.* (2013). Conservation and de novo acquisition of dosage compensation on newly evolved sex chromosomes in Drosophila. Genes & development *27*, 853-858.

Alekseyenko, A.A., Gorchakov, A.A., Zee, B.M., Fuchs, S.M., Kharchenko, P.V., and Kuroda, M.I. (2014). Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. Genes & development *28*, 1445-1460.

Alekseyenko, A.A., Ho, J.W., Peng, S., Gelbart, M., Tolstorukov, M.Y., Plachetka, A., Kharchenko, P.V., Jung, Y.L., Gorchakov, A.A., Larschan, E.*, et al.* (2012). Sequence-specific targeting of dosage compensation in Drosophila favors an active chromatin context. PLoS genetics *8*, e1002646.

Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes & development *20*, 848-857.

Alekseyenko, A.A., Peng, S., Larschan, E., Gorchakov, A.A., Lee, O.K., Kharchenko, P., McGrath, S.D., Wang, C.I., Mardis, E.R., Park, P.J.*, et al.* (2008). A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. Cell *134*, 599-609.

Allahverdi, A., Yang, R., Korolev, N., Fan, Y., Davey, C.A., Liu, C.F., and Nordenskiold, L. (2011). The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association. Nucleic acids research *39*, 1680-1691.

Altmann, R. (1889). Ueber Nucleinsäuren. Archiv für Anatomie und Physiology, Pysiologische Abteilung, Leipzig, 524-536.

Amrein, H., and Axel, R. (1997). Genes expressed in neurons of adult male Drosophila. Cell *88*, 459-469.

Andersen, S.U., Algreen-Petersen, R.G., Hoedl, M., Jurkiewicz, A., Cvitanich, C., Braunschweig, U., Schauser, L., Oh, S.A., Twell, D., and Jensen, E.O. (2007). The conserved cysteine-rich domain of a tesmin/TSO1-like protein binds zinc in vitro and TSO1 is required for both male and female fertility in Arabidopsis thaliana. J Exp Bot *58*, 3657-3670.

Andrews, A.J., and Luger, K. (2011). Nucleosome Structure(s) and Stability: Variations on a Theme. Annu Rev Biophys *40*, 99-117.

Andrews, F.H., Strahl, B.D., and Kutateladze, T.G. (2016). Insights into newly discovered marks and readers of epigenetic information. Nat Chem Biol *12*, 662-668.

Andreyeva, E.N., Belyaeva, E.S., Semeshin, V.F., Pokholkova, G.V., and Zhimulev, I.F. (2005). Three distinct chromatin domains in telomere ends of polytene chromosomes in Drosophila melanogaster Tel mutants. J Cell Sci *118*, 5465-5477.

Ankush Jagtap, P.K., Muller, M., Masiewicz, P., von Bulow, S., Hollmann, N.M., Chen, P.C., Simon, B., Thomae, A.W., Becker, P.B., and Hennig, J. (2019). Structure, dynamics and roX2-lncRNA binding of tandem double-stranded RNA binding domains dsRBD1,2 of Drosophila helicase Maleless. Nucleic acids research, gkz125.

Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. J Exp Med *79*, 137-158.

Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. Genes & development *16*, 3186-3198.

Ballare, C., Lange, M., Lapinaite, A., Martin, G.M., Morey, L., Pascual, G., Liefke, R., Simon, B., Shi, Y., Gozani, O.*, et al.* (2012). Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nature structural & molecular biology *19*, 1257-1265.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res *21*, 381-395.

Bao, X., Cai, W., Deng, H., Zhang, W., Krencik, R., Girton, J., Johansen, J., and Johansen, K.M. (2008). The COOH-terminal domain of the JIL-1 histone H3S10 kinase interacts with histone H3 and is required for correct targeting to chromatin. The Journal of biological chemistry *283*, 32741-32750.

Bao, X., Deng, H., Johansen, J., Girton, J., and Johansen, K.M. (2007). Loss-of-function alleles of the JIL-1 histone H3S10 kinase enhance position-effect variegation at pericentric sites in Drosophila heterochromatin. Genetics *176*, 1355-1358.

Bashaw, G.J., and Baker, B.S. (1995). The msl-2 dosage compensation gene of Drosophila encodes a putative DNA-binding protein whose expression is sex specifically regulated by Sex-lethal. Development *121*, 3245-3258.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243-247.

Beall, E.L., Manak, J.R., Zhou, S., Bell, M., Lipsick, J.S., and Botchan, M.R. (2002). Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. Nature *420*, 833-837.

Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. Trends in genetics : TIG *32*, 29-41.

Becker, P.B., and Workman, J.L. (2013). Nucleosome remodeling and epigenetics. Cold Spring Harb Perspect Biol *5*, a017905.

Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C.*, et al.* (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. Molecular cell *66*, 384-397.

Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation

by enhancing recruitment of the MSL complex in Drosophila melanogaster. Molecular and cellular biology *28*, 3401-3409.

Bell, O., Schwaiger, M., Oakeley, E.J., Lienert, F., Beisel, C., Stadler, M.B., and Schubeler, D. (2010). Accessibility of the Drosophila genome discriminates PcG repression, H4K16 acetylation and replication timing. Nature structural & molecular biology *17*, 894-900.

Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A.*, et al.* (2007). Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. EMBO J *26*, 4974-4984.

Belote, J.M., and Lucchesi, J.C. (1980a). Control of X chromosome transcription by the maleless gene in Drosophila. Nature *285*, 573-575.

Belote, J.M., and Lucchesi, J.C. (1980b). Male-specific lethal mutations of Drosophila melanogaster. Genetics *96*, 165-186.

Bhaumik, S.R., Smith, E., and Shilatifard, A. (2007). Covalent modifications of histones during development and disease pathogenesis. Nature structural & molecular biology *14*, 1008-1016.

Bian, Q., and Belmont, A.S. (2012). Revisiting higher-order and large-scale chromatin organization. Curr Opin Cell Biol *24*, 359-366.

Birchler, J.A. (2016). Parallel Universes for Models of X Chromosome Dosage Compensation in Drosophila: A Review. Cytogenet Genome Res *148*, 52-67.

Blackledge, N.P., Zhou, J.C., Tolstorukov, M.Y., Farcas, A.M., Park, P.J., and Klose, R.J. (2010). CpG islands recruit a histone H3 lysine 36 demethylase. Molecular cell *38*, 179-190.

Boeke, J., Regnard, C., Cai, W., Johansen, J., Johansen, K.M., Becker, P.B., and Imhof, A. (2010). Phosphorylation of SU(VAR)3-9 by the chromosomal kinase JIL-1. PloS one *5*, e10042.

Bone, J.R., Lavender, J., Richman, R., Palmer, M.J., Turner, B.M., and Kuroda, M.I. (1994). Acetylated histone H4 on the male X chromosome is associated with dosage compensation in Drosophila. Genes & development *8*, 96-104.

Bonora, G., and Disteche, C.M. (2017). Structural aspects of the inactive X chromosome. Philos Trans R Soc Lond B Biol Sci *372*, 20160357.

Brayer, K.J., and Segal, D.J. (2008). Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell biochemistry and biophysics *50*, 111-131.

Brien, G.L., Gambero, G., O'Connell, D.J., Jerman, E., Turner, S.A., Egan, C.M., Dunne, E.J., Jurgens, M.C., Wynne, K., Piao, L.*, et al.* (2012). Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. Nature structural & molecular biology *19*, 1273-1281.

Brockdorff, N. (2017). Polycomb complexes in X chromosome inactivation. Philos Trans R Soc Lond B Biol Sci *372*, 20170021.

Brockdorff, N., and Turner, B.M. (2015). Dosage compensation in mammals. Cold Spring Harb Perspect Biol *7*, a019406.

Brown, R.S. (2005). Zinc finger proteins: getting a grip on RNA. Current opinion in structural biology *15*, 94-98.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature methods *10*, 1213-1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol *109*, 21291-21299.

Buscaino, A., Legube, G., and Akhtar, A. (2006). X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. EMBO Rep *7*, 531-538.

Cai, L., Rothbart, S.B., Lu, R., Xu, B., Chen, W.Y., Tripathy, A., Rockowitz, S., Zheng, D., Patel, D.J., Allis, C.D.*, et al.* (2013). An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Molecular cell *49*, 571-582.

Cai, W., Wang, C., Li, Y., Yao, C., Shen, L., Liu, S., Bao, X., Schnable, P.S., Girton, J., Johansen, J.*, et al.* (2014). Genome-wide analysis of regulation of gene expression and H3K9me2 distribution by JIL-1 kinase mediated histone H3S10 phosphorylation in Drosophila. Nucleic acids research *42*, 5456-5467.

Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2010). Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex. The Journal of biological chemistry *285*, 4268-4272.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Celniker, S.E., Dillon, L.A.L., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M.*, et al.* (2009). Unlocking the secrets of the genome. Nature *459*, 927-930.

Chandler, C.H. (2017). When and why does sex chromosome dosage compensation evolve? Ann N Y Acad Sci *1389*, 37-51.

Chen, H., Rossier, C., and Antonarakis, S.E. (1996). Cloning of a human homolog of the Drosophila enhancer of zeste gene (EZH2) that maps to chromosome 21q22.2. Genomics *38*, 30-37.

Cherepanov, P., Ambrosio, A.L., Rahman, S., Ellenberger, T., and Engelman, A. (2005). Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc Natl Acad Sci U S A *102*, 17308-17313.

Chlamydas, S., Holz, H., Samata, M., Chelmicki, T., Georgiev, P., Pelechano, V., Dundar, F., Dasmeh, P., Mittler, G., Cadete, F.T.*, et al.* (2016). Functional interplay between MSL1 and CDK7 controls RNA polymerase II Ser5 phosphorylation. Nature structural & molecular biology *23*, 580-589.

Chou, C.C., Wei, S.Y., Lou, Y.C., and Chen, C. (2017). In-depth study of DNA binding of Cys2His2 finger domains in testis zinc-finger protein. PloS one *12*, e0175051.

Chu, Y., Simic, R., Warner, M.H., Arndt, K.M., and Prelich, G. (2007). Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. EMBO J *26*, 4646-4656.

Chu, Y., Sutton, A., Sternglanz, R., and Prelich, G. (2006). The BUR1 cyclin-dependent protein kinase is required for the normal pattern of histone methylation by SET2. Molecular and cellular biology *26*, 3029-3038.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J.R., and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. Nat Med *11*, 1287-1289.

Clapier, C.R., and Cairns, B.R. (2009). The Biology of Chromatin Remodeling Complexes. Annu Rev Biochem *78*, 273-304.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat Rev Mol Cell Biol *18*, 407-422.

Conrad, T., and Akhtar, A. (2012). Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. Nature reviews Genetics *13*, 123-134.

Conrad, T., Cavalli, F.M., Holz, H., Hallacli, E., Kind, J., Ilik, I., Vaquerizas, J.M., Luscombe, N.M., and Akhtar, A. (2012). The MOF chromobarrel domain controls genome-wide H4K16 acetylation and spreading of the MSL complex. Dev Cell *22*, 610-624.

Copps, K., Richman, R., Lyman, L.M., Chang, K.A., Rampersad-Ammons, J., and Kuroda, M.I. (1998). Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. EMBO J *17*, 5409-5417.

Corona, D.F., Clapier, C.R., Becker, P.B., and Tamkun, J.W. (2002). Modulation of ISWI function by site-specific histone acetylation. EMBO Rep *3*, 242-247.

Cremer, T., and Cremer, C. (1988). Centennial of Wilhelm Waldeyer's introduction of the term "chromosome" in 1888. Cytogenet Cell Genet *48*, 65-67.

Cugusi, S., Ramos, E., Ling, H., Yokoyama, R., Luk, K.M., and Lucchesi, J.C. (2013). Topoisomerase II plays a role in dosage compensation in Drosophila. Transcription *4*, 238-250.

Cutter, A.R., and Hayes, J.J. (2015). A brief review of nucleosome structure. FEBS Lett *589*, 2914-2922.

Cvitanich, C., Pallisgaard, N., Nielsen, K.A., Hansen, A.C., Larsen, K., Pihakaski-Maunsbach, K., Marcker, K.A., and Jensen, E.O. (2000). CPP1, a DNA-binding protein involved in the expression of a soybean leghemoglobin c3 gene. Proc Natl Acad Sci U S A *97*, 8163-8168.

Dahlsveen, I.K., Gilfillan, G.D., Shelest, V.I., Lamm, R., and Becker, P.B. (2006). Targeting determinants of dosage compensation in Drosophila. PLoS genetics *2*, e5.

Dann, G.P., Liszczak, G.P., Bagert, J.D., Muller, M.M., Nguyen, U.T.T., Wojcik, F., Brown, Z.Z., Bos, J., Panchenko, T., Pihl, R.*, et al.* (2017). ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. Nature *548*, 607-611.

Daugaard, M., Baude, A., Fugger, K., Povlsen, L.K., Beck, H., Sorensen, C.S., Petersen, N.H., Sorensen, P.H., Lukas, C., Bartek, J.*, et al.* (2012). LEDGF (p75) promotes DNA-end resection and homologous recombination. Nature structural & molecular biology *19*, 803-810.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9Å Resolution. Journal of molecular biology *319*, 1097-1113.

de Wit, E., Greil, F., and van Steensel, B. (2005). Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. Genome research *15*, 1265-1273.

Demakova, O.V., Kotlikova, I.V., Gordadze, P.R., Alekseyenko, A.A., Kuroda, M.I., and Zhimulev, I.F. (2003). The MSL complex levels are critical for its correct targeting to the chromosomes in Drosophila melanogaster. Chromosoma *112*, 103-115.

Deng, H., Bao, X., Cai, W., Blacketer, M.J., Belmont, A.S., Girton, J., Johansen, J., and Johansen, K.M. (2008). Ectopic histone H3S10 phosphorylation causes chromatin structure remodeling in Drosophila. Development *135*, 699-705.

Deng, H., Bao, X., Zhang, W., Girton, J., Johansen, J., and Johansen, K.M. (2007). Reduced levels of Su(var)3-9 but not Su(var)2-5 (HP1) counteract the effects on chromatin structure and viability in loss-of-function mutants of the JIL-1 histone H3S10 kinase. Genetics *177*, 79-87.

Deng, H., Cai, W., Wang, C., Lerach, S., Delattre, M., Girton, J., Johansen, J., and Johansen, K.M. (2010). JIL-1 and Su(var)3-7 interact genetically and counteract each other's effect on position-effect variegation in Drosophila. Genetics *185*, 1183-1192.

Deng, H., Zhang, W., Bao, X., Martin, J.N., Girton, J., Johansen, J., and Johansen, K.M. (2005). The JIL-1 kinase regulates the structure of Drosophila polytene chromosomes. Chromosoma *114*, 173-182.

Deng, X., Berletch, J.B., Ma, W., Nguyen, D.K., Hiatt, J.B., Noble, W.S., Shendure, J., and Disteche, C.M. (2013). Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. Dev Cell *25*, 55-68.

Deng, X., and Meller, V.H. (2006). roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. Genetics *174*, 1859-1866.

Deshpande, N., and Meller, V.H. (2018). Chromatin That Guides Dosage Compensation Is Modulated by the siRNA Pathway in Drosophila melanogaster. Genetics *209*, 1085-1097.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. The Journal of biological chemistry *285*, 26114-26120.

Disteche, C.M. (2012). Dosage compensation of the sex chromosomes. Annual review of genetics *46*, 537-560.

Disteche, C.M. (2016). Dosage compensation of the sex chromosomes and autosomes. Semin Cell Dev Biol *56*, 9-18.

Doenecke, D., and Karlson, P. (1984). Albrecht Kossel and the discovery of histones. Trends Biochem Sci *9*, 404-405.

Dunlap, D., Yokoyama, R., Ling, H., Sun, H.Y., McGill, K., Cugusi, S., and Lucchesi, J.C. (2012). Distinct contributions of MSL complex subunits to the transcriptional enhancement responsible for dosage compensation in Drosophila. Nucleic acids research *40*, 11281-11291.

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., and Reuter, G. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. Genes & development *18*, 2973-2983.

Eidahl, J.O., Crowe, B.L., North, J.A., McKee, C.J., Shkriabai, N., Feng, L., Plumb, M., Graham, R.L., Gorelick, R.J., Hess, S.*, et al.* (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic acids research *41*, 3924-3936.

Elgin, S.C., and Reuter, G. (2013). Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. Cold Spring Harb Perspect Biol *5*, a017780.

Elrod-Erickson, M., Rould, M.A., Nekludova, L., and Pabo, C.O. (1996). Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. Structure *4*, 1171-1180.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nature Biotechnology *28*, 817-825.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nature methods *9*, 215-216.

Fang, R., Chen, F., Dong, Z., Hu, D., Barbera, A.J., Clark, E.A., Fang, J., Yang, Y., Mei, P., Rutenberg, M.*, et al.* (2013). LSD2/KDM1B and its cofactor NPAC/GLYR1 endow a structural and molecular model for regulation of H3K4 demethylation. Molecular cell *49*, 558-570.

Fauth, T., Muller-Planitz, F., Konig, C., Straub, T., and Becker, P.B. (2010). The DNA binding CXC domain of MSL2 is required for faithful targeting the Dosage Compensation Complex to the X chromosome. Nucleic acids research *38*, 3209-3221.

Fei, J., Ishii, H., Hoeksema, M.A., Meitinger, F., Kassavetis, G.A., Glass, C.K., Ren, B., and Kadonaga, J.T. (2018). NDF, a nucleosome-destabilizing factor that facilitates transcription through nucleosomes. Genes & development *32*, 682-694.

Feller, C., Forne, I., Imhof, A., and Becker, P.B. (2015). Global and specific responses of the histone acetylome to systematic perturbation. Molecular cell *57*, 559-571.

Feller, C., Prestel, M., Hartmann, H., Straub, T., Soding, J., and Becker, P.B. (2012). The MOF-containing NSL complex associates globally with housekeeping genes, but activates only a defined subset. Nucleic acids research *40*, 1509-1522.

Ferrari, F., Alekseyenko, A.A., Park, P.J., and Kuroda, M.I. (2014). Transcriptional control of a whole chromosome: emerging models for dosage compensation. Nature structural & molecular biology *21*, 118-125.

Ferrari, F., Plachetka, A., Alekseyenko, A.A., Jung, Y.L., Ozsolak, F., Kharchenko, P.V., Park, P.J., and Kuroda, M.I. (2013). "Jump start and gain" model for dosage compensation in Drosophila based on direct sequencing of nascent transcripts. Cell reports *5*, 629-636.

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

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J.*, et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212-224.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Molecular and cellular biology *16*, 2802-2813.

Filippova, G.N., Qi, C.F., Ulmer, J.E., Moore, J.M., Ward, M.D., Hu, Y.J., Loukinov, D.I., Pugacheva, E.M., Klenova, E.M., Grundy, P.E.*, et al.* (2002). Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter tts DNA-binding specificity. Cancer Res *62*, 48-52.

Flemming, W. (1882). Zellsubstanz, Kern und Zellteilung (Leipzig: F.C.W. Vogel).

Franke, A., and Baker, B.S. (1999). The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. Molecular cell *4*, 117-122.

Fukunaga, A., Tanaka, A., and Oishi, K. (1975). Maleless, a recessive autosomal mutant of Drosophila melanogaster that specifically kills male zygotes. Genetics *81*, 135-141.

Fulton, D.L., Sundararajan, S., Badis, G., Hughes, T.R., Wasserman, W.W., Roach, J.C., and Sladek, R. (2009). TFCat: the curated catalog of mouse and human transcription factors. Genome biology *10*, R29.

Furuhashi, H., Nakajima, M., and Hirose, S. (2006). DNA supercoiling factor contributes to dosage compensation in Drosophila. Development *133*, 4475-4483.

Gebauer, F., Corona, D.F., Preiss, T., Becker, P.B., and Hentze, M.W. (1999). Translational control of dosage compensation in Drosophila by Sex-lethal: cooperative silencing via the 5' and 3' UTRs of msl-2 mRNA is independent of the poly(A) tail. EMBO J *18*, 6146-6154.

Gebauer, F., Grskovic, M., and Hentze, M.W. (2003). Drosophila sex-lethal inhibits the stable association of the 40S ribosomal subunit with msl-2 mRNA. Molecular cell *11*, 1397-1404.

Gelato, K.A., and Fischle, W. (2008). Role of histone modifications in defining chromatin structure and function. Biol Chem *389*, 353-363.

Gelbart, M.E., Larschan, E., Peng, S., Park, P.J., and Kuroda, M.I. (2009). Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. Nature structural & molecular biology *16*, 825-832.

Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A.*, et al.* (2015). A global reference for human genetic variation. Nature *526*, 68-74.

Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature *512*, 96-100.

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C.*, et al.* (2018). A pathway for mitotic chromosome formation. Science *359*, eaao6135.

Giet, R., and Glover, D.M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. The Journal of cell biology *152*, 669-682.

Gilfillan, G.D., Konig, C., Dahlsveen, I.K., Prakoura, N., Straub, T., Lamm, R., Fauth, T., and Becker, P.B. (2007). Cumulative contributions of weak DNA determinants to targeting the Drosophila dosage compensation complex. Nucleic acids research *35*, 3561-3572.

Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. (2006). Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes & development *20*, 858-870.

Girton, J., Wang, C., Johansen, J., and Johansen, K.M. (2013). The effect of JIL-1 on position-effect variegation is proportional to the total amount of heterochromatin in the genome. Fly (Austin) *7*, 129-133.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature *386*, 44-51.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nature reviews Genetics *17*, 333-351.

Gopalakrishnan, R., Marr, S.K., Kingston, R.E., and Winston, F. (2019). A conserved genetic interaction between Spt6 and Set2 regulates H3K36 methylation. Nucleic acids research, gkz119.

Gossett, A.J., and Lieb, J.D. (2008). DNA Immunoprecipitation (DIP) for the Determination of DNA-Binding Specificity. CSH Protoc *2008*, pdb. prot4972.

Grimaud, C., and Becker, P.B. (2009). The dosage compensation complex shapes the conformation of the X chromosome in Drosophila. Genes & development *23*, 2490-2495.

Grskovic, M., Hentze, M.W., and Gebauer, F. (2003). A co-repressor assembly nucleated by Sex-lethal in the 3'UTR mediates translational control of Drosophila msl-2 mRNA. EMBO J *22*, 5571-5581.

Gu, L., and Walters, J.R. (2017). Evolution of Sex Chromosome Dosage Compensation in Animals: A Beautiful Theory, Undermined by Facts and Bedeviled by Details. Genome Biol Evol *9*, 2461-2476.

Gu, W., Szauter, P., and Lucchesi, J.C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster. Dev Genet *22*, 56-64.

Guo, R., Zheng, L., Park, J.W., Lv, R., Chen, H., Jiao, F., Xu, W., Mu, S., Wen, H., Qiu, J.*, et al.* (2014). BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. Molecular cell *56*, 298-310.

Gurard-Levin, Z.A., Quivy, J.-P., and Almouzni, G. (2014). Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. Annu Rev Biochem *83*, 487-517.

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

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: Contribution to chromatin structure and function. Gene *431*, 1-12.

Henikoff, S., and Smith, M.M. (2015). Histone variants and epigenetics. Cold Spring Harb Perspect Biol *7*, a019364.

Hennig, J., Militti, C., Popowicz, G.M., Wang, I., Sonntag, M., Geerlof, A., Gabel, F., Gebauer, F., and Sattler, M. (2014). Structural basis for the assembly of the Sxl-Unr translation regulatory complex. Nature *515*, 287-290.

Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem Biophys Res Commun *52*, 504-510.

Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J.C. (1997). mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila. EMBO J *16*, 2054-2060.

Holowatyj, A., Yang, Z.Q., and Pile, L.A. (2015). Histone lysine demethylases in Drosophila melanogaster. Fly (Austin) *9*, 36-44.

Huang, C., Yang, F., Zhang, Z., Zhang, J., Cai, G., Li, L., Zheng, Y., Chen, S., Xi, R., and Zhu, B. (2017). Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in Drosophila. Nat Commun *8*, 1649.

Huang, C., and Zhu, B. (2018). Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. Biophysics reports *4*, 170-177.

Ilik, I.A., Maticzka, D., Georgiev, P., Gutierrez, N.M., Backofen, R., and Akhtar, A. (2017). A mutually exclusive stem-loop arrangement in roX2 RNA is essential for X-chromosome regulation in Drosophila. Genes & development *31*, 1973-1987.

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

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931-945.

International Human Genome Sequencing Consortium, Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Inukai, S., Kock, K.H., and Bulyk, M.L. (2017). Transcription factor-DNA binding: beyond binding site motifs. Current opinion in genetics & development *43*, 110-119.

Iuchi, S. (2001). Three classes of C2H2 zinc finger proteins. Cell Mol Life Sci *58*, 625-635.

Izzo, A., Regnard, C., Morales, V., Kremmer, E., and Becker, P.B. (2008). Structure-function analysis of the RNA helicase maleless. Nucleic acids research *36*, 950-962.

Izzo, A., and Schneider, R. (2016). The role of linker histone H1 modifications in the regulation of gene expression and chromatin dynamics. Biochimica et biophysica acta *1859*, 486-495.

Jelinic, P., Pellegrino, J., and David, G. (2011). A novel mammalian complex containing Sin3B mitigates histone acetylation and RNA polymerase II progression within transcribed loci. Molecular and cellular biology *31*, 54-62.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jin, Y., Wang, Y., Johansen, J., and Johansen, K.M. (2000). JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. The Journal of cell biology *149*, 1005-1010.

Jin, Y., Wang, Y., Walker, D.L., Dong, H., Conley, C., Johansen, J., and Johansen, K.M. (1999). JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in Drosophila. Molecular cell *4*, 129-135.

Johansson, A.M., Allgardsson, A., Stenberg, P., and Larsson, J. (2011). msl2 mRNA is bound by free nuclear MSL complex in Drosophila melanogaster. Nucleic acids research *39*, 6428-6439.

Jones, R.S., and Gelbart, W.M. (1993). The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Molecular and cellular biology *13*, 6357-6366.

Joshi, A.A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Molecular cell *20*, 971-978.

Joshi, S.S., and Meller, V.H. (2017). Satellite Repeats Identify X Chromatin for Dosage Compensation in Drosophila melanogaster Males. Curr Biol *27*, 1393-1402.e1392.

Kadlec, J., Hallacli, E., Lipp, M., Holz, H., Sanchez-Weatherby, J., Cusack, S., and Akhtar, A. (2011). Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. Nature structural & molecular biology *18*, 142-149.

Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H.G., Stuckenholz, C., Kelley, R.L., Becker, P.B., and Kuroda, M.I. (2001). Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. EMBO J *20*, 2236-2245.

Kaye, E.G., Booker, M., Kurland, J.V., Conicella, A.E., Fawzi, N.L., Bulyk, M.L., Tolstorukov, M.Y., and Larschan, E. (2018). Differential Occupancy of Two GA-Binding Proteins Promotes Targeting of the Drosophila Dosage Compensation Complex to the Male X Chromosome. Cell reports *22*, 3227-3239.

Kaye, E.G., Kurbidaeva, A., Wolle, D., Aoki, T., Schedl, P., and Larschan, E. (2017). Drosophila Dosage Compensation Loci Associate with a Boundary-Forming Insulator Complex. Molecular and cellular biology *37*, e00253.

Keller, C.I., and Akhtar, A. (2015). The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond. Current opinion in genetics & development *31*, 1-11.

Kelley, R.L., Meller, V.H., Gordadze, P.R., Roman, G., Davis, R.L., and Kuroda, M.I. (1999). Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. Cell *98*, 513-522.

Kelley, R.L., Solovyeva, I., Lyman, L.M., Richman, R., Solovyev, V., and Kuroda, M.I. (1995). Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in Drosophila. Cell *81*, 867-877.

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

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T.*, et al.* (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480-485.

Khorasanizadeh, S. (2004). The Nucleosome. Cell *116*, 259-272.

Kim, D., Blus, B.J., Chandra, V., Huang, P., Rastinejad, F., and Khorasanizadeh, S. (2010). Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. Nature structural & molecular biology *17*, 1027-1029.

Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila. Cell *133*, 813-828.

Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L., and Strahl, B.D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Molecular and cellular biology *25*, 3305-3316.

Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes & development *14*, 2452-2460.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Kossel, A. (1884). Ueber einen peptonartigen Bestandteil des Zellkerns. Hoppe-Seyler's Z Physiol Chem *8*, 511-515.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kruse, J.P., and Gu, W. (2009). MSL2 promotes Mdm2-independent cytoplasmic localization of p53. The Journal of biological chemistry *284*, 3250-3263.

Kumar, G.S., Chang, W., Xie, T., Patel, A., Zhang, Y., Wang, G.G., David, G., and Radhakrishnan, I. (2012). Sequence requirements for combinatorial recognition of histone H3 by the MRG15 and Pf1 subunits of the Rpd3S/Sin3S corepressor complex. Journal of molecular biology *422*, 519-531.

Kuroda, M.I., Hilfiker, A., and Lucchesi, J.C. (2016). Dosage Compensation in Drosophila-a Model for the Coordinate Regulation of Transcription. Genetics *204*, 435-450.

Kusch, T., Florens, L., Macdonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., 3rd, Abmayr, S.M., Washburn, M.P., and Workman, J.L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science *306*, 2084-2087.

Kuzu, G., Kaye, E.G., Chery, J., Siggers, T., Yang, L., Dobson, J.R., Boor, S., Bliss, J., Liu, W., Jogl, G.*, et al.* (2016). Expansion of GA Dinucleotide Repeats Increases the Density of CLAMP Binding Sites on the X-Chromosome to Promote Drosophila Dosage Compensation. PLoS genetics *12*, e1006120.

Laguri, C., Duband-Goulet, I., Friedrich, N., Axt, M., Belin, P., Callebaut, I., Gilquin, B., Zinn-Justin, S., and Couprie, J. (2008). Human mismatch repair protein MSH6 contains a PWWP domain that targets double stranded DNA. Biochemistry *47*, 6199-6207.

Laity, J.H., Dyson, H.J., and Wright, P.E. (2000). DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. Journal of molecular biology *295*, 719-727.

Lam, K.C., Muhlpfordt, F., Vaquerizas, J.M., Raja, S.J., Holz, H., Luscombe, N.M., Manke, T., and Akhtar, A. (2012). The NSL complex regulates housekeeping genes in Drosophila. PLoS genetics *8*, e1002736.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650-665.

Larschan, E., Alekseyenko, A.A., Gortchakov, A.A., Peng, S., Li, B., Yang, P., Workman, J.L., Park, P.J., and Kuroda, M.I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. Molecular cell *28*, 121-133.

Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature *471*, 115-118.

Larschan, E., Soruco, M.M., Lee, O.K., Peng, S., Bishop, E., Chery, J., Goebel, K., Feng, J., Park, P.J., and Kuroda, M.I. (2012). Identification of chromatin-associated regulators of MSL complex targeting in Drosophila dosage compensation. PLoS genetics *8*, e1002830.

Lee, C.G., Chang, K.A., Kuroda, M.I., and Hurwitz, J. (1997). The NTPase/helicase activities of Drosophila maleless, an essential factor in dosage compensation. EMBO J *16*, 2671-2681.

Lee, C.H., Wu, J., and Li, B. (2013). Chromatin remodelers fine-tune H3K36me-directed deacetylation of neighbor nucleosomes by Rpd3S. Molecular cell *52*, 255-263.

Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain. Science *245*, 635-637.

Legube, G., McWeeney, S.K., Lercher, M.J., and Akhtar, A. (2006). X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in Drosophila. Genes & development *20*, 871-883.

Lerach, S., Zhang, W., Bao, X., Deng, H., Girton, J., Johansen, J., and Johansen, K.M. (2006). Loss-of-function alleles of the JIL-1 kinase are strong suppressors of position effect variegation of the wm4 allele in Drosophila. Genetics *173*, 2403-2406.

Levy, S.E., and Myers, R.M. (2016). Advancements in Next-Generation Sequencing. Annual review of genomics and human genetics *17*, 95-115.

Li, B., Howe, L., Anderson, S., Yates, J.R., 3rd, and Workman, J.L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. The Journal of biological chemistry *278*, 8897-8903.

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

Li, F., Parry, D.A., and Scott, M.J. (2005). The amino-terminal region of Drosophila MSL1 contains basic, glycine-rich, and leucine zipper-like motifs that promote X chromosome binding, self-association, and MSL2 binding, respectively. Molecular and cellular biology *25*, 8913-8924.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current opinion in genetics & development *21*, 175-186.

Li, J., Duns, G., Westers, H., Sijmons, R., van den Berg, A., and Kok, K. (2016). SETD2: an epigenetic modifier with tumor suppressor functionality. Oncotarget *7*, 50719-50734.

Li, Y., Cai, W., Wang, C., Yao, C., Bao, X., Deng, H., Girton, J., Johansen, J., and Johansen, K.M. (2013b). Domain requirements of the JIL-1 tandem kinase for histone H3 serine 10 phosphorylation and chromatin remodeling in vivo. The Journal of biological chemistry *288*, 19441-19449.

Lin, L.J., Minard, L.V., Johnston, G.C., Singer, R.A., and Schultz, M.C. (2010). Asf1 can promote trimethylation of H3 K36 by Set2. Molecular and cellular biology *30*, 1116-1129.

Lindehell, H., Kim, M., and 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.

Liu, X., Noll, D.M., Lieb, J.D., and Clarke, N.D. (2005). DIP-chip: rapid and accurate determination of DNA-binding specificity. Genome research *15*, 421-427.

Liu, Y., Lu, C., Yang, Y., Fan, Y., Yang, R., Liu, C.F., Korolev, N., and Nordenskiold, L. (2011). Influence of histone tails and H4 tail acetylations on nucleosome-nucleosome interactions. Journal of molecular biology *414*, 749-764.

Lucchesi, J.C. (2018). Transcriptional modulation of entire chromosomes: dosage compensation. J Genet *97*, 357-364.

Lucchesi, J.C., and Kuroda, M.I. (2015). Dosage compensation in Drosophila. Cold Spring Harb Perspect Biol *7*, a019398.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol *13*, 436-447.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251-260.

Luscombe, N.M., Austin, S.E., Berman, H.M., and Thornton, J.M. (2000). An overview of the structures of protein-DNA complexes. Genome biology *1*, REVIEWS001.

Lv, M., Yao, Y., Li, F., Xu, L., Yang, L., Gong, Q., Xu, Y.Z., Shi, Y., Fan, Y.J., and Tang, Y. (2019). Structural insights reveal the specific recognition of roX RNA by the dsRNA-binding domains of the RNA helicase MLE and its indispensable role in dosage compensation in Drosophila. Nucleic acids research, gky1308.

Lyman, L.M., Copps, K., Rastelli, L., Kelley, R.L., and Kuroda, M.I. (1997). Drosophila male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association. Genetics *147*, 1743-1753.

Maenner, S., Muller, M., and Becker, P.B. (2012). Roles of long, non-coding RNA in chromosome-wide transcription regulation: lessons from two dosage compensation systems. Biochimie *94*, 1490-1498.

Maenner, S., Muller, M., Frohlich, J., Langer, D., and Becker, P.B. (2013). ATP-dependent roX RNA remodeling by the helicase maleless enables specific association of MSL proteins. Molecular cell *51*, 174-184.

Maeshima, K., Rogge, R., Tamura, S., Joti, Y., Hikima, T., Szerlong, H., Krause, C., Herman, J., Seidel, E., DeLuca, J.*, et al.* (2016). Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers. EMBO J *35*, 1115-1132.

Maltby, V.E., Martin, B.J., Schulze, J.M., Johnson, I., Hentrich, T., Sharma, A., Kobor, M.S., and Howe, L. (2012). Histone H3 lysine 36 methylation targets the Isw1b remodeling complex to chromatin. Molecular and cellular biology *32*, 3479-3485.

Marceau, A.H., Felthousen, J.G., Goetsch, P.D., Iness, A.N., Lee, H.W., Tripathi, S.M., Strome, S., Litovchick, L., and Rubin, S.M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun *7*, 12301.

Marin, I. (2003). Evolution of chromatin-remodeling complexes: comparative genomics reveals the ancient origin of "novel" compensasome genes. J Mol Evol *56*, 527-539.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nature Structural &Amp; Molecular Biology *17*, 1272-1278.

Maze, I., Noh, K.-M., Soshnev, A.A., and Allis, C.D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nature reviews Genetics *15*, 259-271.

McDaniel, S.L., Strahl, B.D.J.C., and Sciences, M.L. (2017). Shaping the cellular landscape with Set2/SETD2 methylation. *74*, 3317-3334.

McDowell, K.A., Hilfiker, A., and Lucchesi, J.C. (1996). Dosage compensation in Drosophila: the X chromosome binding of MSL-1 and MSL-2 in female embryos is prevented by the early expression of the Sxl gene. Mech Dev *57*, 113-119.

McElroy, K.A., Kang, H., and Kuroda, M.I. (2014). Are we there yet? Initial targeting of the Male-Specific Lethal and Polycomb group chromatin complexes in Drosophila. Open biology *4*, 140006.

Melcher, M., Schmid, M., Aagaard, L., Selenko, P., Laible, G., and Jenuwein, T. (2000). Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. Molecular and cellular biology *20*, 3728-3741.

Meller, V.H. (2003). Initiation of dosage compensation in Drosophila embryos depends on expression of the roX RNAs. Mech Dev *120*, 759-767.

Meller, V.H., Gordadze, P.R., Park, Y., Chu, X., Stuckenholz, C., Kelley, R.L., and Kuroda, M.I. (2000). Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. Curr Biol *10*, 136-143.

Meller, V.H., and Rattner, B.P. (2002). The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J *21*, 1084-1091.

Meller, V.H., Wu, K.H., Roman, G., Kuroda, M.I., and Davis, R.L. (1997). roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. Cell *88*, 445-457.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J.*, et al.* (2006). Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. Molecular cell *21*, 811-823.

Menon, D.U., Coarfa, C., Xiao, W., Gunaratne, P.H., and Meller, V.H. (2014). siRNAs from an X-linked satellite repeat promote X-chromosome recognition in Drosophila melanogaster. Proc Natl Acad Sci U S A *111*, 16460-16465.

Menon, D.U., and Meller, V.H. (2015). Identification of the Drosophila X chromosome: The long and short of it. RNA Biol *12*, 1088-1093.

Miescher, F. (1871). Ueber die chemische Zusammensetzung der Eiterzellen. Hoppe-Seyler (Hg), Med-Chem Unters *4*, 441-460.

Migeon, B.R. (2017). Choosing the Active X: The Human Version of X Inactivation. Trends in genetics : TIG *33*, 899-909.

Militti, C., Maenner, S., Becker, P.B., and Gebauer, F. (2014). UNR facilitates the interaction of MLE with the lncRNA roX2 during Drosophila dosage compensation. Nat Commun *5*, 4762.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I.*, et al.* (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic acids research *47*, D351-D360.

Moore, S.A., Ferhatoglu, Y., Jia, Y., Al-Jiab, R.A., and Scott, M.J. (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. The Journal of biological chemistry *285*, 40879-40890.

Morales, V., Regnard, C., Izzo, A., Vetter, I., and Becker, P.B. (2005). The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex. Molecular and cellular biology *25*, 5947-5954.

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

Morgunova, E., and Taipale, J. (2017). Structural perspective of cooperative transcription factor binding. Current opinion in structural biology *47*, 1-8.

Morra, R., Smith, E.R., Yokoyama, R., and Lucchesi, J.C. (2008). The MLE subunit of the Drosophila MSL complex uses its ATPase activity for dosage compensation and its helicase activity for targeting. Molecular and cellular biology *28*, 958-966.

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

Mukherjee, A.S., and Beermann, W. (1965). Synthesis of ribonucleic acid by the X-chromosomes of Drosophila melanogaster and the problem of dosage compensation. Nature *207*, 785-786.

Muller, H.J. (1914). A gene for the fourth chromosome in Drosophila. J Exp Zool *17*, 325-336.

Muller, H.J. (1932). Further studies on the nature and causes of gene mutations. In Proceedings of the Sixth International Congress of Genetics (Ithaca, NY), pp. 213-255.

Muller, H.J. (1950). Evidence of the precision of genetic adaptation. Harvey Lecture Series XLIII, 1947–1948 *1*, 165-229.

Musselman, C.A., Lalonde, M.E., Cote, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nature structural & molecular biology *19*, 1218-1227.

Nakahashi, H., Kieffer Kwon, K.R., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S., Yamane, A.*, et al.* (2013). A genome-wide map of CTCF multivalency redefines the CTCF code. Cell reports *3*, 1678-1689.

Nelson, C.J., Santos-Rosa, H., and Kouzarides, T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell *126*, 905-916.

Nielsen, P.R., Nietlispach, D., Buscaino, A., Warner, R.J., Akhtar, A., Murzin, A.G., Murzina, N.V., and Laue, E.D. (2005). Structure of the chromo barrel domain from the MOF acetyltransferase. The Journal of biological chemistry *280*, 32326-32331.

Nolte, R.T., Conlin, R.M., Harrison, S.C., and Brown, R.S. (1998). Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. Proc Natl Acad Sci U S A *95*, 2938-2943.

Ohno, M., Ando, T., Priest, D.G., Kumar, V., Yoshida, Y., and Taniguchi, Y. (2019). Sub-nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs. Cell *176*, 520-534 e525.

Ohno, M., Priest, D.G., and Taniguchi, Y. (2018). Nucleosome-level 3D organization of the genome. Biochem Soc Trans *46*, 491-501.

Ohno, S. (1967). Sex chromosomes and sex-linked genes (Berlin: Springer-Verlag).

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nat Rev Mol Cell Biol *4*, 809-814.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281-300.

Pan, Y., Tsai, C.J., Ma, B., and Nussinov, R. (2010). Mechanisms of transcription factor selectivity. Trends in genetics : TIG *26*, 75-83.

Park, S.W., Kang, Y., Sypula, J.G., Choi, J., Oh, H., and Park, Y. (2007). An evolutionarily conserved domain of roX2 RNA is sufficient for induction of H4-Lys16 acetylation on the Drosophila X chromosome. Genetics *177*, 1429-1437.

Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. Annu Rev Biochem *82*, 81-118.

Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809-817.

Pavletich, N.P., and Pabo, C.O. (1993). Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. Science *261*, 1701-1707.

Pepenella, S., Murphy, K.J., and Hayes, J.J. (2014). Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. Chromosoma *123*, 3-13.

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

Pinheiro, I., and Heard, E. (2017). X chromosome inactivation: new players in the initiation of gene silencing. F1000Res *6*.

Prabhakaran, M., and Kelley, R.L. (2012). Mutations in the transcription elongation factor SPT5 disrupt a reporter for dosage compensation in Drosophila. PLoS genetics *8*, e1003073.

Prabu, J.R., Muller, M., Thomae, A.W., Schussler, S., Bonneau, F., Becker, P.B., and Conti, E. (2015). Structure of the RNA Helicase MLE Reveals the Molecular Mechanisms for Uridine Specificity and RNA-ATP Coupling. Molecular cell *60*, 487-499.

Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. PLoS genetics *8*, e1002717.

Prestel, M., Feller, C., Straub, T., Mitlohner, H., and Becker, P.B. (2010). The activation potential of MOF is constrained for dosage compensation. Molecular cell *38*, 815-826.

Qin, S., and Min, J. (2014). Structure and function of the nucleosome-binding PWWP domain. Trends Biochem Sci *39*, 536-547.

Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat Struct Biol *9*, 217-224.

Quinn, J.J., Ilik, I.A., Qu, K., Georgiev, P., Chu, C., Akhtar, A., and Chang, H.Y. (2014). Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. Nat Biotechnol *32*, 933-940.

Quinn, J.J., Zhang, Q.C., Georgiev, P., Ilik, I.A., Akhtar, A., and Chang, H.Y. (2016). Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes & development *30*, 191-207.

Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: How is it established, and why does it matter? Dev Biol *339*, 258-266.

Raja, S.J., Charapitsa, I., Conrad, T., Vaquerizas, J.M., Gebhardt, P., Holz, H., Kadlec, J., Fraterman, S., Luscombe, N.M., and Akhtar, A. (2010). The nonspecific lethal complex is a transcriptional regulator in Drosophila. Molecular cell *38*, 827-841.

Ramirez, F., Lingg, T., Toscano, S., Lam, K.C., Georgiev, P., Chung, H.R., Lajoie, B.R., de Wit, E., Zhan, Y., de Laat, W.*, et al.* (2015). High-Affinity Sites Form an Interaction Network to Facilitate Spreading of the MSL Complex across the X Chromosome in Drosophila. Molecular cell *60*, 146-162.

Razin, S.V., Borunova, V.V., Maksimenko, O.G., and Kantidze, O.L. (2012). Cys2His2 zinc finger protein family: classification, functions, and major members. Biochemistry Biokhimiia *77*, 217-226.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M.*, et al.* (2017). The Human Cell Atlas. Elife *6*, e27041.

Regnard, C., Straub, T., Mitterweger, A., Dahlsveen, I.K., Fabian, V., and Becker, P.B. (2011). Global analysis of the relationship between JIL-1 kinase and transcription. PLoS genetics *7*, e1001327.

Renda, M., Baglivo, I., Burgess-Beusse, B., Esposito, S., Fattorusso, R., Felsenfeld, G., and Pedone, P.V. (2007). Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. The Journal of biological chemistry *282*, 33336-33345.

Richter, L., Bone, J.R., and Kuroda, M.I. (1996). RNA-dependent association of the Drosophila maleless protein with the male X chromosome. Genes Cells *1*, 325-336.

Rieder, L.E., Koreski, K.P., Boltz, K.A., Kuzu, G., Urban, J.A., Bowman, S.K., Zeidman, A., Jordan, W.T., 3rd, Tolstorukov, M.Y., Marzluff, W.F.*, et al.* (2017). Histone locus regulation by the Drosophila dosage compensation adaptor protein CLAMP. Genes & development *31*, 1494-1508.

Rivera, C.M., and Ren, B. (2013). Mapping human epigenomes. Cell *155*, 39-55.

Robinson, P.J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R.G., and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. Journal of molecular biology *381*, 816-825.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the “30-nm” chromatin fiber: Evidence for a compact, interdigitated structure. Proc Natl Acad Sci U S A *103*, 6506-6511.

Robinson, P.J.J., and Rhodes, D. (2006). Structure of the ‘30nm’ chromatin fibre: A key role for the linker histone. Current opinion in structural biology *16*, 336-343.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234*, 364-368.

Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., and Mann, R.S. (2010). Origins of specificity in protein-DNA recognition. Annu Rev Biochem *79*, 233-269.

Rona, G.B., Eleutherio, E.C.A., and Pinheiro, A.S. (2016). PWWP domains and their modes of sensing DNA and histone methylated lysines. Biophys Rev *8*, 63-74.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochimica et biophysica acta *1839*, 627-643.

Sahakyan, A., Plath, K., and Rougeulle, C. (2017). Regulation of X-chromosome dosage compensation in human: mechanisms and model systems. Philos Trans R Soc Lond B Biol Sci *372*, 20160363.

Samata, M., and Akhtar, A. (2018). Dosage Compensation of the X Chromosome: A Complex Epigenetic Assignment Involving Chromatin Regulators and Long Noncoding RNAs. Annu Rev Biochem *87*, 323-350.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138-141.

Schmahling, S., Meiler, A., Lee, Y., Mohammed, A., Finkl, K., Tauscher, K., Israel, L., Wirth, M., Philippou-Massier, J., Blum, H.*, et al.* (2018). Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. Development *145*, dev163808.

Schmit, F., Cremer, S., and Gaubatz, S. (2009). LIN54 is an essential core subunit of the DREAM/LINC complex that binds to the cdc2 promoter in a sequence-specific manner. FEBS J *276*, 5703-5716.

Schmit, F., Korenjak, M., Mannefeld, M., Schmitt, K., Franke, C., von Eyss, B., Gagrica, S., Hanel, F., Brehm, A., and Gaubatz, S. (2007). LINC, a human complex that is related to pRB-containing complexes in invertebrates regulates the expression of G2/M genes. Cell Cycle *6*, 1903-1913.

Schmitges, F.W., Prusty, A.B., Faty, M., Stutzer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S.*, et al.* (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Molecular cell *42*, 330-341.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J *21*, 1121-1131.

Schunter, S., Villa, R., Flynn, V., Heidelberger, J.B., Classen, A.K., Beli, P., and Becker, P.B. (2017). Ubiquitylation of the acetyltransferase MOF in Drosophila melanogaster. PloS one *12*, e0177408.

Scott, M.J., Pan, L.L., Cleland, S.B., Knox, A.L., and Heinrich, J. (2000). MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. EMBO J *19*, 144-155.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Sharma, S., Cermakova, K., De Rijck, J., Demeulemeester, J., Fabry, M., El Ashkar, S., Van Belle, S., Lepsik, M., Tesina, P., Duchoslav, V.*, et al.* (2018). Affinity switching of the LEDGF/p75 IBD interactome is governed by kinase-dependent phosphorylation. Proc Natl Acad Sci U S A *115*, E7053-E7062.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. Science *311*, 844-847.

Siggers, T., and Gordan, R. (2014). Protein-DNA binding: complexities and multi-protein codes. Nucleic acids research *42*, 2099-2111.

Silva-Sousa, R., and Casacuberta, E. (2013). The JIL-1 kinase affects telomere expression in the different telomere domains of Drosophila. PloS one *8*, e81543.

Silva-Sousa, R., Lopez-Panades, E., Pineyro, D., and Casacuberta, E. (2012). The chromosomal proteins JIL-1 and Z4/Putzig regulate the telomeric chromatin in Drosophila melanogaster. PLoS genetics *8*, e1003153.

Slattery, M., Zhou, T., Yang, L., Dantas Machado, A.C., Gordan, R., and Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. Trends Biochem Sci *39*, 381-399.

Smith, E.R., Allis, C.D., and Lucchesi, J.C. (2001). Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. The Journal of biological chemistry *276*, 31483-31486.

Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Cote, J., and Lucchesi, J.C. (2005). A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Molecular and cellular biology *25*, 9175-9188.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The Drosophila MSL Complex Acetylates Histone H4 at Lysine 16, a Chromatin Modification Linked to Dosage Compensation. Molecular and cellular biology *20*, 312-318.

Smolle, M., Venkatesh, S., Gogol, M.M., Li, H., Zhang, Y., Florens, L., Washburn, M.P., and Workman, J.L. (2012). Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. Nature structural & molecular biology *19*, 884-892.

Song, J.Y., Leung, T., Ehler, L.K., Wang, C., and Liu, Z. (2000). Regulation of meristem organization and cell division by TSO1, an Arabidopsis gene with cysteine-rich repeats. Development *127*, 2207-2217.

Soruco, M.M., Chery, J., Bishop, E.P., Siggers, T., Tolstorukov, M.Y., Leydon, A.R., Sugden, A.U., Goebel, K., Feng, J., Xia, P.*, et al.* (2013). The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. Genes & development *27*, 1551-1556.

Speranzini, V., Pilotto, S., Sixma, T.K., and Mattevi, A. (2016). Touch, act and go: landing and operating on nucleosomes. EMBO J *35*, 376-388.

Spierer, A., Begeot, F., Spierer, P., and Delattre, M. (2008). SU(VAR)3-7 links heterochromatin and dosage compensation in Drosophila. PLoS genetics *4*, e1000066.

Spierer, A., Seum, C., Delattre, M., and Spierer, P. (2005). Loss of the modifiers of variegation Su(var)3-7 or HP1 impacts male X polytene chromosome morphology and dosage compensation. J Cell Sci *118*, 5047-5057.

Stabell, M., Larsson, J., Aalen, R.B., and Lambertsson, A. (2007). Drosophila dSet2 functions in H3-K36 methylation and is required for development. Biochem Biophys Res Commun *359*, 784-789.

Stamatoyannopoulos, J.A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D.M., Groudine, M., Bender, M., Kaul, R., Canfield, T.*, et al.* (2012). An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome biology *13*, 418.

Stec, I., Wright, T.J., van Ommen, G.J., de Boer, P.A., van Haeringen, A., Moorman, A.F., Altherr, M.R., and den Dunnen, J.T. (1998). WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphy gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. Hum Mol Genet *7*, 1071-1082.

Straub, T., and Becker, P.B. (2007). Dosage compensation: the beginning and end of generalization. Nature reviews Genetics *8*, 47-57.

Straub, T., and Becker, P.B. (2011). Transcription modulation chromosome-wide: universal features and principles of dosage compensation in worms and flies. Current opinion in genetics & development *21*, 147-153.

Straub, T., Grimaud, C., Gilfillan, G.D., Mitterweger, A., and Becker, P.B. (2008). The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. PLoS genetics *4*, e1000302.

Straub, T., Zabel, A., Gilfillan, G.D., Feller, C., and Becker, P.B. (2013). Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. Genome research *23*, 473-485.

Strome, S., Kelly, W.G., Ercan, S., and Lieb, J.D. (2014). Regulation of the X chromosomes in Caenorhabditis elegans. Cold Spring Harb Perspect Biol *6*, a018366.

Stuckenholz, C., Meller, V.H., and Kuroda, M.I. (2003). Functional redundancy within roX1, a noncoding RNA involved in dosage compensation in Drosophila melanogaster. Genetics *164*, 1003-1014.

Sudmant, P.H., Rausch, T., Gardner, E.J., Handsaker, R.E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Fritz, M.H.*, et al.* (2015). An integrated map of structural variation in 2,504 human genomes. Nature *526*, 75-81.

Sue, S.C., Chen, J.Y., Lee, S.C., Wu, W.G., and Huang, T.H. (2004). Solution structure and heparin interaction of human hepatoma-derived growth factor. Journal of molecular biology *343*, 1365-1377.

Sun, B., Hong, J., Zhang, P., Dong, X., Shen, X., Lin, D., and Ding, J. (2008). Molecular basis of the interaction of Saccharomyces cerevisiae Eaf3 chromo domain with methylated H3K36. The Journal of biological chemistry *283*, 36504-36512.

Sural, T.H., Peng, S., Li, B., Workman, J.L., Park, P.J., and Kuroda, M.I. (2008). The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. Nature structural & molecular biology *15*, 1318-1325.

Sutou, S., Miwa, K., Matsuura, T., Kawasaki, Y., Ohinata, Y., and Mitsui, Y. (2003). Native tesmin is a 60-kilodalton protein that undergoes dynamic changes in its localization during spermatogenesis in mice. Biol Reprod *68*, 1861-1869.

Suzuki, M., Yamada, T., Kihara-Negishi, F., Sakurai, T., Hara, E., Tenen, D.G., Hozumi, N., and Oikawa, T. (2006). Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. Oncogene *25*, 2477-2488.

Szostak, E., Garcia-Beyaert, M., Guitart, T., Graindorge, A., Coll, O., and Gebauer, F. (2018). Hrp48 and eIF3d contribute to msl-2 mRNA translational repression. Nucleic acids research *46*, 4099-4113.

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

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature structural & molecular biology *14*, 1025-1040.

The Encode Project Consortium, Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J.*, et al.* (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Tremethick, D.J. (2007). Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber. Cell *128*, 651-654.

Turner, B.M., Birley, A.J., and 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.

Uchida, S., Uenoyama, T., and Oishi, K. (1981). Studies on the sex-specific lethals of Drosophila melanogaster. III. A third chromosome male-specific lethal mutant. Jpn J Genet *56*, 523–527.

Ulianov, S.V., Khrameeva, E.E., Gavrilov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M.V., Chertovich, A.*, et al.* (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. Genome research *26*, 70-84.

Urban, J., Kuzu, G., Bowman, S., Scruggs, B., Henriques, T., Kingston, R., Adelman, K., Tolstorukov, M., and Larschan, E. (2017a). Enhanced chromatin accessibility of the dosage

compensated Drosophila male X-chromosome requires the CLAMP zinc finger protein. PloS one *12*, e0186855.

Urban, J.A., Doherty, C.A., Jordan, W.T., 3rd, Bliss, J.E., Feng, J., Soruco, M.M., Rieder, L.E., Tsiarli, M.A., and Larschan, E.N. (2017b). The essential Drosophila CLAMP protein differentially regulates non-coding roX RNAs in male and females. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *25*, 101-113.

Urban, J.A., Urban, J.M., Kuzu, G., and Larschan, E.N. (2017c). The Drosophila CLAMP protein associates with diverse proteins on chromatin. PloS one *12*, e0189772.

van Nuland, R., van Schaik, F.M., Simonis, M., van Heesch, S., Cuppen, E., Boelens, R., Timmers, H.M., and van Ingen, H. (2013). Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin *6*, 12.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution. Nature reviews Genetics *10*, 252-263.

Vaquerizas, J.M., Suyama, R., Kind, J., Miura, K., Luscombe, N.M., and Akhtar, A. (2010). Nuclear pore proteins nup153 and megator define transcriptionally active regions in the Drosophila genome. PLoS genetics *6*, e1000846.

Venkatesh, S., and Workman, J.L. (2013). Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. *2*, 685-700.

Venkatesh, S., and Workman, J.L. (2015). Histone exchange, chromatin structure and the regulation of transcription. Nat Rev Mol Cell Biol *16*, 178-189.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A*., et al.* (2001). The Sequence of the Human Genome. Science *291*, 1304-1351.

Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G*., et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell *142*, 967-980.

Vezzoli, A., Bonadies, N., Allen, M.D., Freund, S.M., Santiveri, C.M., Kvinlaug, B.T., Huntly, B.J., Gottgens, B., and Bycroft, M. (2010). Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. Nature structural & molecular biology *17*, 617-619.

Villa, R., Forne, I., Muller, M., Imhof, A., Straub, T., and Becker, P.B. (2012). MSL2 combines sensor and effector functions in homeostatic control of the Drosophila dosage compensation machinery. Molecular cell *48*, 647-654.

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

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nature Reviews Molecular Cell Biology *13*, 115-126.

Waldeyer, W. (1888). Über Karyokinese und ihre Beziehungen zu den Befruchtungsvorgängen. Arch mikrosk Anat *32*, 1-122.

Wang, C., Cai, W., Li, Y., Deng, H., Bao, X., Girton, J., Johansen, J., and Johansen, K.M. (2011a). The epigenetic H3S10 phosphorylation mark is required for counteracting heterochromatic spreading and gene silencing in Drosophila melanogaster. J Cell Sci *124*, 4309-4317.

Wang, C., Cai, W., Li, Y., Girton, J., Johansen, J., and Johansen, K.M. (2012). H3S10 phosphorylation by the JIL-1 kinase regulates H3K9 dimethylation and gene expression at the white locus in Drosophila. Fly (Austin) *6*, 93-97.

Wang, C., Girton, J., Johansen, J., and Johansen, K.M. (2011b). A balance between euchromatic (JIL-1) and heterochromatic [SU(var)2-5 and SU(var)3-9] factors regulates position-effect variegation in Drosophila. Genetics *188*, 745-748.

Wang, C., Li, Y., Cai, W., Bao, X., Girton, J., Johansen, J., and Johansen, K.M. (2014). Histone H3S10 phosphorylation by the JIL-1 kinase in pericentric heterochromatin and on the fourth chromosome creates a composite H3S10phK9me2 epigenetic mark. Chromosoma *123*, 273-280.

Wang, C.I., Alekseyenko, A.A., LeRoy, G., Elia, A.E., Gorchakov, A.A., Britton, L.M., Elledge, S.J., Kharchenko, P.V., Garcia, B.A., and Kuroda, M.I. (2013). Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. Nature structural & molecular biology *20*, 202-209.

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

Weaver, T.M., Morrison, E.A., and Musselman, C.A. (2018). Reading More than Histones: The Prevalence of Nucleic Acid Binding among Reader Domains. Molecules *23*, E2614.

Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., Tanaka, K., Ren, Y., Xia, Z., Wu, J.*, et al.* (2014). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature *508*, 263-268.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M.*, et al.* (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. Cell *125*, 467-481.

Wolfe, S.A., Grant, R.A., Elrod-Erickson, M., and Pabo, C.O. (2001). Beyond the "recognition code": structures of two Cys2His2 zinc finger/TATA box complexes. Structure *9*, 717-723.

Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct *29*, 183-212.

Woo, H., Dam Ha, S., Lee, S.B., Buratowski, S., and Kim, T. (2017). Modulation of gene expression dynamics by co-transcriptional histone methylations. Experimental &Amp; Molecular Medicine *49*, e326.

Woodcock, C.L., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. I. Evidence for their general occurrence. Exp Cell Res *97*, 101-110.

Woodcock, C.L., Skoultchi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *14*, 17-25.

Wu, H., Zeng, H., Lam, R., Tempel, W., Amaya, M.F., Xu, C., Dombrovski, L., Qiu, W., Wang, Y., and Min, J. (2011a). Structural and histone binding ability characterizations of human PWWP domains. PloS one *6*, e18919.

Wu, L., Zee, B.M., Wang, Y., Garcia, B.A., and Dou, Y. (2011b). 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. Molecular cell *43*, 132-144.

Wuttke, D.S., Foster, M.P., Case, D.A., Gottesfeld, J.M., and Wright, P.E. (1997). Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. Journal of molecular biology *273*, 183-206.

Yadav, T., Quivy, J.-P., and Almouzni, G. (2018). Chromatin plasticity: A versatile landscape that underlies cell fate and identity. Science *361*, 1332-1336.

Yoh, S.M., Lucas, J.S., and Jones, K.A. (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. Genes & development *22*, 3422-3434.

Youdell, M.L., Kizer, K.O., Kisseleva-Romanova, E., Fuchs, S.M., Duro, E., Strahl, B.D., and Mellor, J. (2008). Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36. Molecular and cellular biology *28*, 4915-4926.

Yuan, W., Xie, J., Long, C., Erdjument-Bromage, H., Ding, X., Zheng, Y., Tempst, P., Chen, S., Zhu, B., and Reinberg, D. (2009). Heterogeneous nuclear ribonucleoprotein L Is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity in vivo. The Journal of biological chemistry *284*, 15701-15707.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. The Journal of biological chemistry *286*, 7983-7989.

Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. Cell Res *21*, 564-578.

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., and Johansen, K.M. (2006). The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. Development *133*, 229-235.

Zhang, W., Jin, Y., Ji, Y., Girton, J., Johansen, J., and Johansen, K.M. (2003). Genetic and phenotypic analysis of alleles of the Drosophila chromosomal JIL-1 kinase reveals a functional requirement at multiple developmental stages. Genetics *165*, 1341-1354.

Zheng, S., Villa, R., Wang, J., Feng, Y., Wang, J., Becker, P.B., and Ye, K. (2014). Structural basis of X chromosome DNA recognition by the MSL2 CXC domain during Drosophila dosage compensation. Genes & development *28*, 2652-2662.

Zheng, S., Wang, J., Feng, Y., Wang, J., and Ye, K. (2012). Solution structure of MSL2 CXC domain reveals an unusual Zn3Cys9 cluster and similarity to pre-SET domains of histone lysine methyltransferases. PloS one *7*, e45437.

Zhou, J.C., Blackledge, N.P., Farcas, A.M., and Klose, R.J. (2012). Recognition of CpG island chromatin by KDM2A requires direct and specific interaction with linker DNA. Molecular and cellular biology *32*, 479-489.

Zhou, S., Yang, Y., Scott, M.J., Pannuti, A., Fehr, K.C., Eisen, A., Koonin, E.V., Fouts, D.L., Wrightsman, R., Manning, J.E.*, et al.* (1995). Male-specific lethal 2, a dosage compensation gene of Drosophila, undergoes sex-specific regulation and encodes a protein with a RING finger and a metallothionein-like cysteine cluster. EMBO J *14*, 2884-2895.

5 Table of figures

Figure 1. Structure of the canonical core nucleosome. 21

Figure 2. Models for 30 nm chromatin fibre. 23

Figure 3. Identified post-transcriptional modifications in histones 25

Figure 4. Post-transcriptional modifications of histone tails. 26

Figure 5. CXC domain binding to DNA. 30

Figure 6. C2H2 zinc finger domain binding to DNA. 31

Figure 7. Protein structure of PSIP1 PWWP domain. 33

Figure 8. PWWP domain binding to H3K36me3 nucleosome. 34

Figure 9. Dosage compensation of sex chromosomes. 35

Figure 10. Dosage compensation complex. 37

Figure 11. Protein domain architecture of MSL protein. 39

Figure 12. MSL recognition element. 40

Figure 13. Heterochromatin spreading in JIL-1 mutants. 46

6 Abbreviations

ADP	adenosine diphosphate
ATAC-seq	assay for transposase accessibly chromatin with high-throughput sequencing
ATP	adenosine triphosphate
bp	base pair
ChIP-chip	chromatin immunoprecipitation followed by microarray hybridisation
ChIP-seq	chromatin immunoprecipitation with high-throughput sequencing
CLAMP	Chromatin Linked Adaptor for MSL Proteins
CTD	C-terminal domain
DCC	dosage compensation complex
DIP-seq	DNA immunoprecipitation with high-throughput sequencing
DNA	deoxyribonucleic acid
E(z)	Enhancer of zeste
H3S10ph	phosphorylated histone H3 at serine 10
H3K36me3	trimethylated histone H3 at lysine 36
H4K16ac	acetylated histone H4 at lysine 16
HAS	high affinity site
HAT	histone acetyltransferase
HMT	histone methyltransferase
HP1	heterochromatin protein 1
IP	immunoprecipitation
JASPer	JIL-1 Anchoring and Stabilizing Protein
JJ-complex	JIL-1-JASPer complex
KDM	lysine demethylase
lncRNA	long non-coding RNA
MBT	malignant brain tumor
MLE	maleless
MOF	males absent on the first

MRE	MSL recognition element
MRG15	MORF-related gene 15
mRNA	messenger ribonucleic acid
MSL	male-specific-lethal
NFR	nucleosome free region
NGS	next generation sequencing
NMR	nuclear magnetic resonance
PHD	plant homeodomain
PionX	Pioneering-sites-on-the-X
PTM	post-transcriptional modification
PWM	position weight matrix
RING	really interesting new gene
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RNAi	RNA interference
roX	RNA on the X
siRNA	small interfering RNA
SRI	Set2-Rpb1 interaction
SUMO	Small Ubiquitin-related Modifier
Su(var)	suppressor of variegation
SXL	Sex lethal
TAD	topologically associating domain
TF	transcription factor
ZnF	zinc finger

7 Danksagung

Viele tolle Menschen haben mir in den letzten fünf Jahren geholfen meine nun endlich vorliegende Doktorarbeit fertigzustellen. Ihnen allen gilt mein tiefster Dank!

Besonders herzlich will ich mich bei Peter Becker bedanken, dafür dass du mir dein Vertrauen entgegen gebracht und mich als Doktorand der ersten QBM-Generation in deiner Arbeitsgruppe aufgenommen hast. Während der gesamten Zeit habe ich viel Unterstützung und Beratung von dir erhalten, auch warst du mir immer ein Vorbild des integren Wissenschaftlers. Du hast mir den nötigen Freiraum gegeben meinen eigenen Weg einzuschlagen und mich in die Bioinformatik einzuarbeiten, nur dadurch konnte ich mich zu einem eigenständigen Wissenschaftler entwickeln. Vielen Dank!

Ganz vielen lieben Dank gilt auch Catherine Regnard, für deine immerwährende Unterstützung und Betreuung, sowie herausragende Zusammenarbeit. Die vielen unerlässlichen Diskussionen mit dir über die ganzen Experimente, Daten und Auswertungen haben meine Arbeit immer bereichert. Auch während schwierigeren Phasen meiner Doktorarbeit warst du für mich da. Vielen Dank für deine Freundschaft!

Liebsten Dank natürlich auch an Nadine Harrer, Lisa Harpprecht, Johanna Ludwigsen und Silke Krause für die besten Mittags- und Kaffeepausen, sowie für eure Freundschaft auch außerhalb des Labors.

Zusätzlich will ich mich noch besonders bei Silke Krause für deine phantastische technische Unterstützung in allen Belangen des Laboralltags bedanken.

Dem Becker-Labor gilt mein herzlicher Dank für eure herausragenden fachlichen Diskussionen, tolle Arbeitsatmosphäre und vieles mehr!

Mein bester Dank gilt Tamás Schauer für deine unschätzbare Hilfe in der Bioinformatik, sowie Tobias Straub für deine großartigen Bioinformtikkurse, die Etablierung des CompBio Clubs und Bereitstellung des High Performance Computational Clusters.

Ich will mich auch herzlich bei Stefan Krebs und LaFuGa für euren großartigen Next Generation Sequencing Service danken. Auch Ignasi Forné, Axel Imhof und dem Zentrallabor für Proteinanalytik für euren Massenspektrometrie-Service gilt mein bester Dank.

Vielen Dank an die Graduiertenschule für Quantitative Biowissenschaften München für finanzielle Unterstützung, das wertvolle Kursangebot und die vielen weiteren Veranstaltungen.

Vielen Dank an alle Kollaborationspartner für eure wichtigen Beiträge zu den Veröffentlichungen aus denen diese Doktorarbeit hervorgeht.

Zum Schluss will ich mich noch von ganzem Herzen bei meinen Eltern für eure fortwährende und uneingeschränkte Unterstützung bedanken. Ihr seid die Besten!

Liebe Sabrina, deine bedingungslose Liebe und dein grenzenloser Beistand sind unermesslich. Danke für unsere kleine Familie! Ich liebe dich wahnsinnig dolle!