Chapter 7
roX RNAs and Genome Regulation in Drosophila Melanogaster

S. Kiran Koya and Victoria H. Meller

Abstract Organisms with dimorphic sex chromosomes suffer a potentially lethal imbalance in gene expression in one sex. Addressing this fundamental problem can be considered the first, and most essential, aspect of sexual differentiation. In the model organisms Drosophila, Caenorhabditis elegans, and mouse, expression from X-linked genes is modulated by selective recruitment of chromatin-modifying complexes to X chromatin. In both flies and mammals, large noncoding RNAs have a central role in recruitment and activity of these complexes. This review will summarize current knowledge of the function of the noncoding roX genes in this process in Drosophila. Identification of an autosomal function for the roX RNAs raises intriguing questions about the origin of the modern dosage compensation system in flies.

7.1 Introduction

7.1.1 Genome Regulation and Large, Noncoding RNAs

Control of gene expression is central to life in all organisms. In addition to local gene regulation, many eukaryotes rely on coordinated control large chromatin domains. These clusters of coregulated genes can be as large as an entire chromosome. While the mechanisms that coordinate control of groups of genes are often poorly understood, the frequent association of large, noncoding RNAs (lncRNAs) with this process suggests that RNA is extremely well suited for regional chromatin regulation. The most dramatic example of this is sex chromosome dosage compensation in flies and mammals. Many diploid species, such as Caenorhabditis elegans, Drosophila, and mammals, have dimorphic sex chromosomes. Females and
C. elegans hermaphrodites have two X chromosomes, but males have a single X chromosome (XY or XO). In these species, the Y chromosome is gene poor, and the genes present on it are expressed only in testes. The resulting imbalance in the ratio of X to autosomal gene expression is potentially lethal to one sex (Gupta et al. 2006; Nguyen and Distech 2006). Several independently evolved strategies to balance X-linked gene expression between the sexes, a process called dosage compensation, have arisen (Lucchesi et al. 2005). Drosophila males increase transcription from their single X chromosome. This increase requires transcript produced from the X-linked \( \text{roX} \) \((\text{RNA on the X} \ 1, \ -2)\) genes. Mammalian females silence transcription from most genes on one of their two X chromosomes. \( \text{Xist} \) \((\text{X inactive specific transcript})\) is a lncRNA produced from the \( \text{X inactivation center} \) \((\text{Xic})\) [reviewed by Plath et al. (2002) and Chap. 3]. Production of \( \text{Xist} \) induces silencing of the X chromosome on which it is situated, one of the two X chromosomes present in females. In spite of striking differences in dosage compensation between flies and mammals, both employ lncRNAs that regulate this process and are necessary for identification of X chromatin. This convergence of function suggests that lncRNAs are particularly well suited for the regulation of broad chromatin domains. This review will explore the regulatory role the \( \text{roX} \) transcripts in Drosophila.

### 7.2 Noncoding RNAs in Drosophila: A Wealth of Transcripts with Few Known Functions

The Drosophila melanogaster genome consists of about 15,000 genes. While only a few hundreds are currently annotated as noncoding RNAs, this group is poorly understood, and annotations of noncoding transcripts lag the rest of the genome (Tweedie et al. 2009). The best-studied noncoding RNAs participate in translation and have well-defined functions that are determined by the structure of the RNA in question. But, in addition, there are numerous long, spliced, and polyadenylated transcripts that appear similar to mRNAs but lack significant open reading frames (Tupy et al. 2005). Identification of these began over two decades ago, but just a handful of the predicted lncRNAs in the fly genome have been studied in any detail. Of these, the \( \text{roX} \) RNAs are perhaps the best understood. The potent regulatory effects of the \( \text{roX} \) genes raise the question of whether some of the many transcripts with no known function may have similar actions in genomic regulation.

### 7.3 \( \text{roX} \) RNAs and Dosage Compensation

To overcome the potentially lethal imbalance in gene expression caused by hemizygosity of the X chromosome, male flies increase expression from almost all genes on their single X chromosome. This ensures a constant ratio of X to autosomal gene
product in both sexes. To achieve this, a complex of protein and roX RNA, termed
the Male Specific Lethal complex (MSL complex or dosage compensation com-
plex, DCC), is recruited to the X chromosome with exquisite selectivity, as illu-
strated in Fig. 7.1a, b. The MSL complex alters expression by modifying chromatin
within the body of transcribed genes. roX1 and roX2 are polyadenylated, noncoding
RNAs that are dissimilar in size and sequence (Amrein and Axel 1997; Meller et al.
1997). The major forms of roX1 are almost 4 kb, but the most abundant form of
roX2 is only 500 bp (Park et al. 2005). In spite of their dramatic difference in size
and sequence, roX1 and roX2 are redundant for all known functions. Mutation of
either roX gene alone has no phenotype, but simultaneous mutation of both is male-
lethal. Chromosome preparations from roX1 roX2 males display reduced X locali-
zation of the proteins in the MSL complex, and these proteins now can be observed
binding at ectopic sites throughout the genome (Fig. 7.1c). In contrast, females
mutated for both roX genes display no detectable phenotype and are fully viable
(Deng et al. 2009; Meller and Rattner 2002).

Although both roX genes are situated on the X chromosome, transcripts from
autosomal roX transgenes will assemble with the MSL complex, bind to the
X chromosome, and rescue roX1 roX2 males (Meller and Rattner 2002). roX RNA
can therefore travel through the nucleoplasm to regulate a chromosome in
trans to its site of transcription. This suggests that the roX genes act in a fundamen-
tally different way than Xist, whose action appears limited to its chromosome of
origin, a feature necessitated by the need to protect one X chromosome from
inactivation. However, roX also has the ability to direct binding of the MSL

---

Fig. 7.1 The roX transcripts localize to the X chromosome and are necessary for X chromosome
recognition. (a) roX1 coats the X chromosome in a male salivary gland. In situ hybridization to an
antisense roX1 probe is detected by alkaline phosphatase staining (purple). (b) MSL1 localization
in a polytenic preparation from a wild type male reveals exclusive localization to the X chromo-
some. Anti-MSL1 is detected by Texas Red, DNA by DAPI. (c) MSL1 localization in roX1 roX2
males is no longer exclusive to the X chromosome (X) but appears at a number of ectopic autosomal
sites.
complex to autosomal chromatin in cis to roX transgenes (Kageyama et al. 2001; Kelley et al. 1999). While the mechanisms that underlie the ability to recruit the MSL complex in cis remain speculative, all studies to date suggest that recruitment is determined by the ratio of MSL protein to roX RNA (Kelley et al. 2008; Oh et al. 2003; Park et al. 2002). High levels of MSL protein are proposed to allow formation of intact complexes as roX RNA is being transcribed, favoring localization close to the site of transcription. In contrast, when protein levels are low and roX transcription is high, roX will be released from its site of synthesis before assembly of the complex, eliminating the preference for local binding. Many questions remain about the precise molecular mechanisms by which roX RNAs act, but it is clear that the roX genes are central to X chromosome targeting.

7.4 Proteins of the MSL Complex

The roX RNAs assemble with five proteins, collectively known as the Male-Specific Lethals (MSLs; reviewed by (Gelbart and Kuroda 2009; Mendjan and Akhtar 2007)). These are MSL1, -2, and -3 (Male Specific Lethal 1, -2, and -3), MLE (Maleless), and MOF (Males absent on the first). All of the MSL proteins are necessary for dosage compensation. Mutation of any one of the msl genes causes male lethality as third instar larvae or pupae. In spite of the male-limited role of the MSL complex, most of the MSL proteins are present in both sexes, albeit at lower levels in females (Chang and Kuroda 1998; Lyman et al. 1997). MSL2 is the sole member of the complex whose expression is limited to males (Bashaw and Baker 1995; Kelley et al. 1995; Zhou et al. 1995). Translation of MSL2 mRNA is blocked by the Sexlethal protein, which is present only in females (SXL; Cline and Meyer 1996; Gebauer et al. 1998). Ectopic expression of MSL2 in females leads to formation of intact MSL complexes that bind both female X chromosomes, causing female lethality (Kelley et al. 1995). Female lethality is presumably due to elevated expression from both X chromosomes. This supports the idea that maintenance of the correct ratio of X to autosomal gene products is critical for normal development.

MSL2 and MSL1 are essential for all chromatin binding by the rest of the complex (Li et al. 2005; Lyman et al. 1997). MSL1 and MSL2 interact with each other, and this protein–protein interaction has been postulated to form a joint DNA-binding surface, although biochemical studies have yet to confirm this (Copps et al. 1998; Li et al. 2008; Rodriguez et al. 2007). MSL1 also serves as a scaffold for assembly of the other MSL proteins, as carboxy-terminal regions of MSL1 interact with MOF and MSL3 (Morales et al. 2004; Scott et al. 2000). MOF and MSL3 also interact with each other, and both proteins have been reported to bind RNA (Akhtar et al. 2000; Buscaino et al. 2003). MLE is the only member of the complex whose association with the other MSL proteins appears to be indirect (Copps et al. 1998). MLE is an RNA and DNA helicase of the DEAH subfamily (Kuroda et al. 1991). MLE association with the polytene X chromosome is RNA dependent (Richter et al. 1996). MLE can be coimmunoprecipitated with roX2 RNA from SL2 cells.
In early embryos, \textit{roX1} stability depends on maternally deposited MLE (Meller 2003). Taken together, these observations suggest that MLE is tethered to the MSL complex through \textit{roX} RNA, and, in the absence of MLE, \textit{roX} RNAs are not integrated into the MSL complex, subjecting them to rapid degradation. Localization of MOF and MSL3 has also been reported to be sensitive to RNase treatment (Akhtar et al. 2000; Buscaino et al. 2003). \textit{roX} RNAs are thus believed to play a major role in the assembly of the MSL complex, and their continued presence may be necessary for stable binding to the X chromosome.

While elimination of MSL1 or MSL2 results in loss of all chromatin binding by the remaining member of the complex, elimination of MLE, MSL3, or MOF leaves residual MSL proteins bound at a subset of X-linked sites. These are proposed to be recruitment sites from which the MSL complex can spread to nearby genes (Kelley et al. 1999). Indeed, recent studies have identified a short sequence motif enriched at sites of strong MSL3-independent binding, termed MSL recognition elements (MRE; Alekseyenko et al. 2008; Straub et al. 2008). Current models propose that the complete pattern of MSL binding along the X chromosome involves attraction of the MSL complex to sites containing MREs, followed by spreading into transcribed genes situated nearby (Gelbart and Kuroda 2009). Examination of MSL binding at high resolution revealed enrichment in the body and 3’ ends of actively transcribed genes (Larschan et al. 2007). This pattern is similar to that of the cotranscriptional H3K36me3 mark, and this observation is explained by the finding that the MSL3 chromodomain binds H3K36me3 (Sural et al. 2008). Mutation of conserved residues in the MSL3 chromodomain disrupts normal spreading of the MSL complex into transcribed genes. Taken together, these studies support an elegant model that explains the local distribution of the MSL complex on the X chromosome. What this model fails to explain is how the MSL complex is limited to the X chromosome. All transcribed genes share enrichment for H3K36me3, and MREs are only modestly enriched on the X chromosome. It thus seems unlikely that these are the only factors directing localization of the MSL complex. A recent study determined that during interphase, regions of the X chromosome with high affinity for the MSL complex are closer together than regions with little or no binding (Grimaud and Becker 2009). This suggests that interphase chromosome architecture might be a factor in selective recognition of X chromatin.

MOF is a histone acetyltransferase specific for lysine 16 on H4 (H4Ac16) (Akhtar and Becker 2000; Hilfiker et al. 1997; Smith et al. 2000). Histone acetylation in general is thought to reduce the strength of histone–DNA interactions, making DNA more accessible. The H4Ac16 modification specifically prevents tight packing of nucleosomes, and this may contribute to elevated expression, as well as the slightly decondensed character of the male X chromosome (Shogren-Knaak et al. 2006). The distribution of H4Ac16 enrichment is similar to that of the MSL complex, being more pronounced in the 3’ ends of genes and coding regions than on promoters of transcribed genes (Kind et al. 2008). As chromatin modification by the MSL complex occurs mainly within the body of genes, it is likely that enhanced transcription is due to facilitation of elongation, rather than initiation of
transcription (Smith et al. 2001). Modulation of a general property of RNA pol II, such as speed or processivity, would explain how the MSL complex achieves a uniform two-fold increase in expression of thousands of genes with disparate expression patterns and regulatory regions.

A second chromatin modification is enriched on the male X chromosome and depends on the MSL complex. The JIL-1 kinase is an essential protein required in both sexes, yet partial loss of function alleles affect males more severely than females, and a genetic study suggests a role for JIL-1 in compensation of an X-linked gene (Jin et al. 1999; Lerach et al. 2005). JIL-1 phosphorylates serine 10 on histone 3 (H3pS10), a mark that is associated with open chromatin structure and increased gene expression (Wang et al. 2001). JIL-1 localizes to interband regions on all chromosomes and is enriched on the male X chromosome. On the X, JIL-1 colocalizes with MSL proteins and, under some conditions, it may immuno-precipitate with the MSL complex, suggesting a possible molecular interaction (Jin et al. 2000). However, it remains unclear if JIL-1 enrichment on the X chromosome is due to a direct interaction with the MSL complex or if it is the consequence of MSL complex action, for example, a response to elevated transcription or chromatin modification by the MSL complex.

### 7.5 Separate Domains of roX1 Regulate X-Localization and Histone Modification

At least one roX transcript is essential for targeting the intact MSL complex to the male X chromosome (Deng and Meller 2006; Meller and Rattner 2002). In roX1 roX2 males, the proteins of the MSL complex still colocalize but are no longer exclusive to the X chromosome. Although MOF is present at these ectopic autosomal sites in roX1 roX2 males, H4Ac16 modification at these sites is low, suggesting that roX association with the MSL complex is necessary for full MOF activity (Deng and Meller 2006). Interestingly, the 3′ end of roX1 contains a stem loop that is necessary for roX1 function, but deletion of this portion of the transcript has a relatively mild effect on X-localization (Stuckenholz et al. 2003). In addition, short repeats in the 3′ end of roX1 are also present in roX2 and in the roX genes of numerous related species (Franke and Baker 1999; Kelley et al. 2008). The presence of these repeats appears to regulate activity of the complex (Kelley et al. 2008; Park et al. 2007). While it is tempting to speculate that roX RNA is the allosteric regulator of MOF, other MSL proteins also influence MOF activity. Interaction of MOF with a subcomplex of MSL1 and MSL3 increases the efficiency and substrate specificity of MOF (Morales et al. 2004). This emphasizes the point that the normal activity of MOF requires assembly of the intact MSL complex.

In contrast to the function of 3′ roX1 sequences, deletions removing significant portions of the 5′ end affect localization of the complex to the X chromosome. This region comprises almost 1.5 kb and lacks obvious repetitive sequences or secondary structures of high stability. A scanning deletion analysis that removed ~300 bp
portions of roX1 failed to identify essential elements in the 5' end (Stuckenholz et al. 2003). However, deletions removing large portions of this region reduce X chromosome binding of the MSL complex, although mutants retaining even a very small portion of the 5' end support partial dosage compensation (Deng and Meller 2008; Deng et al. 2005). Together, these studies support the idea that separate regions of roX1 direct MSL complex localization and the chromatin-modifying activity of the complex. This is reminiscent of the distribution of function in Xist. Short, tandem stem loops are necessary for Xist-mediated chromatin silencing in mice, but painting of the X chromosome is directed by several large segments of Xist that may work cooperatively to ensure X recognition (Wutz et al. 2002).

7.6 Ancestral Origins of Complexes that Dosage Compensate Sex Chromosomes

The sex chromosomes of mammals, C. elegans, and flies are unrelated to each other in evolutionary origin. Indeed, the de novo origin of differentiated sex chromosomes has occurred repeatedly in different animal lineages (Bull 1985). In accordance with this independent origin, sex chromosome dosage compensation has arisen independently many times. While each system for dosage compensation achieves the goal of maintaining an appropriate ratio of X to autosomal gene products, each has adopted a completely different strategy to do so. All three systems have developed through recruitment of preexisting chromatin regulatory complexes (Table 7.1). For example, the DCC of C. elegans is related to, and shares subunits with, the condensin complex that compacts chromosomes and enables normal segregation during mitosis and meiosis (Chan et al. 2004; Csankovszki et al. 2009). In accordance with this, some mutations that disrupt C. elegans dosage compensation also disrupt meiosis and mitosis (Hagstrom et al. 2002; Lieb et al. 1996). Silencing of an X chromosome in female mice is a complex process that takes place over several days during early embryogenesis, but an early event is recruitment of the Polycomb group 2 and 1 (Pcg2, Pcg1) complexes (Schoeftner et al. 2006; Zhao et al. 2008). Pcg2 deposits the silencing H3K27 trimethylation mark, and Pcg1 ubiquitinates H2 on K119 (Kohlmaier et al. 2004; Plath et al. 2003). These modifications may contribute to

Table 7.1 Dosage compensation recruits existing chromatin-modifying complexes for novel functions

<table>
<thead>
<tr>
<th>Organism</th>
<th>Compensation machinery</th>
<th>Compensatory function</th>
<th>Ancestral complex and function</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. elegans</td>
<td>DCC complex</td>
<td>Downregulation of X-linked genes</td>
<td>Condensin mitotic and meiotic chromosome condensation</td>
</tr>
<tr>
<td>Mammals</td>
<td>Prc1, Prc2</td>
<td>Inactivation of X chromosome</td>
<td>Prc1, Prc2 developmentally stable repression</td>
</tr>
<tr>
<td>Drosophila</td>
<td>MSL complex</td>
<td>Upregulation of X-linked genes</td>
<td>?</td>
</tr>
</tbody>
</table>
the stability of X inactivation. In addition to their role in X inactivation, both complexes continue to function in epigenetic repression throughout the genome during mammalian development (Bernstein et al. 2007).

While the ancestral functions of the proteins that achieve dosage compensation in Drosophila remain to be fully defined, homologs of MOF, MSL1, MSL2, and MSL3 have been found in organisms as diverse as yeast and mammals (Eisen et al. 2001; Smith et al. 2005; Marin 2003; Sanjuan and Marin 2001). With the exception of MLE, the mammalian homologs associate with each other, suggesting that the modern MSL complex of flies has an ancient origin. Human MOF (hMOF) is notable as it participates in multiple complexes, and these are responsible for the majority of H4KAc16 modification in mammalian cells (Cai et al. 2010; Mendjan et al. 2006; Smith et al. 2005). Although the precise molecular function of hMOF-containing complexes is not clear, depletion of hMOF affects DNA repair, possibly by disruption of damage signaling (Gupta et al. 2005; Taipale et al. 2005). MOF also participates in multiple complexes in flies, which may allow it to serve as a general regulator of chromatin at promoters, although this finding remains controversial (Mendjan et al. 2006; Gelbart et al. 2009; Kind et al. 2008). Recent work in our laboratory suggests a different autosomal role for MOF, and other MSL proteins, in flies.

7.7 Regulation of Heterochromatic Genes by roX and a Subset of MSL Proteins

In addition to reduced expression of X-linked genes in roX1 roX2 males, several hundred autosomal genes situated in heterochromatic regions are also misregulated (Deng et al. 2009). Regions containing misregulated genes include the entire 4th chromosome. The 4th chromosome has several peculiarities, including its small size, lack of recombination, and possible evolutionary kinship with the X chromosome (Larsson and Meller 2006; Riddle and Elgin 2006). However, the fact that the 4th chromosome is enriched for heterochromatin is the feature that it shares with the other autosomal genes that depend on roX RNA for full expression. Unexpectedly, this feature of heterochromatic gene regulation is limited to males (Deng et al. 2009). Analysis of expression in msl mutants revealed that MSL1, MSL3, MLE, and MOF are also required for full expression of heterochromatic and 4th-linked genes in males. However, no misregulation of these autosomal genes is observed in msl2 mutants, indicating that the intact MSL complex is not involved (Deng et al. 2009). Because MSL2 is the sole member of the MSL complex that is strictly male-limited, it remains unclear how the sex-specificity of heterochromatic gene regulation is maintained. As MSL1 and MSL2 are postulated to work together to target the MSL complex to the X chromosome, it appears likely that MSL2 is dedicated for the recognition of the X chromosome (Li et al. 2008; Rodriguez et al. 2007). The X chromosome is about twofold enriched for MREs in comparison to the autosomes, but interestingly, MREs are depleted from the 4th chromosome (Alekseyenko et al. 2008). This reinforces the idea that although regulation of X-linked and
heterochromatic genes requires overlapping sets of molecules, recognition of these two groups occurs by different mechanisms. To explain these findings, we have proposed that a second complex composed of \textit{roX} RNA and subset of MSL proteins is responsible for the modulation of chromatin at autosomal heterochromatic sites in males.

It is tempting to speculate that regulation of heterochromatic genes reflects an ancestral function of the members of the MSL complex. Heterochromatic genes are situated in a difficult environment and have long been thought to utilize specialized regulatory mechanisms (Yasuhara and Wakimoto 2006). However, the limitation to males suggests a process that coevolved with the modern sex chromosomes of flies. One possibility is that the highly differentiated sex chromosomes create nuclear environments that are sufficiently different to require a dedicated regulatory system in one sex. As the \textit{Drosophila} Y chromosome is large and entirely heterochromatic, it is plausible that it alters the balance of chromatin proteins throughout the nucleus (Weiler and Wakimoto 1995). Taking into account the multiple functions of \textit{roX} RNA in genome regulation, we present a hypothetical model for the origin of \textit{roX}-dependent complexes in Fig. 7.2. In human cells, homologs of MLE and \textit{roX} have

![Fig. 7.2 Proposed origin \textit{roX} RNA complexes in \textit{Drosophila}. As homologs of MSL1, MSL2, MSL3, and MOF associate in many organisms, this association may represent the ancestral form of the complex in flies (top). We speculate that acquisition of MLE and \textit{roX} was an early event in formation of the modern MSL complex, followed by coevolution of MSL1 and MSL2 (left), a feature that may determine X recognition. Loss or replacement of MSL2 may enable complex members to acquire autosomal functions such as regulation of genes in heterochromatic environments (right). While the association of \textit{roX} and MSL proteins in the intact MSL complex (left) is well established, the presence of a subcomplex lacking MSL2 (right) has yet to be established.](image)
not been identified in association with the other hMSL proteins (Cai et al. 2010; Smith et al. 2005; Taipale et al. 2005). Because of this, we have chosen to model the acquisition of roX and MLE as an early step in the evolution of modern roX-containing complexes of flies. Rapid co-evolution of the MSL2 and MSL1 interaction domains may enable recognition of MREs, likely to be an essential function of the MSL complex that dosage compensates the X chromosome (left). MSL2 is the sole member of this complex lacking a heterochromatic role and is thus anticipated to be dedicated to X recognition. We speculate that loss or replacement of MSL2 has enabled the remaining MSL proteins and roX RNA to be recruited for a new purpose, regulation of heterochromatic genes in males. While it appears logical that a subset of MSL proteins and the roX RNAs form a second complex, the existence of this has yet to be demonstrated. While studies in flies have identified multiple MOF-containing complexes, the technique used, affinity purification followed by mass spectrometric analysis, would not reveal a minor contribution of a subcomplex lacking MSL2 (Mendjan et al. 2006). The mechanism by which autosomal genes are regulated by the roX RNAs remains to be fully elucidated.

7.8 Conclusions

The large noncoding roX RNAs have a central role in sex chromosome dosage compensation in flies, where they fulfill a role with similarities to that of Xist during mammalian dosage compensation. roX transcripts assemble with the MSL proteins to form a complex that displays exclusive X chromosome binding. Situation of the roX genes on the X chromosome facilitates X recognition through the ability of the roX genes to attract the MSL complex in cis. These observations created the impression that the roX RNAs were dedicated to identification and modification of the X chromosome, but this idea has been revised by the unexpected discovery of a role for roX in expression of heterochromatic genes. It appears that both regulatory systems may be necessitated by the presence of highly differentiated sex chromosomes.

References


