N-Terminus of *Drosophila Melanogaster* MSL1 Is Critical for Dosage Compensation

Valentin Babosha[1, 2], Natalia Klimenko[2], Anastasia Revel-Muroz[2], Evgeniya Tikhonova[1, 2], Pavel Georgiev[1,]*, Oksana Maksimenko[2]*

[1]Department of the Control of Genetic Processes, Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilova St., Moscow 119334, Russia;
[2]Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilova St., Moscow 119334, Russia

*Corresponding author:
Oksana Maksimenko: maksog@mail.ru
Pavel Georgiev: georgiev_p@mail.ru

Key words: sex determination, transcription factor, MSL2, MSL3, RNA-protein interaction, MSL complex

Abstract

The male-specific dosage compensation complex (DCC), which consists of five proteins and two non-coding roX RNAs, is necessary for the transcriptional enhancement of X-linked genes to compensate for the sex chromosome monosomy in *Drosophila* XY males compared with XX females. The MSL1 and MSL2 proteins form the heterotetrameric core of DCC and are critical for the specific recruitment of the DCC to the high-affinity “entry” sites (HAS) on the X chromosome. In this study, we demonstrated that the N-terminal region of MSL1 is critical for stability and functions of MSL1. Amino acid deletions and substitutions in the N-terminal region

of MSL1 strongly affect both the interaction with roX2 RNA and DCC binding to HAS on the X chromosome. In particular, substitution of the conserved N-terminal amino-acids 3-7 in MSL1 ($MSL1^{GS}$) affects dosage compensation similar to the inactivation of genes encoding roX RNAs. In addition, $MSL1^{GS}$ binds to promoters such as $MSL1^{WT}$ but does not co-bind with MSL2 and MSL3 to X chromosomal HAS. However, overexpression of MSL2 partially restores the functional activity of $MSL1^{GS}$ in dosage compensation. Thus, the interaction of MSL1 with roX RNA is critical for the efficient assembly of DCCs on HAS of the male X chromosome.

Introduction

At present, it is unknown how transcription-regulating complexes bind only to certain chromatin regions that do not have a pronounced sequence specificity relative to other genome sites. Dosage compensation in *Drosophila* is a striking example of the specific recruitment of transcription complexes (Lucchesi, 2018). In particular, dosage compensation is performed by increasing the level of gene expression from the X chromosome of males (X/Y) compared with the gene expression from the X chromosome of females (X/X), despite having two X chromosomes. The multisubunit dosage compensation complex (DCC), which binds only to the male X chromosome (Ferrari et al., 2014; Kuroda et al., 2016; Lucchesi, 2018; Samata and Akhtar, 2018), is responsible for the observed gene activation. The DCC consists of five proteins, male-specific lethal (MSL)1, MSL2, MSL3, males-absent on the first protein (MOF), and maleless (MLE). Two non-coding RNAs are included, namely roX1 (3.7 kb) and roX2 (0.6 kb), which perform partially redundant functions (Kim et al., 2018; Kuroda et al., 2016; Meller and Rattner, 2002; Samata and Akhtar, 2018). The proteins MSL1, MSL3, MOF, and MLE are also expressed in females. These proteins and are involved in the regulation of gene expression processes that are unrelated to X chromosome dosage compensation (Samata and Akhtar, 2018). The MSL2 protein is male-specific, and the expression of the *msl-2* gene is controlled by the sex lethal (SXL) protein, which binds to *msl-2* mRNA in females and blocks protein translation

(Beckmann et al., 2005). Therefore, MSL2 protein is thought to play a key role in the specific binding of the DCC to the X chromosome in males.

The N-terminus of MSL2 contains a really interesting new gene (RING) domain, which functions as a ubiquitin E3 ligase (Villa et al., 2012; Wu et al., 2011). MSL2 has been shown to induce the ubiquitination of itself as well as other DCC proteins. The MSL1 protein functions as a scaffold for the assembly of the DCC (Gorman et al., 1993; Kadlec et al., 2011; Palmer et al., 1993; Scott et al., 2000), as the N-terminal coiled-coil domain (Hallacli et al., 2012) of MSL1 facilitates homodimerization and interacts with the N-terminal RING domain of MSL2. The PEHE domain located in the C-terminal region of MSL1 is responsible for interactions with MSL3 and MOF (Kadlec et al., 2011; Scott et al., 2000). MSL1 and MSL2 form the structural core of the DCC, whereas MLE—an ATP-dependent RNA/DNA helicase belonging to the DEAD box family—interacts with the roX1 and roX2 RNAs to induce their unwinding (Ilik et al., 2017; Maenner et al., 2013; Meller et al., 2000). This interaction allows both RNAs to bind MSL2 and possibly MSL1 (Muller et al., 2020), combining MLE, MSL2, and other MSL proteins into a single complex.

An incomplete MSL1-MSL2 subcomplex can bind to a reproducible set of several hundreds of sites on the X chromosome, referred to as the primary chromatin entry sites (CESs) (Alekseyenko et al., 2008) or high-affinity sites (HAS) (Straub et al., 2008). A GA-rich motif within the HAS/CES, which serves as an MSL recognition element (MRE), is important for DCC targeting (Alekseyenko et al., 2008; Straub et al., 2008). Structural analysis and an *in vitro* genome-wide DNA binding assay showed that MSL2, through its CXC and proline/basic-residue-rich domains, specifically recognizes and binds to a subclass of HAS-named pioneering sites on the X (PionX) (Fauth et al., 2010; Villa et al., 2016; Zheng et al., 2014).

The zinc-finger protein chromatin-linked adaptor for MSL proteins (CLAMP) was found to be significant for the specific recruitment of the MSL complex to the male X chromosome (Eggers et al., 2023; Larschan et al., 2012; Soruco et al., 2013). CLAMP is an essential protein in

Drosophila that binds thousands of GA-rich sequences genome-wide in both sexes (Kuzu et al., 2016; Soruco et al., 2013; Urban et al., 2017). In particular, CLAMP binds to GA repeats in HAS and is involved in the organization of open chromatin in these genome regions (Albig et al., 2019; Urban et al., 2017). CLAMP has an N-terminal unstructured dimerization domain (Tikhonova et al., 2022a) and promotes long-range chromatin interactions that mediate the dosage compensation of the male X chromosome (Jordan and Larschan, 2021). CLAMP also directly interacts with the MSL2 protein (Albig et al., 2019; Eggers et al., 2023; Tikhonova et al., 2022b, 2019). The CLAMP interaction and CXC domains of MSL2 have partially redundant functions for the recruitment of the MSL complex to the X chromosome (Tikhonova et al., 2019). A previous study (Li et al., 2005) showed that the N-terminal 265 amino acids region of MSL1 could bind to approximately 30 sites on polytene X chromosomes in females that exogenously express MSL2, and the first 26 aa are essential for such binding. The authors suggested that the N-terminal sequences of MSL1 can bind to DNA and facilitate the specific recruitment of MSL to the male X chromosome. However, further studies examining the mechanisms of DCC recruitment to the male X chromosome have concentrated on MSL2 as this is the only DCC protein that is exclusively expressed in males [see review (Kuroda et al., 2016; Lucchesi, 2018; Samata and Akhtar, 2018)]. Although MSL1 represents a core protein of the DCC (Gorman et al., 1993; Palmer et al., 1993; Scott et al., 2000), MSL1 also interacts with gene promoters independently of the DCC (Chlamydas et al., 2016, p. 2; Straub et al., 2013).

Here, we tested the functional role of different regions in the N-terminal 1-85 aa region of MSL1. Deletion of the 1-15 aa results in the instability of MSL1. Unexpectedly, we found that the simultaneous replacement from 3 to 7 aa from the N-terminus or deletion from 41 to 65 aa of MSL1 strongly influences the binding of the MSL complex on the male X chromosome. These mutations also affect the interaction of MSL1 with roX2 RNA. Thus, the N-terminal region of MSL1 performs key functions in the recruitment of the DCC to the X chromosome.

Results

The N-terminal 85 aa amino acids of MSL1 are critical for the activity of the MSL complex

The principal goal of the work is to study the functional role of the N-terminal 85 aa MSL1 followed by the coiled-coil domain involved in homodimerization and interaction with MSL2. With the exception of the coiled-coil domain, the first 15 amino acids represent the most conserved region within the N-terminus of MSL1 among Drosophilidae (***Figure 1A***). However, among different families of Diptera (***Figure 1 – figure supplement 1***), a high level of homology was only observed for the coiled-coil region, which is required for interactions with MSL2 (Hallacli et al., 2012). The N-terminal amino acids are also conserved among representatives within each of the Diptera families ***(Figure 1 – figure supplement 1***), but not across families. For example, the MSL1 of *Anopheles* mosquitoes, which are closely related to *Drosophila melanogaster*, features a completely different N-terminal domain.

To evaluate the functional effects of N-terminal domain deletions in the MSL1 protein *in vivo*, we created transgenic flies expressing wild-type ($MSL1^{WT}$) and mutant variants of MSL1 ($MSL1^{\Delta 1\text{-}85}$, $MSL1^{\Delta 1\text{-}15}$, $MSL1^{\Delta 8\text{-}20}$, $MSL1^{\Delta 24\text{-}39}$, $MSL1^{\Delta 41\text{-}85}$, $MSL1^{\Delta 41\text{-}65}$, and $MSL1^{\Delta 66\text{-}85}$) (***Figure 1A***). The MSL1 variants were expressed under the control of a strong *Ubi-p63E* (Ubi) promoter (***Figure 1 – figure supplement 2***). To avoid the influence of position effects on the expression of the MSL1 variants, all transgenes were inserted into the same genomic location (86Fb) on the third chromosome using a φC31-based integration system (Bischof et al., 2007).

To determine the role played by the N-terminal sequence during the specific recruitment of DCC in males, we crossed obtained transgenes into the null *msl-1*$^{-}$ background (the *msl-1*$^{-}$ background corresponds to the *msl-1*$^{\gamma 269}$/*msl-1*^{L60} heterozygote) (Palmer et al., 1993). As shown previously, the main band detected in immunoblot analysis corresponds to a 170-kD protein that is considerably larger than the predicted 1039 aa protein (Palmer et al., 1994). Anomalous electrophoretic migration was explained by specific clusters of proline, acidic, or basic residues as well as potential post-translational modifications in particular ubiquitination. MSL1 is

expressed 2.5–3 times stronger under the control of the Ubi promoter in the transgenic line than in wild-type males (y^1w^{1118} line) (***Figure 1B***, ***Figure 1 – figure supplement 3A***). Increased expression of MSL1 leads to the appearance of more MSL2, which is stabilized in a complex with MSL1. Similar results were obtained in females. In the absence of MSL2, the expression of MSL1 was lower in females compared to males (***Figure 1B***).

To compare the expression of the MSL1 variants in transgenic lines, we examined the levels of MSL1 expression in 2–3-day-old *msl-1*$^-$ females, which are homozygous for the transgenes expressing MSL1 variants. Immunoblot analysis showed that all MSL1 variants, with the exception of MSL1$^{\Delta 1\text{-}85}$ and MSL1$^{\Delta 1\text{-}15}$, were expressed in transgenic females at nearly equivalent levels (***Figure 1C***). Surprisingly, MSL1$^{\Delta 1\text{-}85}$ and MSL1$^{\Delta 1\text{-}15}$ could not be detected by immunoblotting, suggesting the instability of these MSL1 variants.

Homozygous *msl-1*$^-$ females remained viable, whereas males died during the late larva and early pupa stages (Gorman et al., 1993; Palmer et al., 1993). The transgenes, expressed MSL1^{WT} and MSL1$^{\Delta 66\text{-}85}$, compensated for the viability of *msl-1*$^-$ males, indicating that these MSL1 variants are functional (***Figure 1D***, ***Figure 1 – source data 1***). In MSL1$^{\Delta 24\text{-}39}$ and MSL1$^{\Delta 8\text{-}20}$ lines, male viability was reduced suggesting that dosage compensation is partially affected. At the end, all other tested mutant MSL1 proteins (MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 1\text{-}85}$, MSL1$^{\Delta 41\text{-}65}$ and MSL1$^{\Delta 41\text{-}85}$) failed to compensate for the *msl-1*$^-$ mutation in males (***Figure 1D***). Taken together, these results suggested that at least two regions (1-15 aa and 41-65 aa) in the N-terminus of MSL1 are critical for the functional activity of the MSL complex.

We then asked how MSL2 expression might affect the stability of MSL1 variants. To answer this question, we measured MSL1, and MSL2 proteins in *msl-1*$^-$*; Ubi:msl-2*^{WT}*-FLAG/ Ubi:msl-1** (***Figure 1E***). Again, immunoblotting showed that amount of MSL1 was almost completely eliminated in females expressed either MSL1$^{\Delta 1\text{-}15}$ and MSL2, or MSL1$^{\Delta 1\text{-}85}$ and MSL2 (***Figure 1E***). All other MSL1 variants were present at the same level. Interestingly, there

is no noticeable difference in the amount of MSL2 in all samples. Thus, the N-terminal 1-15 aa are essential for stability of MSL1 independently of MSL2 presence.

The N-terminus of MSL1 regulates the binding of the MSL complex to the X chromosome in females expressing MSL2

We next examined the extent to which deletions in the N-terminal region of MSL1 would affect the recruitment efficiency of the MSL complex to the male X chromosome. Immunostaining of polytene chromosomes from the salivary glands of *Drosophila* larvae allows the visualization of proteins on interphase chromatin and has been used extensively to study dosage compensation (Dahlsveen et al., 2006; Demakova et al., 2003; Li et al., 2005; Lyman et al., 1997; Meller and Rattner, 2002; Palmer et al., 1994). In the MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 41\text{-}85}$, MSL1$^{\Delta 41\text{-}65}$ and MSL1$^{\Delta 1\text{-}85}$ mutant males, rare larvae surviving to 2-3 stages had the polytene chromosomes with poor morphology. For this reason, we used a previously described sensitive model system to study the factors required for the recruitment of the DCC to the X chromosome in females (Chang and Kuroda, 1998; Li et al., 2008; Morra et al., 2011; Palmer et al., 1994; Zhou et al., 1995). It was shown (Tikhonova et al., 2019), that MSL1 and MSL2 were specifically recruited to the X chromosome in transgenic females expressing MSL2^{WT}-FLAG under the control of the Ubi promoter inserted into 86Fb (*Ubi:msl-2^{WT}-FLAG*).

When the MSL2^{WT}-FLAG was expressed in the *msl-1*$^{-}$ background, we did not observe the binding of the DCC to the female X chromosome (data not shown). However, the binding of MSL1 and MSL2 to the X chromosome was detected in *msl-1*$^{-}$ females when MSL2^{WT}-FLAG was expressed simultaneously with either MSL1^{WT} or MSL1$^{\Delta 66\text{-}85}$ (***Figure 2, figure 2 – figure supplement 1***). These results further supported that the 66–85-aa region is not essential for MSL1 function in dosage compensation.

The expression of all MSL1 variants with deletions of the first 15-aa region (MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 1\text{-}85}$) resulted in the near-complete loss of DCC binding to the X chromosome in

females (***Figure 2, figure 2 – figure supplement 1***). This result is expected as immunoblotting showed a strong decrease of MSL1 level in MSL1$^{\Delta 1\text{-}15}$/MSL2^{WT} and MSL1$^{\Delta 1\text{-}85}$/MSL2^{WT} females (***Figure 1E***). Previous studies (Dahlsveen et al., 2006; Demakova et al., 2003; Lyman et al., 1997; Meller and Rattner, 2002; Palmer et al., 1994) showed that the inactivation of MSL3, MLE, or roX RNAs resulted in the recruitment of the MSL complex to a small proportion of bands that correspond to the main HAS, including regions of the *roX1* (3F) and *roX2* (10C) genes. In polytene chromosomes in the MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 1\text{-}85}$ lines, we found only two intense and several faint MSL1/MSL2 bands (***figure 2 – figure supplement 2***). Interestingly, additional MSL1/MSL2 bands were identified on the polytene autosomes, suggesting that the MSL proteins began to bind to lower affinity sites on autosomes (***figure 2 – figure supplement 1***). Thus, the deletion of the N terminal 15 aa affects not only stability of MSL1 but also specificity of MSL binding to the X chromosome.

The deletion of the 41–85 aa or 41-65 aa regions did not affect the stability of MSL1, but at the same time significantly reduced the binding of MSL proteins to the X chromosome. MSL1$^{\Delta 41\text{-}85}$ and MSL2 bind to several bands on the X chromosome, many of which can also be identified as weak bands in the MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 1\text{-}85}$ lines (***Figure 2, figure 2 – figure supplement 1,2***). In comparison with MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 1\text{-}85}$, MSL proteins bind more efficiently with the MSL1$^{\Delta 41\text{-}85}$ and MSL1$^{\Delta 41\text{-}65}$ polytene chromosomes. However, the binding of the MSL complex to the X chromosome in MSL1 mutants was much weaker than previously suggested for a minimal MSL complex lacking either MSL3 or MLE (Dahlsveen et al., 2006). As the 66–85-aa deletion did not affect activity of MSL1, we concluded that the 41-65-aa region is critical for MSL1 function.

Deletion of 24-39 aa, which separates the 1–15 aa and 41–65 aa regions in MSL1, only moderately affects binding of the MSL2 and MSL1$^{\Delta 24\text{-}39}$ proteins on polytene chromosomes (***Figure 2, figure 2 – figure supplement 1***). This observation indicates that the 1–15 aa and 41–65 aa regions have at least partially independent functions in recruiting of the MSL complex to

the X chromosome in males. In MSL1$^{\Delta 8\text{-}20}$ female larvae patterns of MSL1 and MSL2 binding to polytene chromosome are the same as in larvae expressed MSL1$^{\Delta 24\text{-}39}$ (Figure 2, Figure S4). Thus, a deletion from 8 aa up to 20 aa only partially affects the functions of MSL1, in contrast to the complete inactivation of this protein by deleting the terminal 15 aa (MSL1$^{\Delta 1\text{-}15}$).

The N terminal 3-7 amino acids of MSL1 are critical for dosage compensation

Our results showed that MSL1$^{\Delta 1\text{-}15}$ is unstable, while MSL1$^{\Delta 8\text{-}20}$ only partially affects the specific recruitment of MSL to the X chromosome suggesting that the most terminal amino-acids are essential for functions. To test the role of the N-terminal amino acids (3-7) in the stability and functions of MSL1, we made several amino acids substitution in this region. The N-terminal region contains well-conserved basic and aromatic amino acid residue (***Figure 1A***). We used data about the previously characterized MSL1 mutants (Li et al., 2005) in those three conserved basic amino acids, lysine 3, arginine 4, and lysine 6, or two conserved aromatic amino acids, phenylalanine 5 and tryptophan 7, were replaced by alanine. In our work we replaced basic amino acids by serine, aromatic amino acids – by alanine or glycine (***Figure 3A***). In result we obtained three different transgenes: MSL1^{3S} (K3S R4S K6S), MSL1^{2A} (F5A W7A), and MSL1^{GS} (K3S R4S F5G K6S W7G).

To evaluate the functional effects of N-terminal amino acids of the MSL1 protein *in vivo*, we created transgenic flies expressing mutants MSL1^{2A}, MSL1^{3S} and MSL1^{GS} under control of the Ubi promoter. All MSL1 variants were targeted by HAx3 epitope at the C-termini (***Figure 3A***). MSL1^{2A} and MSL1^{3S} variants were expressed at levels comparable to MSL1^{WT} in 2–3-day-old *msl-1*$^{-}$ females that are homozygous for the MSL1 transgenes (***Figure 3B, C***). At the same time, there was an approximately 1.5–2-fold decrease in the MSL1^{GS} protein compared to MSL1^{WT}.

The transgenes, expressed MSL1^{2A} and MSL1^{3S}, almost completely compensated for the viability of *msl-1*$^{-}$ males, indicating that these MSL1 variants are functional (***Figure 3D, Figure***

3 – source data 1). In the *msl-1*$^{-}$ line expressing MSL1^{GS}, the viability of males was strongly reduced. At the same time, increasing in MSL2 and a simultaneous decreasing in MSL1^{GS} expression partially restore viability of MSL1^{GS}/MSL2 males (***Figure 3E, Figure 3 – source data 2***). In similar way, viability of MSL1$^{\Delta 8\text{-}20}$/MSL2 males also is close to wild-type, but MSL1$^{\Delta 41\text{-}65}$/MSL2 and MSL1$^{\Delta 1\text{-}15}$/ MSL2 males are still lethal.

According to the results obtained on male, MSL1^{2A} and MSL1^{3S} showed binding patterns like MSL1^{WT} on the larva polytene chromosome, while the MSL1^{GS} protein failed to effectively bind on the X chromosome in complex with MSL2 (***Figure 3F, figure 3 – figure supplement 1,2***). There are also binding sites for only MSL2 or MSL1^{GS} on the autosomes and the X chromosome. Taken the obtained results suggest that the N-terminal amino acids 3-7 additively contribute to specific recruitment of MSL on the X chromosome.

Next, we asked whether MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 8\text{-}20}$, MSL1$^{\Delta 24\text{-}39}$, MSL1$^{\Delta 41\text{-}65}$, MSL1^{GS}, MSL1$^{\Delta 1\text{-}85}$, MSL1$^{\Delta 41\text{-}85}$ variants exhibit similar stability when expressed in Drosophila male-like S2 cells. All MSL1 variants targeted by HAx3 were co-expressed with MSL2-FLAG in S2 cells (***Figure 4A***). MSL1^{WT} was stable and effectively immunoprecipitated with MSL2 (***Figure 4B***). Similar results were obtained with MSL1$^{\Delta 8\text{-}20}$, MSL1$^{\Delta 24\text{-}39}$, MSL1$^{\Delta 41\text{-}85}$, MSL1^{2A}, MSL1^{3S} and MSL1^{GS} (***Figure 4B***). Like in Drosophila, MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 1\text{-}85}$ were unstable and expressed at very low level. These results confirm that the specific combinations of amino-acids at the N-terminus are required for stability of the MSL1 protein.

The results in S2 cells showed that MSL1$^{\Delta 41\text{-}85}$ and MSL1^{GS} can co-immunoprecipitate with MSL2, as does MSL1^{WT} (***Figure 4B***). These results suggest that mutant MSL1$^{\Delta 41\text{-}65}$ and MSL1^{GS} may be part of the MSL complex. Since S2 cells contain an endogenous MSL1 protein, it seems likely that the mutated MSL1 variants form heterodimers with wild-type one. We then asked if these mutant protein variants could interact with endogenous MSL1 for specific recruitment to the X chromosome. To distinguish between WT and MSL1 mutants, we generated transgenic line expressing MSL1$^{\Delta 41\text{-}85}$-HA tagged with the HAx3 epitope. In transgenic males

and females expressed MSL2 and MSL1 variants in heterozygous state, we observed complete co-localization between MSL2 and MSL1 variants marked by HAx3 epitope (***Figure 4C***). Thus, in the presence of wild-type MSL1, $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ variants enter the MSL complex, in which they specifically bind to the male X chromosome. These results suggest that $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ effectively dimerize with the wild-type MSL1 via the coiled-coil domain.

Two regions in the N-terminal domain are required for interaction with roX2 RNA

Our results suggest that $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ can interact with wild-type MSL1 and form the complex. Previously it was suggested that the N-terminal part of MSL1 (191 aa) contributes for interaction with roX RNAs (Muller et al., 2020). It can be assumed that the ineffectiveness of MSL1 variants in dosage compensation is associated with their inability to interact with roX RNAs. To test this hypothesis, we obtained transgenic *msl-1*$^{-}$ (*msl-1*^{L60}/ *msl-1*$^{\gamma 269}$) females expressed either of MSL1 variants ($MSL1^{WT}$-HA, $MSL1^{GS}$-HA and $MSL1^{\Delta 41\text{-}85}$-HA) and roX2 RNA. To express roX2, we used previously obtained *roX2* gene under the control of the Ubi promoter in the 86Fb site (Tikhonova et al., 2022b). The *roX2* transgene was expressed at the same level in all females expressed different MSL1 variants (***Figure 5A***).

The total extracts obtained from the 2-3-days old adult females (***Figure 5***) were used for immunoprecipitation with MSL1 antibodies. In all cases antibodies efficiently precipitate the tested MSL1 variants (***Figure 5B***). *roX2* RNA was detected (***Figure 5C***) in precipitate obtained from $MSL1^{WT}$ extract (positive control) and was not found in precipitate obtained from *msl-1*$^{-}$ extract (negative control). In the case of $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ precipitates, *roX2* RNA was not detected as in the negative control. The obtained results suggest that $MSL1^{WT}$ interacts with roX2 in contrast to the $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ mutants. These results are consistent with the hypothesis that failure to interact with roX RNAs is the primary cause of functional inactivation of $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$.

The mutants MSL1^{GS} and MSL1$^{\Delta 41\text{-}85}$ disrupt recruitment of the MSL complex to HAS

For a more complete understanding of the role played by the N-terminal region of MSL1 in the recruitment of the MSL complex to the X chromosome, we tested the binding of the MSL1, MSL2, and MSL3 proteins by chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq) with a standard protocol for chromatin fragmentation based on sonication. We obtained MSL protein profiles from two biological replicates of 3-day-old adult males and females.

As the binding profiles of the MSL complex have been well-studied in cell lines, we compared MSL1, MSL2, and MSL3 binding efficiency at the HAS (***Figure 6 – figure supplement 1***) with previously published data from S2 cells (Schauer et al., 2017; Straub et al., 2013). The ChIP-seq data from adult y^1w^{1118} (WT) males showed MSL binding to approximately 90% of the previously identified HAS from three different studies (Alekseyenko et al., 2008; Ramirez et al., 2015; Straub et al., 2008) (***Figure 6 – source data 1***). As a control, the MSL1 and MSL3 proteins bind to autosomal sites but not to HAS in females (***Figure 6 – figure supplement 1A***). These results indicate that most HAS are common between S2 cells and adult males.

Next, we compared the binding profiles of MSL1, MSL2, and MSL3 proteins obtained in *msl-1*$^{-}$; *Ubi:msl-2*^{WT}*-FLAG*/*Ubi:msl-1*^{WT} females (M1[WT]+MSL2) and *msl-1*$^{-}$; *Ubi:msl-1*^{WT}/*Ubi:msl-1*^{WT} males (M1[WT]) (Figure 6A). We found a good correlation between the obtained profiles, demonstrating that MSL2 expressed in females induces the binding of the MSL complex to the same HAS regions. We observed stronger binding of the MSL1, MSL2, and MSL3 proteins in M1[WT] (*msl-1*$^{-}$; *Ubi:msl-1*^{WT}/*Ubi:msl-1*^{WT}) males compared with those in control y^1w^{1118} males (***Figure 6A, Figure 6 – figure supplement 1B***), which might be a consequence of the higher concentrations of MSL proteins in the transgenic line due to being expressed under the control of the Ubi promoter (***Figure 1 – figure supplement 3***). In M1[WT]+MSL2 females, the MSL proteins also bind to most HAS, but with lower efficiency (***Figure 6ABC***). These results are consistent with previous studies demonstrating that the ectopic

expression of MSL2 in females is sufficient to trigger the correct targeting of MSL proteins to the X chromosome (Valsecchi et al., 2018).

To test the role played by the N-terminal region of MSL1, we compared the binding of MSL proteins in females expressing MSL2 combined with $MSL1^{WT}$, $MSL1^{\Delta 1\text{-}15}$, $MSL1^{\Delta 41\text{-}85}$, and $MSL1^{GS}$ (***Figure 7A***). As shown above, MSL2 and $MSL1^{WT}$ effectively bind to most HAS in females. MSL3 was associated with extended regions surrounding the HAS, consistent with previously reported data (Larschan et al., 2007; Sural et al., 2008). In $MSL1^{\Delta 1\text{-}15}$, $MSL1^{\Delta 41\text{-}85}$ and $MSL1^{GS}$ flies binding of MSL1, MSL2 and MSL3 to X chromosomal HAS is strongly reduced (***Figure 7A-E***, ***Figure 7 – figure supplement 1A***). At the same time, we noticed that there are HAS in which MSL2 continues to bind effectively when MSL1 in $MSL1^{\Delta 1\text{-}15}$, $MSL1^{\Delta 41\text{-}85}$ and $MSL1^{GS}$ flies is strongly reduced (***Figure 7 – figure supplement 1B***).

To investigate the observed variations in MSL protein localization at various genomic sites, we calculated the average logFC between normalized (RPKM) test signals and nonspecific IgG signals for female fly lines in two types of regions: MSL1∩MSL2 overlapping peaks and MSL1 alone peaks. Regions were further hierarchically clusterized using Ward's algorithm and subdivided with a chosen number of clusters based on the binding profiles of the MSL1, MSL2, and MSL3 proteins (***Figure 8 – source data 1***). Clusters 1 and 2 include binding sites (HAS) for the MSL complex on the X chromosome (***Figure 8A, Figure 7 – figure supplement 1, Figure 8 – figure supplement 1A***). These sites are highly enriched by the MSL1, MSL2, and MSL3 proteins in $MSL1^{WT}$ flies, suggesting the binding of the complete MSL complex. The cluster 1 (n = 133) includes sites to which all MSL proteins bind in $MSL1^{WT}$ flies. However, this binding is lost in flies expressing any of the MSL1 mutants (***Figure 8A, Figure 7 – figure supplement 1A, Figure 8 – figure supplement 1A***). Most of these sites are HAS (n = 130). It seems likely that only the MSL1-MSL2 complex can effectively bind to these sites. In cluster 2 (n = 87), the MSL2 protein binds to sites with a similar efficiency regardless of the co-expressed MSL1 variant (***Figure 8A, Figure 7 – figure supplement 1B, Figure 8 – figure supplement 1A***). At the

same time, the binding of MSL1^{GS} and MSL3 is considerably reduced, suggesting that these proteins are not efficiently associated with MSL2 at these sites. MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 41\text{-}85}$, and MSL3 are not bound to these sites, suggesting that MSL2 binds these sites alone or in cooperation with MLE. This group of sites fully corresponds to HAS.

The other three clusters include sites on all chromosomes that are bound by the MSL1 protein (***Figure 8B, Figure 8 – figure supplement 1B***). Previously, it was shown that MSL1 binds to promoters independently of the MSL complex (Chlamydas et al., 2016; Hallacli et al., 2012; Straub et al., 2013). Accordingly, MSL2 binding was weak or absent at all sites included in these clusters. MSL1^{GS} binds to all sites with the same efficiency as MSL1^{WT}, confirming that this mutant protein retains its functionality. Cluster 3 (n = 253) includes most of the MSL1 binding sites to which MSL3 does not or weakly binds (***Figure 8B, Figure 8 – figure supplement 1***). The sites are predominantly located in the promoter regions (n=199). The MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 41\text{-}85}$ are not associated with these sites. The absence of MSL1$^{\Delta 1\text{-}15}$ at this group of sites can be explained by the low stability of this mutant protein. As MSL1$^{\Delta 41\text{-}85}$ is expressed more strongly than MSL1^{GS}, it seems likely that the 41-85 aa region facilitates the interaction of MSL1 with transcription factors associated with these promoters.

The remaining two clusters (***Figure 8B***) combine sites in which the MSL1 and MSL3 proteins colocalize. The MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 41\text{-}85}$ proteins do not bind to the sites of cluster 4 (n = 34) (***Figure 8B, Figure 8 – figure supplement 1B***). MSL3 binds to sites with equal efficiency in flies expressing all MSL1 variants, suggesting that MSL3 binds independently of MSL1. At the end, in cluster 5 (n = 33), sites are combined to which the MSL1$^{\Delta 1\text{-}15}$ and MSL1$^{\Delta 41\text{-}85}$ are bound but with weaker efficiency than MSL1^{GS} or MSL1^{WT} (***Figure 8B, Figure 8 – figure supplement 1B***). Again, the binding of MSL3 is similar in flies expressed different variants of MSL1. The sites of clusters 4 and 5 are predominantly located in the promoter regions (n=30 for cluster 4, and n=19 for cluster 5).

Discussion

This study focused on examining the role played by the MSL1 protein in recruiting the DCC in *Drosophila* to HAS. By UV-XL assay, it was shown that the N-terminal 191 aa of MSL1 and MSL2 are direct roX2 interactors (Muller et al., 2020). Here, we found that the 15 N-terminal aa are critical for stability of the MSL1 protein. The substitution of the N-terminal amino-acids 3–7 (**KRFKW**) on the **SSGSG** in $MSL1^{GS}$ slightly reduces the stability of the mutant protein and strongly affects the interaction of $MSL1^{GS}$ with roX2. The 41-85 aa are also critical for MSL1 function and interaction with roX2 RNA *in vivo*. Both regions are separated by a linker that is less essential for MSL1 activity.

As functions of roX1 and roX2 are redundant (Meller and Rattner, 2002; Stuckenholz et al., 2003), it seems likely that roX1 interacts with the same N-terminal regions of MSL1 as roX2. Further study is required to understand the mechanisms for how N-terminal sequences of MSL1 determine the highly specific interaction with roX RNAs. In accordance with the key role played in the recruitment of the DCC to the male X chromosome, the N-terminal 15 amino acids of MSL1 are highly conserved among Drosophilidae. However, MSL1 orthologs in other families of Diptera have completely different N-terminal sequences despite the presence of high homology in the dimeric coiled-coil region. Even between *Anopheles gambiae* and *Drosophila melanogaster*, the X chromosome is thought to have formed independently from the same ancestral chromosome (Jiang et al., 2015; Pease and Hahn, 2012; Toups and Hahn, 2010).

The generally accepted model has suggested that most MSL complexes initially bind with numerous HAS on the male X chromosome (Kageyama et al., 2001; Kelley et al., 1997; Meller et al., 2000; Park et al., 2003; Park and Kuroda, 2001). Most HAS contain low-complexity GAGA motifs, referred to as the MRE (Alekseyenko et al., 2008; Straub et al., 2008). The CLAMP protein binds to MREs and is responsible for chromatin opening (Albig et al., 2019; Rieder et al., 2019). MSL2 contributes to the specific recognition of HAS through the direct

binding of GAGA motifs (Villa et al., 2016) and interactions with CLAMP (Albig et al., 2019; Eggers et al., 2023; Tikhonova et al., 2019). Subsequently, the MSL complexes spread from the HAS to active X chromosomal genes through distance chromatin contacts, which are formed by preexisting chromosome architecture (Prayitno et al., 2019; Schauer et al., 2017). roX RNAs guide the MSL complex assembly and its spread from HAS into flanking chromatin (Ankush Jagtap et al., 2019; Ilik et al., 2017; Lv et al., 2019; Muller et al., 2020; Park et al., 2002; Valsecchi et al., 2021; Villa et al., 2021). It was shown that expression of roX RNA on autosome in presence of MSL proteins excess can induce the local spreading of MSL complex from the site of the transgene localization (Kelley et al., 2008; Oh et al., 2003; Park et al., 2002).

Recent study has shown that low-complexity C-terminal domains of MSL2 and roX RNAs form a stably condensed state that induces local “trapping” of the MSL complex (Valsecchi et al., 2021). Experimental evidence shows that roX2 RNA plays an instructive role by modulating the MSL complex binding specificity (Villa et al., 2021). For example, Rieder et al. (Rieder et al., 2019) showed that at the initiation of embryogenesis, the DCC first binds to all chromosomes before becoming concentrated on the X chromosome. Weak interactions with chromatin sites on autosomes likely lead to the rapid dissociation of the DCC, resulting in the gradual concentration of the DCC on HAS. It was shown that complete MSL complexes preferentially interact with the sites on the X chromosome, while the MSL1 and MSL2 proteins can independently bind to many promoters located at all chromosomes (Chlamydas et al., 2016; Hallacli et al., 2012; Straub et al., 2013; Valsecchi et al., 2018).

The $MSL1^{GS}$ is the best for testing the functional role of the MSL1-roX interaction because $MSL1^{\Delta 1\text{-}15}$ is unstable and $MSL1^{\Delta 41\text{-}85}$ loses its ability to interact effectively with autosomal promoters. Thus, the N-terminal region between 41 and 85 aa is significant not only for interaction with RNA but also because plays an additional role in attracting MSL1 to active promoters independently from the MSL complex. Previous studies have shown that MSL1 alone or in combination with MSL3 or MOF also associates with many autosomal promoters

independent of MSL2 (Chlamydas et al., 2016; Straub et al., 2013). Our ChIP seq data suggest that MSL1 and MSL3 are recruited to the promoters mainly independent of each other. In contrast with the recruitment of the MSL complex to HAS on the X chromosome, the binding of MSL2 to autosomal promoters does not require the MSL1 protein (Valsecchi et al., 2018). Thus, MSL proteins can function autonomously or as a part of alternative transcription complexes.

Our results with the $MSL1^{GS}$ mutant agree with the model that the interaction of MSL1 with roX RNAs is critical for the assembly of the MSL complex and its specific recruitment on the X chromosome. $MSL1^{GS}$ effectively binds to autosomal promoters like $MSL1^{WT}$ but disrupts recruitment of the MSL complex on the X chromosomal HAS. As a result, males expressing $MSL1^{GS}$ have very low viability, as shown earlier for males that have both roX genes inactivated (*roX*$^{-}$). In *roX*$^{-}$ males, MSL1 and MSL2 bind only to the strongest regions of the X chromosome and begin to bind to autosomal regions (Meller and Rattner, 2002; Park et al., 2003; Stuckenholz et al., 2003). However, overexpression of MSL1 and MSL2 can overcome a lack of roX RNAs (Oh et al., 2003). Consequently, the co-binding of MSL1 and MSL2 to the X chromosome is partially restored, and the viability of males is increased. We also found that the overexpression of MSL2 partially restores the viability of males expressed $MSL1^{GS}$. Probably, a high concentration of MSL2 increases the likelihood of interaction with $MSL1^{GS}$ without the help of roX RNAs.

The $MSL1^{GS}$ and $MSL1^{\Delta 41\text{-}85}$ variants can efficiently heterodimerize with $MSL1^{WT}$, suggesting that efficient MSL1 dimerization does not requires roX RNAs. Thus, the interaction of MSL2 with the MSL1 dimer is a critical step in the organization of the MSL complex, which is accelerated in the presence of roX RNA.

The structure of the complex has only been determined for human MSL1 and MSL2 (Hallacli et al., 2012), for which the interaction does not require an RNA mediator (Wu et al., 2011). In contrast, mutagenic analysis showed that in Drosophila, MSL2 preferentially interacts with the MSL1 coiled-coil domain via its RING finger (Copps et al., 1998). Thus, it is possible

that interaction between MSL1 and MSL2 in Drosophila is weaker and requires roX RNAs for efficient association. Similar to the MSL1^{GS} mutant, it has also been shown (Hallacli et al., 2012) that a mutant MSL1 dimer that has lost the ability to interact with MSL2 cannot effectively bind to the X chromosome.

Generally, our results are consistent with the model (Valsecchi et al., 2021; Villa et al., 2021) according to which roX RNAs provide efficient assembly of the MSL complex, which has a high affinity for HAS on the X chromosome. This process occurs through specific interactions of roX RNAs with MSL1, MSL2, MLE, and possibly MSL3 and MOF. The inability of MSL1^{GS} to interact with roX RNA disrupts the formation of the MSL complex. Consequently, individual MSL proteins independently bind to regulatory elements (mainly promoters) or integrate into alternative transcription complexes. In general, the functions of roX RNA resemble the activity of the X inactive-specific transcript (Xist) in mammals, which specifically interacts with repression complexes involved in X chromosome inactivation (Brockdorff et al., 2020; Jacobson et al., 2022). In addition, Xist specifically targets repression complexes to the inactive X chromosome. However, for Drosophila, the significance of roX RNAs in specific MSL recruitment to the X chromosome remains an open question requiring further examination.

Materials and methods

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit anti-MSL1[423-1030]	This lab	
rabbit anti-MSL2[421-540]	This lab	
rabbit anti-MSL3	This lab	
mouse monoclonal anti-FLAG clone M2	Sigma, USA	F3165

mouse monoclonal anti-HA clone HA-7	Sigma, USA	H9658
anti-lamin Dm0 clone ADL84.12	DSHB, USA	ADL84.12
anti-FLAG-HRP antibodies	Sigma, USA	A8592
goat anti-mouse Alexa Fluor 555	Thermo Fisher Scientific, USA	A28180
goat anti-rabbit Alexa Fluor 488	Thermo Fisher Scientific, USA	A11008
Chemicals, peptides, and recombinant proteins		
Pierce™ 16% Formaldehyde (w/v), Methanol-free	Thermo Fisher Scientific, USA	28906
CNBr-activated Sepharose	Cytiva, UK	GE17-0430-01
Aminolink Plus Coupling Resin	Thermo Fisher Scientific, USA	20505
TRI-reagent	MRC, USA	TR 118
Ribonucleoside Vanadyl Complex	NEB, USA	S1402S
Nuclease-free BSA	Sigma, USA	126609
Qubit dsDNA HS Assay Kit	Life Technologies Corporation	Q32851
Phase Lock Gel™, QuantaBio - 2302830, Phase Lock Gel Heavy	VMR, USA	10847-802
NEBNext Ultra II DNA Library Prep Kit for Illumina	NEB, USA	E7645L
Ampure Xp beads	Beckman Coulter, USA	A63881
Diamant TaqA Hot-start polymerase	Belbiolab, Russia	E-TAP
Dynabeads Protein A	Thermo Fisher Scientific, USA	10008D
DAPI	Applichem, Germany	A1001
MACSFectin	Miltenyi Biotec, USA	130-098-410
SFX-Insect Cell Culture Media	HyClone, USA	SH30278

anti-HA magnetic beads	Thermo Fisher Scientific, USA	88836
anti-FLAG magnetic beads	Sigma, USA	M8823
mouse IgG magnetic beads	NEB, USA	S1431
Halt Protease Inhibitor Cocktail	Thermo Fisher Scientific, USA	78438
Calbiochem Complete Protease Inhibitor Cocktails V	Merck, USA	539137
Calbiochem Complete Protease Inhibitor Cocktails VII	Merck, USA	539138
Crescendo Western Blotting substrate	Merck, USA	WBLUR0100
ChIP DNA Clean&Concentrator kit	Zymo Research, USA	D5205
Experimental models: Organisms/strains		
P{ry[+t7.2]=hsp70-FLP}1, y[1] w[*]; M{3xP3-RFP.attP}ZH-86Fb; M{RFP[3xP3.PB] GFP[E.3xP3]=vas-int.B}ZH-102D	Bloomington Drosophila Stock Center	23648
w[1118]; sna[Sco]/CyO, P[ActGFP.w-]CC2	Bloomington Drosophila Stock Center	9325
y[1] w[1118]	Bloomington Drosophila Stock Center	6598
Ubi:msl-2WT-FLAG	Tikhonova et al., 2019	
msl-1L60 /CyO	donated by M. Kuroda	
msl-1γ269 cn1bw1/CyO	donated by J. Lucchesi	
*Ubi-msl-1** transgenes are described in Materials and Methods	This work	

S2 cells	Drosophila genomics resource center	DGRC Stock 181 ; https://dgrc.bio.indiana.edu//stock/181 ; RRID:CVCL_Z992
Oligonucleotides		
Oligonucleotides used are listed in "Supplement table 1"	Evrogen, Lytech, DNA synthesis	N/A
Software and algorithms		
ImageJ 1.54f and Fiji bundle 2.14.0	Schindelin et al., 2012	fiji.sc
cutadapt	Martin, 2011	
Bowtie version 2	Langmead and Salzberg, 2012	
deepTools	Ramirez et al., 2014	
Picard	http://broadinstitute.github.io/picard/	
MACS version 2	Zhang et al., 2008	
R version 4.2.1	http://www.r-project.org	
ChIPpeakAnno package	Zhu et al., 2010	
ChIPseeker package	Yu et al., 2015	
pheatmap package.	https://github.com/raivokolde/pheatmap	
Gviz	Helaers et al., 2011	
Other		
Nikon Eclipse Ti fluorescent microscope with Nikon DS-Qi2 digital camera	Nikon, Japan	
Focused-ultrasonicator Chemidoc MP Imaging system	Covaris Bio-Rad	Covaris ME220, 500506 12003154

Real-Time PCR System	Thermo Fisher Scientific	StepOne Plus, Quantstudio 6Flex
Bioruptor	Diagenode, USA	B01020001
100 µm MACS SmartStrainers	Miltenyi Biotec, USA	130-110-917
70 µm MACS SmartStrainers	Miltenyi Biotec, USA	130-110-916
Bioanalyzer 2100	Agilent, USA	G2939B

Plasmid construction

PCR directed mutagenesis was used to make constructs with deletion variants of MSL1 (corresponding primers are given in Supplement table 1). Different full-sized variants of MSL1 were cloned into an expression vector containing *attB* site for φC31-mediated recombination, *Ubi-p63E* promoter with its 5'UTR, 3'UTR with SV40 polyadenylation signal, intronless *yellow* gene as a reporter for detection of transformants (***Figure 1 – figure supplement 2***).

Fly crosses and transgenic lines

Drosophila strains were grown at 25°C and 75% humidity under standard culture conditions. The transgenic constructs were injected into preblastoderm embryos using the φC31-mediated site-specific integration system at locus 86Fb (Bischof et al., 2007). The emerging adults were crossed with the *y ac* w^{1118} flies, and the progeny carrying the transgene in the 86Fb region were identified by y^{+} pigmented cuticle.

The *msl-1*$^{-}$ background corresponds to *msl-1*$^{\gamma 269}$/*msl-1*^{L60} heterozygote. The *msl-1*^{L60} */CyO* (Chang and Kuroda, 1998) and *msl-1*$^{\gamma 269}$ $cn^{1}bw^{1}$*/CyO* (Palmer et al., 1993) stocks with the null *msl-1* mutations were kindly donated by M. Kuroda and J. Lucchesi. The *w[1118]; sna[Sco]/CyO, P[ActGFP.w*$^{-}$*]CC2* stock no. 9325 was obtained from the Bloomington stock center. To access the viability of males carrying MSL1 variants (MSL1*, MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 41\text{-}}$

^{85}, and MSL1$^{\Delta 1\text{-}85}$), 2-3 days old virgin *msl-1^{L60}/CyO*, GFP; *Ubi-msl-1*/ Ubi-msl-1** females were crossed with 2-3 days old *msl-1$^{\gamma 269}$/CyO*, GFP; *Ubi-msl-1*/Ubi-msl-1** males. *Ubi-msl-1** is any transgene expressing one of the tested MSL1 variants. *msl-1^{L60}/msl-1$^{\gamma 269}$*; *Ubi-msl-1*/Ubi-msl-1** males and *msl-1^{L60}/msl-1$^{\gamma 269}$*; *Ubi-msl-1*/Ubi-msl-1** females were counted after hatching. The percentage of male viability with different *Ubi-msl-1** transgenes was estimated by taking the viability of females as 100%. The viability of males carrying MSL1 variants (MSL1*, MSL1^{WT}, MSL1$^{\Delta 66\text{-}85}$ and MSL1$^{\Delta 24\text{-}39}$) was tested in *msl-1^{L60}/ msl-1$^{\gamma 269}$*; *Ubi-msl-1*/Ubi-msl-1** lines.

To assess viability of females expressed MSL2 and one of MSL1 variants (MSL*), 2-3 days old virgin *Ubi-msl-1*/Ubi-msl-1** females were crossed with 2-3 days old *Ubi:msl-2^{WT}-FLAG/TM6,T* viability was estimated by taking the viability of *Ubi-msl-1*/TM6,Tb* females as 100%.

To assess role of MSL1 variants in binding of MSL complex to polytene chromosomes, 2-3 days old virgin *msl-1^{L60}/msl-1$^{\gamma 269}$*; *Ubi-msl-1*/Ubi-msl-1** females were crossed with 2-3 days old *msl-1^{L60}/CyO, GFP; Ubi:msl-2^{WT}-FLAG/ TM6,Tb* males. *msl-1^{L60}/msl-1$^{\gamma 269}$*; *Ubi-msl-1*/ Ubi:msl-2^{WT}-FLAG* larvae were selected for further study.

b males. *Ubi:msl-2^{WT}-FLAG* is transgene (86Fb) expressing the MSL2 cDNA under the control of the Ubi-p63E promoter at the 86Fb site (Tikhonova et al., 2019). The percentage of *Ubi:msl-2^{WT}-FLAG /Ubi-msl-1** female

Protein extract preparation

Twenty adult flies were cooled in liquid nitrogen, homogenized for 30 sec with a pestle in 200 µL of extraction buffer (20 mM HEPES, pH 7.5, 100 mM KCl, 5% glycerol, 10 mM EDTA, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 1 mM DTT, 5 mM PMSF, and 1:100 Calbiochem Complete Protease Inhibitor Cocktails VII and V) and incubated on ice for 10 min. The suspension was sonicated in a Bioruptor (Diagenode, USA) for 3 min on setting H, 15 sec ON/45

sec OFF. Then 4x SDS-PAGE sample buffer was added to the homogenate. Extracts were incubated for 10 min at 100°C, centrifuged at 16000 x g for 5 min, and loaded on a 6% SDS-PAGE gel.

Immunostaining of polytene chromosomes

Drosophila 3rd instar larvae were cultured at 18°C under standard conditions. Western blotting showed that the expression of MSL1 and MSL2 under control of the *Ubi-p63E* promoter is significantly increased at 18°C compared to 25°C (Figure S2B). Polytene chromosome staining was performed as described (Murawska and Brehm, 2012). The following primary antibodies were used: rabbit anti-MSL1 at 1:500 dilution, rabbit anti-MSL2 at 1:500 dilution, mouse anti-HA at 1:50 dilution, mouse anti-FLAG at 1:100 dilution. The secondary antibodies were Alexa Fluor 488 or 555 goat anti-rabbit and anti-mouse, at 1:2,000 dilution (Invitrogen). The polytene chromosomes were co-stained with DAPI (AppliChem). Images were acquired on the Nikon Eclipse T*i* fluorescent microscope using Nikon DS-Qi2 digital camera, processed with ImageJ 1.54f and Fiji bundle 2.14.0. 3-4 independent staining and 4-5 samples of polytene chromosomes were performed with each MSL1-expressing transgenic line. Images of stretched chromosomes are obtained as follows: based on the DAPI channel, a chromosome trace is created using the "Segmented line" tool, the thickness is selected ('Edit-Selection-Properties'), saved as 'Selection' in .roi format. This file is opened on top of the desired channel image, and the 'Edit-Selection-Straighten' commands are applied.

Co-immunoprecipitation assay

Drosophila S2 cells grown in SFX medium (HyClone) at 25°C were co-transfected by plasmid expressing MSL1*-3xHA and plasmid expressing MSL2-3xFLAG with MACSFectin (Miltenyi Biotec), as recommended by the manufacturer. After transfection, the cells were incubated for 48 h, washed once with cold 1×PBS, and resuspended in 20 packed cell volumes of hypotonic lysis

buffer (20 mM Tris-HCl, pH 7.4, with 10 mM KCl, 10 mM MgCl2, 2 mM EDTA, 10% glycerol, 1% Triton X-100, 1 mM DTT, 0.5 mM PMSF and Halt Protease Inhibitor Cocktail). After incubation on ice for 10 min, the cells were sonicated in a Bioruptor (Diagenode, USA) for 2 min on setting L, 15 sec ON/45 sec OFF, NaCl was added to a final concentration of 420 mM, and incubation on ice continued for 60 min, with periodic mixing. Sonication was repeated as above to reduce viscosity, cell debris was pelleted by centrifugation at 10,000 g for 30 min at 4°C, and the supernatant was collected for immunoprecipitation with anti-FLAG, anti-HA and mouse IgG magnetic beads equilibrated in incubation buffer-150 (20 mM Tris-HCl, pH 7.4, with 150 mM NaCl, 10 mM MgCl2, 1 mM EDTA, 1 mM EGTA, 10% glycerol, and 0.1% NP-40). The protein extract (50 μg protein) was adjusted to a volume of 200 μL with buffer-150, mixed with magnetic beads (15 μL), and incubated on a rotary shaker overnight at 4°C. The beads were then washed with five portions of buffer-150, resuspended in SDS-PAGE loading buffer, boiled, and analyzed by immunoblotting with anti-FLAG-HRP antibodies (Sigma) antibodies. Proteins were detected using the Crescendo Western Blotting substrate (Millipore).

RNA Immunoprecipitation assay

To prepare the samples, 100 mg of adult flies were ground in a mortar in liquid nitrogen and resuspended in 0.5 mL of VNC-hypotonic lysis buffer (20 mM Tris-HCl, pH 7.4, with 10 mM KCl, 10 mM MgCl2, 10% glycerol, 1% Triton X-100, 1 mM DTT, 0.5 mM PMSF, 2 mM Ribonucleoside Vanadyl Complex, and 1xHalt Protease Inhibitor Cocktail). The suspension was homogenized in a Dounce (loose pestle, 20 strokes) homogenizers and filtered through 100 μm MACS SmartStrainers (Miltenyi Biotec, United States). After incubation on ice for 10 min, the suspension was sonicated in a Bioruptor (Diagenode, USA) for 3 min on setting L (15 sec ON/45 sec OFF). NaCl was added to a final concentration of 420 mM, and the suspension was incubated on ice for an additional 60 minutes with periodic mixing. Sonication was repeated as above to reduce viscosity. Cell debris was pelleted by centrifugation at 10,000 g for 30 min at

4°C, and the supernatant was collected for subsequent immunoprecipitation. The RNA-protein extracts were diluted to 150 mM of NaCl with VNC-incubation buffer-0 (20 mM Tris-HCl, pH 7.4, 10 mM MgCl2, 10% glycerol, 0.1% NP-40, 0.5 mM PMSF, 2 mM Ribonucleoside Vanadyl Complex, and 1xHalt Protease Inhibitor Cocktail). Rabbit anti-MSL1 (1:200) and non-specific IgG were incubated for 1 hour at room temperature with 20 µL aliquots of Protein A Dynabeads (Thermo Fisher, USA) mixed with 200 µL of PBST. The antibody–Dynabead complexes were then washed and equilibrated in VNC-incubation buffer-150 (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10 mM MgCl2, 10% glycerol, 0.1% NP-40, 0.5 mM PMSF, 2 mM Ribonucleoside Vanadyl Complex, and 1xHalt Protease Inhibitor Cocktail). The RNA-protein extracts (10% aliquots were stored as input material for subsequent protein and RNA extractions) were mixed with magnetic beads and incubated on a rotary shaker overnight at 4°C. The beads were then washed with four portions of VNC-incubation buffer-150 (10% aliquots of supernatants after immunoprecipitation were stored as output material for subsequent protein and RNA extractions). Half of the beads were resuspended in SDS-PAGE loading buffer, boiled, and analysed by immunoblotting with anti-HA-HRP antibodies (Thermo Fisher, USA). Proteins were detected using the Crescendo Western Blotting substrate (Millipore). The other half of the beads were resuspended in the TRI reagent for subsequent RNA extraction and cDNA synthesis.

RT-PCR

Total RNA was isolated using the TRI reagent (Molecular Research Center, United States) according to the manufacturer's instructions. RNA was treated with two units of Turbo DNase I (Ambion) for 30 min at 37°C to eliminate genomic DNA. The synthesis of cDNA was performed according to the manufacturer's instructions using 1 µg of RNA or entire sample volume after immunoprecipitation, 100 U of EpiScript reverse transcriptase (LGC Biosearch Technologies, UK), and 10 pM of random hexamers as a primer. The amounts of specific cDNA fragments were quantified by real-time PCR. At least three independent measurements were made for each

RNA sample. Relative levels of mRNA expression were calculated in the linear amplification range by calibration to a standard genomic DNA curve to account for differences in primer efficiencies. Individual expression values were normalized with reference to *RpL32* mRNA.

ChIP-seq

Chromatin was prepared from two- to three-day-old adult flies. Samples of 500 mg each of adult flies were ground in a mortar in liquid nitrogen and resuspended in 10 mL of buffer A (15 mM HEPES-KOH, pH 7.6, 60 mM KCl, 15 mM NaCl, 13 mM EDTA, 0.1 mM EGTA, 0.15 mM spermine, 0.5 mM spermidine, 0.5% NP-40, 0.5 mM DTT) supplemented with 0.5 mM PMSF and Calbiochem Complete Protease Inhibitor Cocktail V. The suspension was then homogenized in a Potter (10 strokes) and subsequently in a Dounce (tight pestle, 15 strokes) homogenizers and filtered through 70 µm MACS SmartStrainers (Miltenyi Biotec, United States). The homogenate was cross-linked with 1% formaldehyde for 15 min at room temperature. Cross-linking was stopped by adding glycine to a final concentration of 125 mM. The nuclei were washed with three 10 mL portions of wash buffer (15 mM HEPES-KOH, pH 7.6, 60 mM KCl, 15 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 0.1% NP-40, protease inhibitors) and one 5-mL portion of nuclear lysis basic buffer (15 mM HEPES, pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 1% Triton X-100, 0.5 mM DTT, 0.1% sodium deoxycholate, and protease inhibitors), and resuspended in 1 mL of nuclear lysis buffer (15 mM HEPES, pH 7.6, 140 mM NaCl, 1 mM EDTA, 0.1 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.5% SLS, 0.1% SDS, 0.5 mM PMSF, Calbiochem Complete Protease Inhibitor Cocktail V). The suspension was incubated for 30 min at 4°C and then sonicated with Covaris ME220 focused-ultrasonicator (50 alternating 15-s ON and 45-s OFF intervals, at 6°C, peak power 75.0, duty % factor 25, cycles/burst 1000), and 50 µL aliquots were used to test the extent of sonication and to measure DNA concentration. Chromatin was then transferred to 1.5 ml-eppendorf tubes and centrifuged at 14000 rpm for 10 min at 4 °C. Finally, supernatants were pooled, aliquoted and frozen in liquid nitrogen.

For immunoprecipitation chromatin was pre-cleared with Protein A Dynabeads (Thermo Fisher, USA), blocked with 0.5% BSA. Corresponding antibodies were incubated for 1 hour at room temperature with 20 µL aliquots of Protein A (rabbit anti-MSL1, 1:200; anti-MSL2, 1:100; anti-MSL3, 1:200; rabbit non-specific IgG) mixed with 200 µL of PBST. Then antibody–Dynabead complexes were washed and equilibrated in nuclear lysis buffer. Chromatin samples containing 10–20 µg of DNA equivalent in 200 µL nuclear lysis buffer (2 µL aliquots of pre-cleared chromatin as input material) were incubated overnight at 4°C with antibody–Dynabead complexes. After 3 rounds of washing with lysis buffer supplemented with 500 mM NaCl and TE buffer (10 mM Tris-HCl, pH 8; 1 mM EDTA), the DNA was eluted with elution buffer (50 mM Tris-HCl, pH 8.0; 1 mM EDTA, 1% SDS), treated with RNase A for 30 min and proteinase K overnight, incubated at 65 °C for 6 hours and extracted with ChIP DNA Clean&Concentrator kit (Zymo Research, USA).

The ChIP-seq libraries were prepared with NEBNext® Ultra™ II DNA Library Prep kit, as described in the manufacturer's instructions. In short, eluted DNA was end-repaired and terminal adenosine residues were added using the NEBNext reagents. Indexed adapters were ligated, after which the material was size selected at ~200-600 bp with Ampure XP beads (Beckman Coulter). PCR amplification was performed using NEB primers for 15-16 cycles using the Q5 Hot Start HiFi PCR Master Mix (NEB). The PCR-amplified library was purified using Ampure XP beads and its quality was assessed on a Bioanalyzer 2100 system (Agilent). Diluted libraries were clustered on a pair-read flowcell and sequenced using a NovaSeq 6000 system (Illumina) (using «Genetico» facility).

ChIP-seq data processing and sequence analysis

All ChIP-seq raw data were presented as two biological replicates with paired-end reads (except MSL1 M1[wt]♂), for which only one biological replicate was obtained). Trimming and mapping were performed using cutadapt software (Martin, 2011) and Bowtie version 2 (Langmead and

Salzberg, 2012), as described previously (Sabirov et al., 2021). The dm6 version of the Drosophila melanogaster genome was used as a reference genome. After merging replicates, coverage tracks (BedGraph) were obtained using the deepTools (Ramirez et al., 2014) bamCoverage function, with bin widths of 10 bp and extendReads option and normalized by reads per kilobase of transcript, per million mapped reads (RPKM). Raw and processed data were deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE243396 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE243396; secure token for reviewer: ebalmcwmdbarjkv).

Peak calling was performed for MSL proteins from flyes of ♂M1[wt] and ♀ M1[wt]+MSL2 lines, as described previously (Sabirov et al., 2021). Briefly:

1) duplicates were removed with Picard (http://broadinstitute.github.io/picard/);

2) blacklist filtration was performed;

3) peak calling was performed using MACS version 2 against nonspecific IgG (Zhang et al., 2008), with a soft p-value threshold of 1×10^{-2} for further irreproducible discovery rate (IDR) analysis.

Reproducibility was assessed using the IDR pipeline, with p-value thresholds of 0.05 for true replicates and 0.01 for pseudoreplicates. All samples showed ideal reproducibility [both the rescue ratio (RR) and self-consistency ratio (SR) were less than 2; thus, an “optimal” set of highly-reproduced peaks was chosen for each sample.

Further analysis was performed in R version 4.2.1 (http://www.r-project.org). Colocalization analysis was performed using the ChIPpeakAnno package (Zhu et al., 2010). Average signal calculation and heatmaps were constructed using the ChIPseeker package (Yu et al., 2015).

HAS regions for which ChIP-seq signal enrichment was calculated were obtained by combining HAS variants from three different studies (Alekseyenko et al., 2008; Ramirez et al., 2015; Straub et al., 2008). Only HAS regions located on the X chromosome were considered in this analysis, resulting in 301 HAS regions (***Figure 6 – source data 1***). For ***Figure 6A*** and ***Figure 6 – figure***

supplenet 1A visualization 10% autosomal regions from ♂M1[wt] with highest log p-value were chosen for every MSL protein (MSL1 - 671, MSL2 - 288, MSL3 - 185 peaks).

To investigate MSL binding for female fly lines at different genomic locations, a list of unique MSL1, MSL2 peaks for ♀ M1[wt]+MSL2 fly line was generated using findOverlapsOfPeaks function (maxgap = 500). This list was subdivided into two groups, according to the mean logFC between protein signal and control at 500 bp region around the center of each peak: 'MSL1 alone' (MSL1 > threshold, MSL2 < threshold, threshold=0.2) and 'MSL1 and MSL2 overlapping' (MSL1 > threshold, MSL2 > threshold, threshold=0.4). For 'MSL1 alone' only peaks not overlapping with HAS were chosen (AUT+X) and for 'MSL1 MSL2 overlapping' X-associated peaks were selected (HAS+X). Regions were hierarchically clustered using hclust (method= 'ward.D2') and subdivided into a chosen number of clusters (***Figure 8 – source data 1***) using cutree function from package stats. Cluster visualization was made with the pheatmap package (https://github.com/raivokolde/pheatmap). Gviz (Helaers et al., 2011) were used for genomic track visualization.

Acknowledgments

We thank Mitzi Kuroda (Harvard Medical School), and Erica Larschan (Brown University) for kindly providing fly strains. We thank Farhod Hasanov and Aleksander Parshikov for fly injection. This study was performed using the equipment of the Center for Precision Genome Editing and Genetic Technologies for Biomedicine of IGB RAS supported by the Ministry of Science and Education of the Russian Federation. Some cartoons were created with BioRender.com.

Funding

This work was supported by the grant 21-14-00211 from Russian Science Foundation. Chip-Seq analysis was supported by grant 075-15-2019-1661 from the Ministry of Science and Higher

Education of the Russian Federation.

Funding for open access charge: Russian Science Foundation and Ministry of Science and Education of the Russian Federation.

Author contributions

Valentin Babosha, Conceptualization, Data curation, Formal analysis, Validation, Investigation; Natalia Klimenko, Data curation, Formal analysis, Investigation; Anastasia Revel-Muroz, Data curation, Formal analysis, Investigation; Evgeniya Tikhonova, Data curation, Investigation; Pavel Georgiev, Conceptualization, Writing - review and editing; Oksana Maksimenko, Conceptualization, Formal analysis, Supervision, Funding acquisition, Investigation, Resources, Writing - original draft, Writing - review and editing

Data availability

NGS data have been deposited at the NCBI Gene Expression Omnibus (GEO) under accession number GSE243396.

References

Albig C, Tikhonova E, Krause S, Maksimenko O, Regnard C, Becker PB. 2019. Factor cooperation for chromosome discrimination in Drosophila. *Nucleic Acids Research* **47**:1706–1724. doi:10.1093/nar/gky1238

Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ, Kuroda MI. 2008. A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. *Cell* **134**:599–609. doi:10.1016/j.cell.2008.06.033

Ankush Jagtap PK, Muller M, Masiewicz P, von Bulow S, Hollmann NM, Chen PC, Simon B, Thomae AW, Becker PB, 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* **47**:4319–4333. doi:10.1093/nar/gkz125

Beckmann K, Grskovic M, Gebauer F, Hentze MW. 2005. A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. *Cell* **122**:529–40. doi:10.1016/j.cell.2005.06.011

Bischof J, Maeda RK, Hediger M, Karch F, Basler K. 2007. An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proceedings of the National Academy of Sciences of the United States of America* **104**:3312–7. doi:10.1073/pnas.0611511104

Brockdorff N, Bowness JS, Wei G. 2020. Progress toward understanding chromosome silencing by Xist RNA. *Genes Dev* **34**:733–744. doi:10.1101/gad.337196.120

Chang KA, Kuroda MI. 1998. Modulation of MSL1 abundance in female Drosophila contributes to the sex specificity of dosage compensation. *Genetics* **150**:699–709.

Chlamydas S, Holz H, Samata M, Chelmicki T, Georgiev P, Pelechano V, Dundar F, Dasmeh P, Mittler G, Cadete FT, Ramirez F, Conrad T, Wei W, Raja S, Manke T, Luscombe NM, Steinmetz LM, Akhtar A. 2016. Functional interplay between MSL1 and CDK7 controls RNA polymerase II Ser5 phosphorylation. *Nature structural & molecular biology* **23**:580–9. doi:10.1038/nsmb.3233

Copps K, Richman R, Lyman LM, Chang KA, Rampersad-Ammons J, Kuroda MI. 1998. Complex formation by the Drosophila MSL proteins: role of the MSL2 RING finger in protein complex assembly. *The EMBO journal* **17**:5409–17. doi:10.1093/emboj/17.18.5409

Dahlsveen IK, Gilfillan GD, Shelest VI, Lamm R, Becker PB. 2006. Targeting determinants of dosage compensation in Drosophila. *Plos Genetics* **2**:e5. doi:10.1371/journal.pgen.0020005

Demakova OV, Kotlikova IV, Gordadze PR, Alekseyenko AA, Kuroda MI, Zhimulev IF. 2003. The MSL complex levels are critical for its correct targeting to the chromosomes in Drosophila melanogaster. *Chromosoma* **112**:103–15. doi:10.1007/s00412-003-0249-1

Eggers N, Gkountromichos F, Krause S, Campos-Sparr A, Becker PB. 2023. Physical interaction between MSL2 and CLAMP assures direct cooperativity and prevents competition at composite binding sites. *Nucleic Acids Research* gkad680. doi:10.1093/nar/gkad680

Fauth T, Muller-Planitz F, Konig C, Straub T, Becker PB. 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**:3209–21. doi:10.1093/nar/gkq026

Ferrari F, Alekseyenko AA, Park PJ, Kuroda MI. 2014. Transcriptional control of a whole chromosome: emerging models for dosage compensation. *Nature structural & molecular biology* **21**:118–25. doi:10.1038/nsmb.2763

Gorman M, Kuroda MI, Baker BS. 1993. Regulation of the sex-specific binding of the maleless dosage compensation protein to the male X chromosome in Drosophila. *Cell* **72**:39–49. doi:10.1016/0092-8674(93)90048-u

Hallacli E, Lipp M, Georgiev P, Spielman C, Cusack S, Akhtar A, Kadlec J. 2012. Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in Drosophila. *Molecular Cell* **48**:587–600. doi:10.1016/j.molcel.2012.09.014

Helaers R, Bareke E, De Meulder B, Pierre M, Depiereux S, Habra N, Depiereux E. 2011. gViz, a novel tool for the visualization of co-expression networks. *BMC Res Notes* **4**:452. doi:10.1186/1756-0500-4-452

Ilik IA, Maticzka D, Georgiev P, Gutierrez NM, Backofen R, 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. doi:10.1101/gad.304600.117

Jacobson EC, Pandya-Jones A, Plath K. 2022. A lifelong duty: how Xist maintains the inactive X chromosome. *Current Opinion in Genetics & Development* **75**:101927. doi:10.1016/j.gde.2022.101927

Jiang X, Biedler JK, Qi Y, Hall AB, Tu Z. 2015. Complete Dosage Compensation in Anopheles stephensi and the Evolution of Sex-Biased Genes in Mosquitoes. *Genome biology and evolution* **7**:1914–24. doi:10.1093/gbe/evv115

Jordan W, Larschan E. 2021. The zinc finger protein CLAMP promotes long-range chromatin interactions that mediate dosage compensation of the Drosophila male X-chromosome. *Epigenetics & Chromatin* **14**:29. doi:10.1186/s13072-021-00399-3

Kadlec J, Hallacli E, Lipp M, Holz H, Sanchez-Weatherby J, Cusack S, Akhtar A. 2011. Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nature structural & molecular biology* **18**:142–9. doi:10.1038/nsmb.1960

Kageyama Y, Mengus G, Gilfillan G, Kennedy HG, Stuckenholz C, Kelley RL, Becker PB, Kuroda MI. 2001. Association and spreading of the Drosophila dosage compensation complex from a discrete roX1 chromatin entry site. *The EMBO journal* **20**:2236–45. doi:10.1093/emboj/20.9.2236

Kelley RL, Lee O-K, Shim Y-K. 2008. Transcription rate of noncoding roX1 RNA controls local spreading of the Drosophila MSL chromatin remodeling complex. *Mechanisms of Development* **125**:1009–1019. doi:10.1016/j.mod.2008.08.003

Kelley RL, Wang J, Bell L, Kuroda MI. 1997. Sex lethal controls dosage compensation in Drosophila by a non-splicing mechanism. *Nature* **387**:195–9. doi:10.1038/387195a0

Kim M, Faucillion ML, Larsson J. 2018. RNA-on-X 1 and 2 in Drosophila melanogaster fulfill separate functions in dosage compensation. *Plos Genetics* **14**:e1007842. doi:10.1371/journal.pgen.1007842

Kuroda MI, Hilfiker A, Lucchesi JC. 2016. Dosage Compensation in Drosophila-a Model for the Coordinate Regulation of Transcription. *Genetics* **204**:435–450. doi:10.1534/genetics.115.185108

Kuzu G, Kaye EG, Chery J, Siggers T, Yang L, Dobson JR, Boor S, Bliss J, Liu W, Jogl G, Rohs R, Singh ND, Bulyk ML, Tolstorukov MY, Larschan E. 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. doi:10.1371/journal.pgen.1006120

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**:357–9. doi:10.1038/nmeth.1923

Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ, Kuroda MI. 2007. MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Molecular Cell* **28**:121–33. doi:10.1016/j.molcel.2007.08.011

Larschan E, Soruco MM, Lee OK, Peng S, Bishop E, Chery J, Goebel K, Feng J, Park PJ, Kuroda MI. 2012. Identification of chromatin-associated regulators of MSL complex targeting in Drosophila dosage compensation. *Plos Genetics* **8**:e1002830. doi:10.1371/journal.pgen.1002830

Li F, Parry DA, Scott MJ. 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**:8913–24. doi:10.1128/MCB.25.20.8913-8924.2005

Li F, Schiemann AH, Scott MJ. 2008. Incorporation of the noncoding roX RNAs alters the chromatin-binding specificity of the Drosophila MSL1/MSL2 complex. *Molecular and Cellular Biology* **28**:1252–64. doi:10.1128/MCB.00910-07

Lucchesi JC. 2018. Transcriptional modulation of entire chromosomes: dosage compensation. *J Genet* **97**:357–364.

Lv M, Yao Y, Li F, Xu L, Yang L, Gong Q, Xu YZ, Shi Y, Fan YJ, 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* **47**:3142–3157. doi:10.1093/nar/gky1308

Lyman LM, Copps K, Rastelli L, Kelley RL, Kuroda MI. 1997. Drosophila male-specific lethal-2 protein: structure/function analysis and dependence on MSL-1 for chromosome association. *Genetics* **147**:1743–53.

Maenner S, Muller M, Frohlich J, Langer D, Becker PB. 2013. ATP-dependent roX RNA remodeling by the helicase maleless enables specific association of MSL proteins. *Molecular Cell* **51**:174–84. doi:10.1016/j.molcel.2013.06.011

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**:10–12.

Meller VH, Gordadze PR, Park Y, Chu X, Stuckenholz C, Kelley RL, Kuroda MI. 2000. Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. *Current biology□: CB* **10**:136–43. doi:10.1016/s0960-9822(00)00311-0

Meller VH, Rattner BP. 2002. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *The EMBO journal* **21**:1084–91. doi:10.1093/emboj/21.5.1084

Morra R, Yokoyama R, Ling H, Lucchesi JC. 2011. Role of the ATPase/helicase maleless (MLE) in the assembly, targeting, spreading and function of the male-specific lethal (MSL) complex of Drosophila. *Epigenetics & chromatin* **4**:6. doi:10.1186/1756-8935-4-6

Muller M, Schauer T, Krause S, Villa R, Thomae AW, Becker PB. 2020. Two-step mechanism for selective incorporation of lncRNA into a chromatin modifier. *Nucleic Acids Research* **48**:7483–7501. doi:10.1093/nar/gkaa492

Murawska M, Brehm A. 2012. Immunostaining of Drosophila polytene chromosomes to investigate recruitment of chromatin-binding proteins. *Methods Mol Biol* **809**:267–77. doi:10.1007/978-1-61779-376-9_18

Oh H, Park Y, Kuroda MI. 2003. Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. *Genes & development* **17**:1334–9. doi:10.1101/gad.1082003

Palmer MJ, Mergner VA, Richman R, Manning JE, Kuroda MI, Lucchesi JC. 1993. The male-specific lethal-one (msl-1) gene of Drosophila melanogaster encodes a novel protein that associates with the X chromosome in males. *Genetics* **134**:545–57.

Palmer MJ, Richman R, Richter L, Kuroda MI. 1994. Sex-specific regulation of the male-specific lethal-1 dosage compensation gene in Drosophila. *Genes & development* **8**:698–706. doi:10.1101/gad.8.6.698

Park Y, Kelley RL, Oh H, Kuroda MI, Meller VH. 2002. Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. *Science* **298**:1620–3. doi:10.1126/science.1076686

Park Y, Kuroda MI. 2001. Epigenetic aspects of X-chromosome dosage compensation. *Science* **293**:1083–5. doi:10.1126/science.1063073

Park Y, Mengus G, Bai X, Kageyama Y, Meller VH, Becker PB, Kuroda MI. 2003. Sequence-specific targeting of Drosophila roX genes by the MSL dosage compensation complex. *Molecular Cell* **11**:977–86.

Pease JB, Hahn MW. 2012. Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Molecular biology and evolution* **29**:1645–53. doi:10.1093/molbev/mss010

Prayitno K, Schauer T, Regnard C, Becker PB. 2019. Progressive dosage compensation during Drosophila embryogenesis is reflected by gene arrangement. *EMBO reports* **20**:e48138. doi:10.15252/embr.201948138

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research* **42**:W187-91. doi:10.1093/nar/gku365

Ramirez F, Lingg T, Toscano S, Lam KC, Georgiev P, Chung HR, Lajoie BR, de Wit E, Zhan Y, de Laat W, Dekker J, Manke T, Akhtar A. 2015. High-Affinity Sites Form an Interaction Network to Facilitate Spreading of the MSL Complex across the X Chromosome in Drosophila. *Molecular Cell* **60**:146–62. doi:10.1016/j.molcel.2015.08.024

Rieder LE, Jordan WT, Larschan EN. 2019. Targeting of the Dosage-Compensated Male X-Chromosome during Early Drosophila Development. *Cell reports* **29**:4268-4275 e2. doi:10.1016/j.celrep.2019.11.095

Sabirov M, Kyrchanova O, Pokholkova GV, Bonchuk A, Klimenko N, Belova E, Zhimulev IF, Maksimenko O, Georgiev P. 2021. Mechanism and functional role of the interaction between CP190 and the architectural protein Pita in Drosophila melanogaster. *Epigenetics & chromatin* **14**:16. doi:10.1186/s13072-021-00391-x

Samata M, 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. doi:10.1146/annurev-biochem-062917-011816

Schauer T, Ghavi-Helm Y, Sexton T, Albig C, Regnard C, Cavalli G, Furlong EE, Becker PB. 2017. Chromosome topology guides the Drosophila Dosage Compensation Complex for target gene activation. *EMBO reports*. doi:10.15252/embr.201744292

Scott MJ, Pan LL, Cleland SB, Knox AL, Heinrich J. 2000. MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in Drosophila. *The EMBO journal* **19**:144–55. doi:10.1093/emboj/19.1.144

Soruco MM, Chery J, Bishop EP, Siggers T, Tolstorukov MY, Leydon AR, Sugden AU, Goebel K, Feng J, Xia P, Vedenko A, Bulyk ML, Park PJ, Larschan E. 2013. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes & development* **27**:1551–6. doi:10.1101/gad.214585.113

Straub T, Grimaud C, Gilfillan GD, Mitterweger A, Becker PB. 2008. The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. *Plos Genetics* **4**:e1000302. doi:10.1371/journal.pgen.1000302

Straub T, Zabel A, Gilfillan GD, Feller C, Becker PB. 2013. Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. *Genome Research* **23**:473–85. doi:10.1101/gr.146407.112

Stuckenholz C, Meller VH, Kuroda MI. 2003. Functional Redundancy Within *roX1* , a Noncoding RNA Involved in Dosage Compensation in *Drosophila melanogaster*. *Genetics* **164**:1003–1014. doi:10.1093/genetics/164.3.1003

Sural TH, Peng S, Li B, Workman JL, Park PJ, Kuroda MI. 2008. The MSL3 chromodomain directs a key targeting step for dosage compensation of the Drosophila melanogaster X chromosome. *Nature structural & molecular biology* **15**:1318–25. doi:10.1038/nsmb.1520

Tikhonova E, Fedotova A, Bonchuk A, Mogila V, Larschan EN, Georgiev P, Maksimenko O. 2019. The simultaneous interaction of MSL2 with CLAMP and DNA provides redundancy in the initiation of dosage compensation in Drosophila males. *Development* **146**. doi:10.1242/dev.179663

Tikhonova E, Mariasina S, Arkova O, Maksimenko O, Georgiev P, Bonchuk A. 2022a. Dimerization Activity of a Disordered N-Terminal Domain from Drosophila CLAMP Protein. *IJMS* **23**:3862. doi:10.3390/ijms23073862

Tikhonova E, Mariasina S, Efimov S, Polshakov V, Maksimenko O, Georgiev P, Bonchuk A. 2022b. Structural basis for interaction between CLAMP and MSL2 proteins involved in the specific recruitment of the dosage compensation complex in *Drosophila*. *Nucleic Acids Research* **50**:6521–6531. doi:10.1093/nar/gkac455

Toups MA, Hahn MW. 2010. Retrogenes reveal the direction of sex-chromosome evolution in mosquitoes. *Genetics* **186**:763–6. doi:10.1534/genetics.110.118794

Urban J, Kuzu G, Bowman S, Scruggs B, Henriques T, Kingston R, Adelman K, Tolstorukov M, Larschan E. 2017. Enhanced chromatin accessibility of the dosage compensated Drosophila male X-chromosome requires the CLAMP zinc finger protein. *Plos One* **12**:e0186855. doi:10.1371/journal.pone.0186855

Valsecchi CIK, Basilicata MF, Georgiev P, Gaub A, Seyfferth J, Kulkarni T, Panhale A, Semplicio G, Manjunath V, Holz H, Dasmeh P, Akhtar A. 2021. RNA nucleation by MSL2 induces selective X chromosome compartmentalization. *Nature* **589**:137–142. doi:10.1038/s41586-020-2935-z

Valsecchi CIK, Basilicata MF, Semplicio G, Georgiev P, Gutierrez NM, Akhtar A. 2018. Facultative dosage compensation of developmental genes on autosomes in Drosophila and mouse embryonic stem cells. *Nature communications* **9**:3626. doi:10.1038/s41467-018-05642-2

Villa R, Forne I, Muller M, Imhof A, Straub T, Becker PB. 2012. MSL2 combines sensor and effector functions in homeostatic control of the Drosophila dosage compensation machinery. *Molecular Cell* **48**:647–54. doi:10.1016/j.molcel.2012.09.012

Villa R, Jagtap PKA, Thomae AW, Campos Sparr A, Forne I, Hennig J, Straub T, Becker PB. 2021. Divergent evolution toward sex chromosome-specific gene regulation in Drosophila. *Genes & development*. doi:10.1101/gad.348411.121

Villa R, Schauer T, Smialowski P, Straub T, Becker PB. 2016. PionX sites mark the X chromosome for dosage compensation. *Nature* **537**:244–248. doi:10.1038/nature19338

Wu L, Zee BM, Wang Y, Garcia BA, Dou Y. 2011. The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation. *Molecular Cell* **43**:132–44. doi:10.1016/j.molcel.2011.05.015

Yu G, Wang LG, He QY. 2015. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**:2382–3. doi:10.1093/bioinformatics/btv145

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**:R137. doi:10.1186/gb-2008-9-9-r137

Zheng S, Villa R, Wang J, Feng Y, Becker PB, Ye K. 2014. Structural basis of X chromosome DNA recognition by the MSL2 CXC domain during Drosophila dosage compensation. *Genes & development* **28**:2652–62. doi:10.1101/gad.250936.114

Zhou S, Yang Y, Scott MJ, Pannuti A, Fehr KC, Eisen A, Koonin EV, Fouts DL, Wrightsman R, Manning JE, 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. *The EMBO journal* **14**:2884–95.

Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC bioinformatics* **11**:237. doi:10.1186/1471-2105-11-237

FIGURES

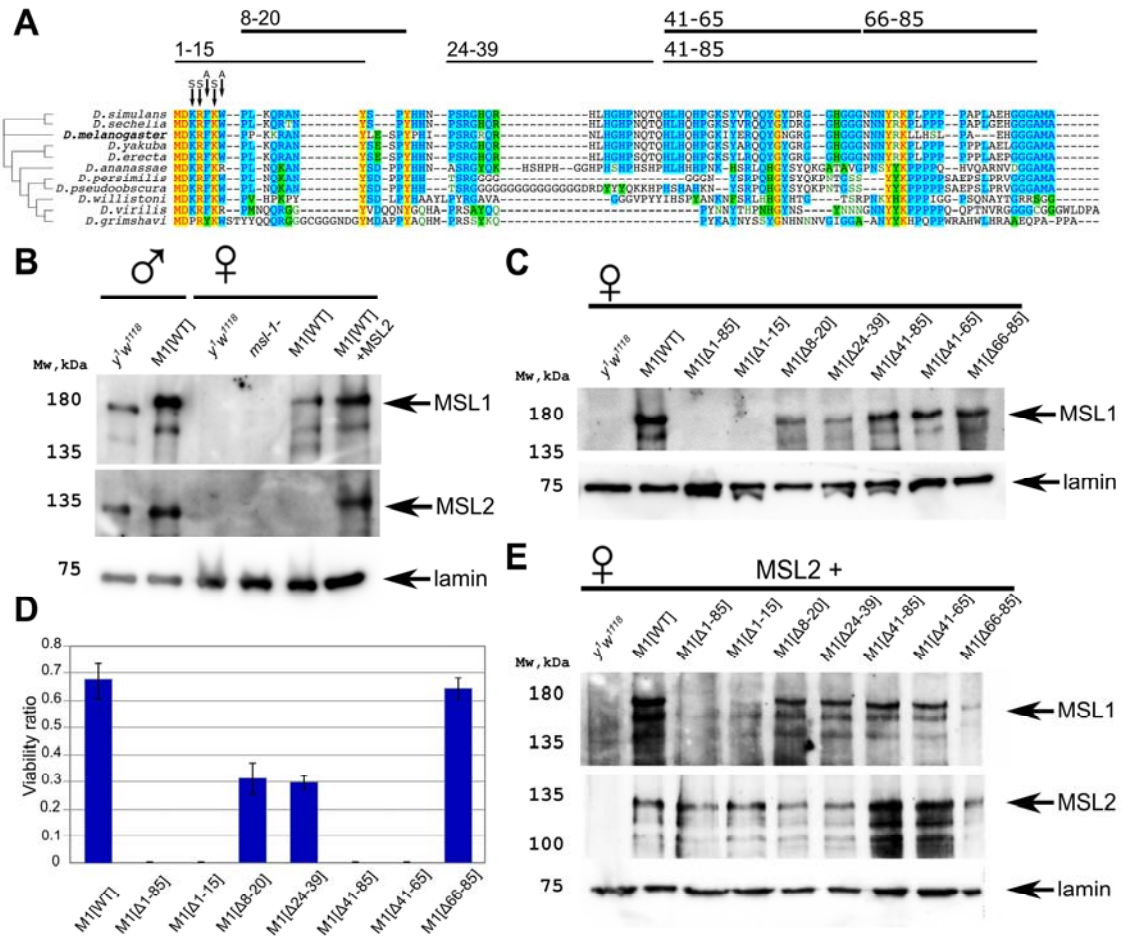

Figure 1. Testing the functional activity of MSL1 mutants. (A) Schematic presentation of the MSL1 deletion proteins expressed in flies. Alignment of the N-terminal domain (1-85 aa) of MSL1 among Drosophilidae. **(B)** Immunoblot analysis of protein extracts prepared from adult males y^1w^{1118} (control) and *msl-1*$^-$*; Ubi:msl-1*^{WT}*/ Ubi:msl-1*^{WT}*(M1[WT]),* and from adult females y^1w^{1118} (control), *msl-1*$^-$ *(msl-1*^{L60}*/ msl-1*$^{\gamma 269}$*)*, M1[WT] (*msl-1*$^-$; *Ubi:msl-1*^{WT}*/Ubi:msl-1*^{WT}) and M1[WT]+MSL2 (*msl-1*$^-$; *Ubi:msl-2*^{WT}*-FLAG / Ubi:msl-1*^{WT}). Immunoblot analysis was performed using anti-MSL1, anti-MSL2 and anti-lamin Dm0 (internal control) antibodies. **(C)** Immunoblot analysis of protein extracts prepared from adult females expressing different MSL1 protein variants (MSL1^{WT}, MSL1$^{\Delta 1\text{-}85}$ MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 8\text{-}20}$ MSL1$^{\Delta 24\text{-}39}$, MSL1$^{\Delta 41\text{-}85}$, MSL1$^{\Delta 41\text{-}65}$, MSL1$^{\Delta 66\text{-}85}$) and y^1w^{1118} females (control). Immunoblot analysis was performed using anti-MSL1, and anti-lamin Dm0 (internal control) antibodies. **(D)** Viability of males expressing MSL1 variants in the

msl-1-null background (marked as M1[*]). Viability (as a relative percentage) of *msl-1*$^{-}$*(msl-1*^{L60}*/ msl-1*$^{\gamma 269}$*); Ubi:msl-1*/ Ubi:msl-1** males to *msl-1*$^{-}$*; Ubi:msl-1*/Ubi:msl-1** females obtained in the progeny of crosses between females and males with the *msl-1*$^{-}$*/* CyO, GFP*; Ubi:msl-1*/Ubi:msl-1** genotype. *Ubi:msl-1** is any transgene expressing one of the tested MSL1 variants. The results are expressed as the mean of three independent crosses, with error bars showing standard deviations. (**E**) Immunoblot analysis of protein extracts prepared from adult female flies expressing MSL2-FLAG in combination with different MSL1 protein variants (MSL1^{WT}, MSL1$^{\Delta 1\text{-}85}$ MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 8\text{-}20}$ MSL1$^{\Delta 24\text{-}39}$, MSL1$^{\Delta 41\text{-}85}$, MSL1$^{\Delta 41\text{-}65}$, MSL1$^{\Delta 66\text{-}85}$) and *y*^{1}*w*^{1118} females (control). Immunoblot analysis was performed using anti-MSL1, anti-MSL2, and anti-lamin Dm0 (internal control) antibodies.

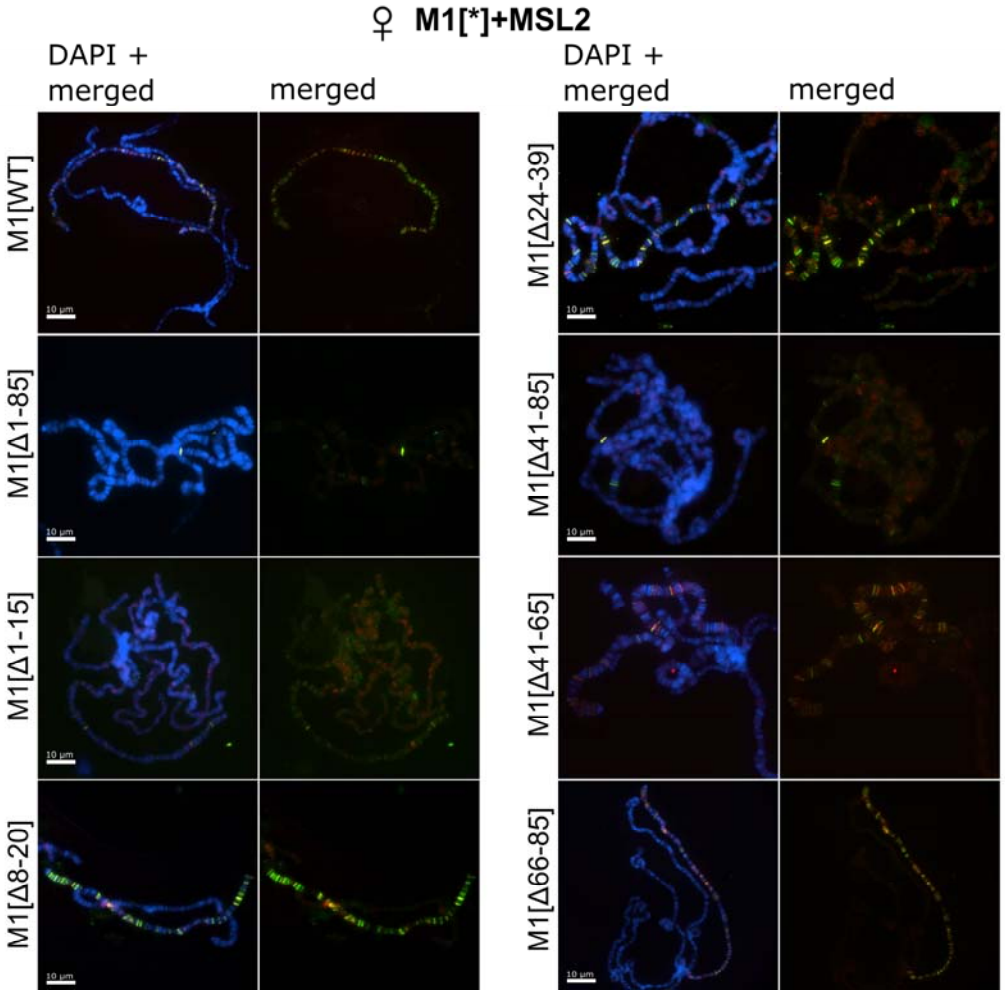

Figure 2. Testing the role of MSL1 mutants in recruiting the MSL complex on the X chromosome in a female model system. Distribution of the MSL complex on the polytene chromosomes from 3rd-day female larvae expressing both MSL1* variants and the MSL2^{WT}-FLAG protein. Panels show the immunostaining of proteins using rabbit anti-MSL1 antibody (green) and mouse anti-FLAG antibody (red). DNA was stained with DAPI (blue). M1[WT](♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1^{WT}*); M1[Δ1-85] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 1\text{-}85}$*); M1[Δ1-15] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 1\text{-}15}$*); M1[Δ8-20] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 8\text{-}20}$*); M1[Δ24-39] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 24\text{-}39}$*); M1[Δ41-85] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 41\text{-}85}$*); M1[Δ41-65] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 66\text{-}85}$*); M1[Δ66-85] (♀ *msl-1^{-}; Ubi:msl-2^{WT}-FLAG/ Ubi:msl-1$^{\Delta 66\text{-}85}$*).

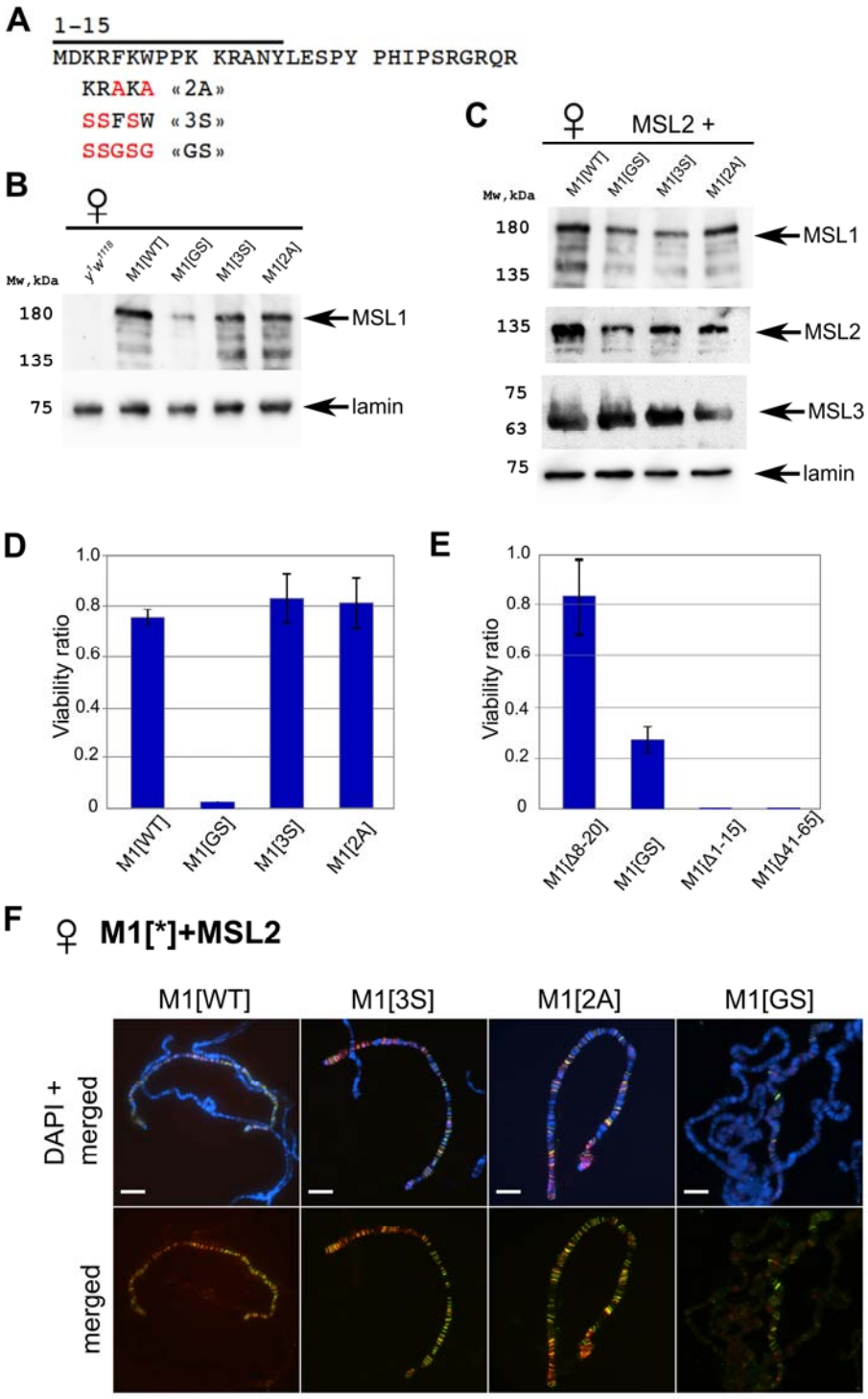

Figure 3. Testing the functional role of N-terminal amino acids 3-7 of MSL1. (A) Schematic presentation of the mutations at the N-terminus of MSL1. **(B)** Immunoblot analysis of protein extracts prepared from adult female flies expressing different MSL1* variants in the *msl-1*$^{-}$ (*msl-1*^{L60}/ *msl-1*$^{\gamma 269}$) background (M1[WT]: MSL1^{WT}; M1[2A]: MSL1$^{F5A\ W7A}$, M1[3S]: MSL1^{K3S}

$^{R4S\ K6S}$; M1[GS]: MSL1$^{K3S\ R4S\ F5G\ K6S\ W7G}$). **(C)** Immunoblot analysis of protein extracts prepared from adult female flies expressing MSL2-FLAG in combination with different MSL1* variants. Immunoblot analysis performed using anti-MSL1, anti-MSL2, anti-MSL3 and anti-lamin Dm0 (internal control) antibodies. **(D)** Viability of males expressing MSL1 variants (marked as M1[*]) in the *msl-1*$^{-}$ (*msl-1*^{L60}*/ msl-1*$^{\gamma 269}$) background. Viability (as a relative percentage) of *msl-1*$^{-}$*; Ubi:msl-1*/ Ubi:msl-1** males to *msl-1*$^{-}$*; Ubi:msl-1*/Ubi:msl-1** females obtained in the progeny of cross between females and males with *msl-1*$^{-}$*/* CyO, GFP*; Ubi:msl-1*/Ubi:msl-1** genotype. *Ubi:msl-1** is any transgene expressing one of the tested MSL1* variants. **(E)** Viability of males expressing MSL2^{WT} and MSL1* variants (marked as M1[*]) in the *msl-1*$^{-}$ background. Viability (as a relative percentage) of *msl-1*$^{-}$*; Ubi:msl-2*^{WT}*-FLAG /Ubi:msl-1** males to *msl-1*$^{-}$*; Ubi:msl-2*^{WT}*-FLAG /Ubi:msl-1** females obtained in the progeny of cross between males *msl-1*$^{-}$*/* CyO, GFP*; Ubi:msl-1*/Ubi:msl-1** and females *msl-1*$^{-}$*; Ubi:msl-2/Ubi:msl-2*. The results are expressed as the mean of three independent crosses, with error bars showing standard deviations. (**F**) Distribution of the MSL complex on the polytene chromosomes from 3rd-day female larvae expressing both MSL1* variants and the MSL2-FLAG protein. Panels show the merged results of immunostaining of MSL1 (green, rabbit anti-MSL1 antibody) and MSL2 (red, mouse anti-FLAG antibody). DNA was stained with DAPI (blue).

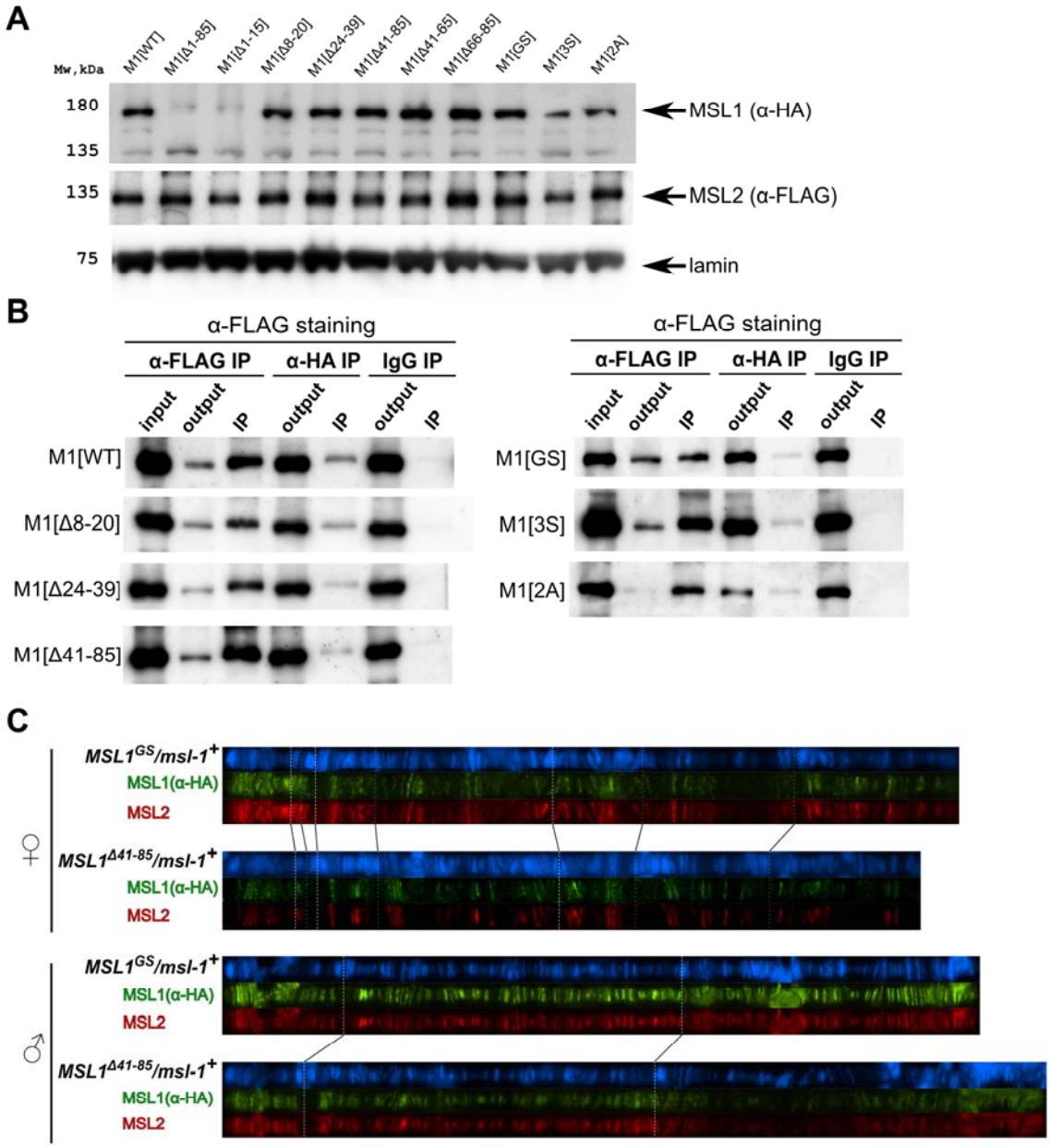

Figure 4. Testing the functional activity of mutant variants of the MSL1 protein. (A) Immunoblot analysis of protein extracts prepared from S2 cells expressing MSL2-FLAG and different MSL1-HAx3 protein variants (M1[WT]: $MSL1^{WT}$; M1[Δ*]: $MSL1^{\Delta*}$; M1[2A]: $MSL1^{2A}$, M1[3S]: $MSL1^{3S}$; M1[GS]: $MSL1^{GS}$). Immunoblot analysis performed using anti-HA (MSL1-HAx3 variants), anti-FLAG (MSL2-FLAG) and anti-lamin Dm0 (internal control) antibodies. **(B)** Protein extracts from Drosophila S2 cells co-transfected with different MSL1-

HAx3 variants and MSL2-FLAG were immunoprecipitated with antibodies against FLAG or HA or nonspecific mouse IgG as a negative control. The immunoprecipitates were analyzed by immunoblotting for the presence of FLAG-tagged proteins in immunoprecipitated samples. **(C)** Distribution of the MSL1*-HA and MSL2-FLAG on the polytene chromosomes from 3rd-day female or male larvae expressing one of MSL1*-HA variants and MSL2-FLAG at the wild-type background (*msl-1*$^{+}$*; Ubi:msl-2*^{WT}*-FLAG /Ubi:msl-1*)*. Panels show the merged results of immunostaining of MSL1*-HA (green, mouse anti-HA antibody) and MSL2 (red, rabbit anti-MSL2 antibody). DNA was stained with DAPI (blue).

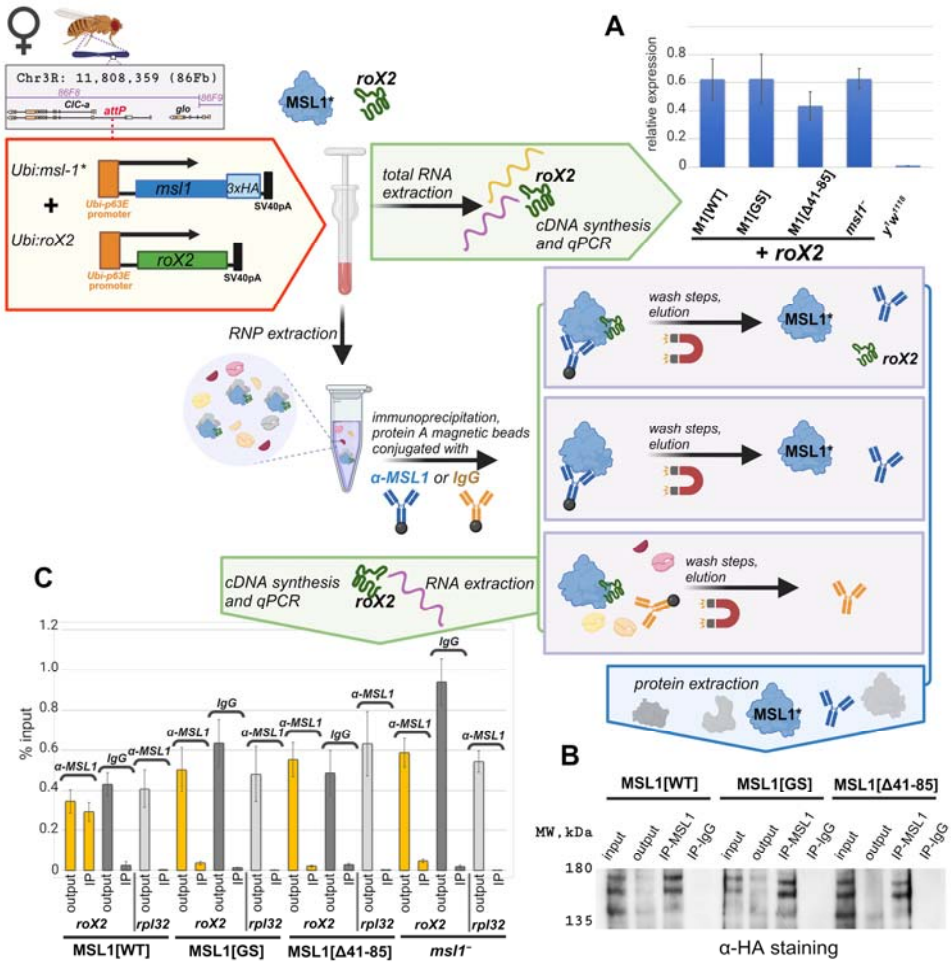

Figure 5. Testing role of the N-terminal region in recruiting of roX2 by MSL1. Extraction of RNA and RNP from adult flies, immunoprecipitation followed by RNA extraction were prepared as described in the Materials and Methods section and briefly schematically explained in the drawing. **(A)** Expression levels of the roX2 RNA in females of the y^1w^{1118}, M1[WT] (*msl-1⁻; Ubi:msl-1^{WT}-HA/ Ubi:roX2*), M1[GS] (*msl-1⁻; Ubi:msl-1^{GS}-HA/ Ubi:roX2*), M1[Δ41-85] (*msl-1⁻; Ubi:msl-1$^{\Delta 41\text{-}85}$-HA/ Ubi:roX2*) and *msl1⁻*(*msl-1⁻;Ubi:roX2/TM6,Tb*). Individual transcript levels were determined by RT-qPCR with primers for *roX2* gene normalized relative to *RpL32* for the amount of input cDNA. The error bars show standard deviations of triplicate measurements. **(B)** RNP extracts from adult females were immunoprecipitated with antibodies against MSL1 (IP-MSL1) or nonspecific rabbit IgG (IP-IgG) as a negative control. The efficiency of immunoprecipitation was tested by immunoblot analysis for the presence of HA-

tagged MSL1 protein in immunoprecipitate samples. **(C)** Total RNA was extracted from immunoprecipitates (α-MSL1 or IgG) and analyzed for the presence of *roX2* RNA by RT-PCR. *RpL32* was used as negative control. The results of immunoprecipitations are presented as the percentage of input cDNA. 'output' – supernatant after immunoprecipitation; 'IP' – immunoprecipitated sample. The error bars indicate SDs from three independent biological samples.

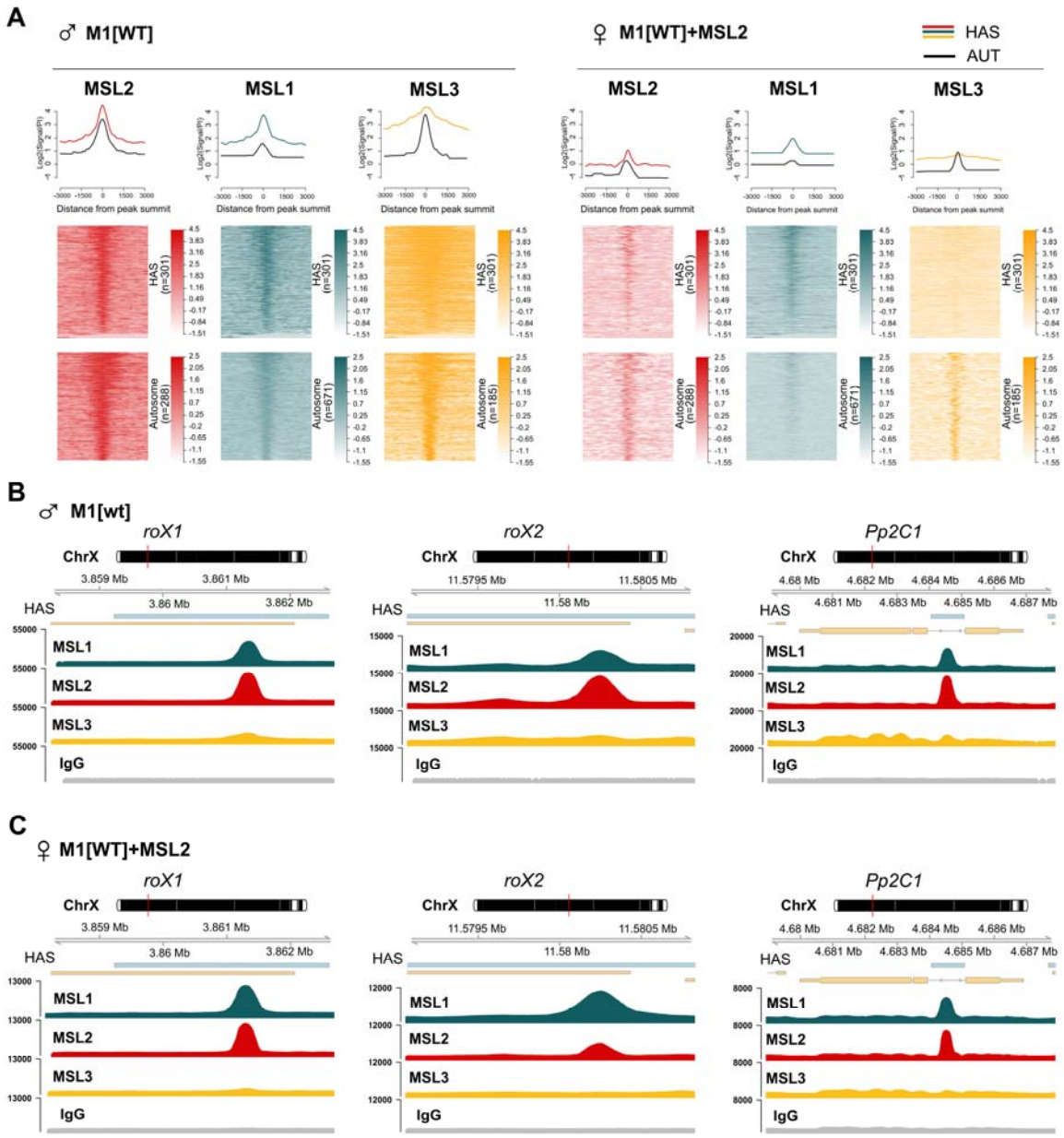

Figure 6. Comparisons of the binding patterns of MSL proteins between 2-3-day adult males and females expressing MSL2. (A) Average signal (top) and heatmap (bottom) showing the log fold-change (FC) between normalized (RPKM) test signals and nonspecific IgG signals for HAS and autosomal regions (see Materials and Methods). LogFC was calculated after smoothing signals using the Daniell kernel with kernel size 50 and the addition of a pseudocount. On the heatmaps, the peaks are ranked according to the average logFC in the M1[wt] male sample. ♂ M1[wt] indicates 2–3-day-old *msl-1*$^{-}$*; Ubi:msl-1*^{WT}*/ Ubi:msl-1*^{WT} males. ♀M1[wt]+MSL2 indicates 2–3-day-old *msl-1*$^{-}$*; Ubi:msl-2* ^{WT}*-FLAG/ Ubi:msl-1*^{WT} females. **(B** and

C) Enrichment of MSL1 (green), MSL2 (red), and MSL3 (yellow) signals associated with the X-linked genes encoding *Pp2C1* and the non-coding RNAs *roX1* and *roX2* in **(B)** M1[WT] males and (**C)** M1[WT]+MSL2 females.

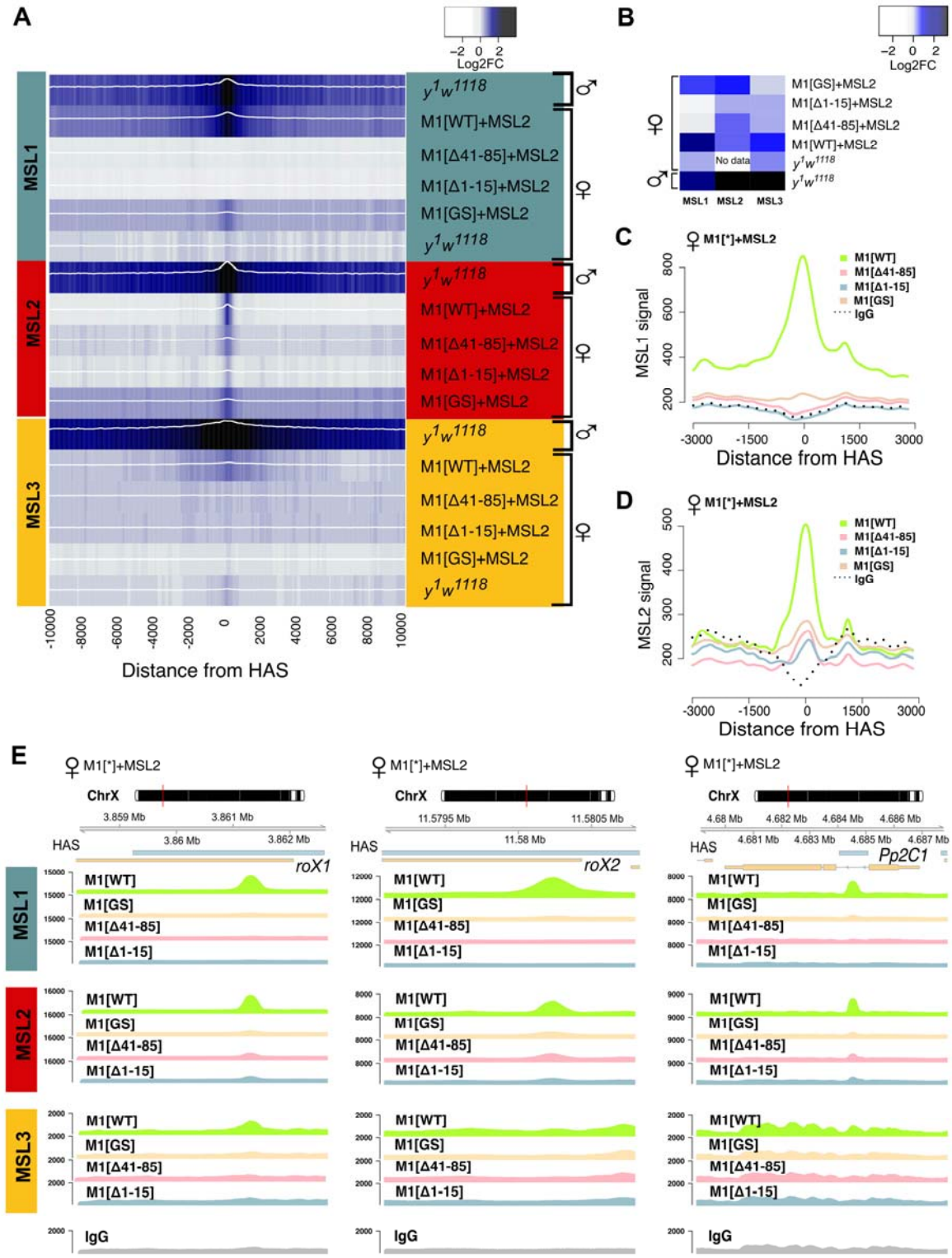

Figure 7. Role of the N-terminal regions in the recruitment of the dosage compensation complex to the HAS. To study the functional role of the N-terminal region of MSL1 for the recruitment of the dosage compensation complex, we compared the profiles of MSL1, MSL2, and MSL3 binding in 2-3-day females expressing MSL2 and one of four MSL1 variants

(MSL1^{WT}, MSL1^{GS}, MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 41\text{-}85}$). *y^{1}w^{1118}* (wild-type males and females), ♀M1[WT]+MSL2 (*msl-1^{-}; Ubi:msl-2^{WT}-FLAG/Ubi:msl-1^{WT}* females), ♀M1[Δ41-85]+MSL2 (*msl-1^{-}; Ubi:msl-2^{WT}-FLAG/Ubi:msl-1$^{\Delta 41\text{-}85}$* females), ♀M1[Δ1-15]+MSL2 (*msl-1^{-}; Ubi:msl-2^{WT}-FLAG/Ubi:msl-1$^{\Delta 1\text{-}15}$* females), ♀M1[GS]+MSL2 (*msl-1^{-}; Ubi:msl-2^{WT}-FLAG/Ubi:msl-1^{GS}* females). **(A**) Average log fold-change between normalized (RPKM) test signals and nonspecific IgG signals in HAS (see Materials and Methods). Average log fold-change was calculated after smoothing signals using the Daniell kernel with kernel size 50 and the addition of a pseudocount. Next, the tracks were aggregated using 100-bp bins. (**B)** Heatmap showing average log fold-change between normalized (RPKM) test signals and nonspecific IgG signals in region ± 500 bps from HAS centers. (**C** and **D)** Average signal (RPKM) for **(C)** MSL1 protein and **(D)** MSL2 protein in different fly lines at HAS (see Methods). **(E)** Enrichment of MSL1, MSL2, and MSL3 signals associated with the X-linked genes encoding *Pp2C1* and the non-coding RNAs *roX1* and *roX2* in females expressing MSL2 and one of four MSL1 variants (MSL1^{WT}, MSL1^{GS}, MSL1$^{\Delta 1\text{-}15}$, MSL1$^{\Delta 41\text{-}85}$).

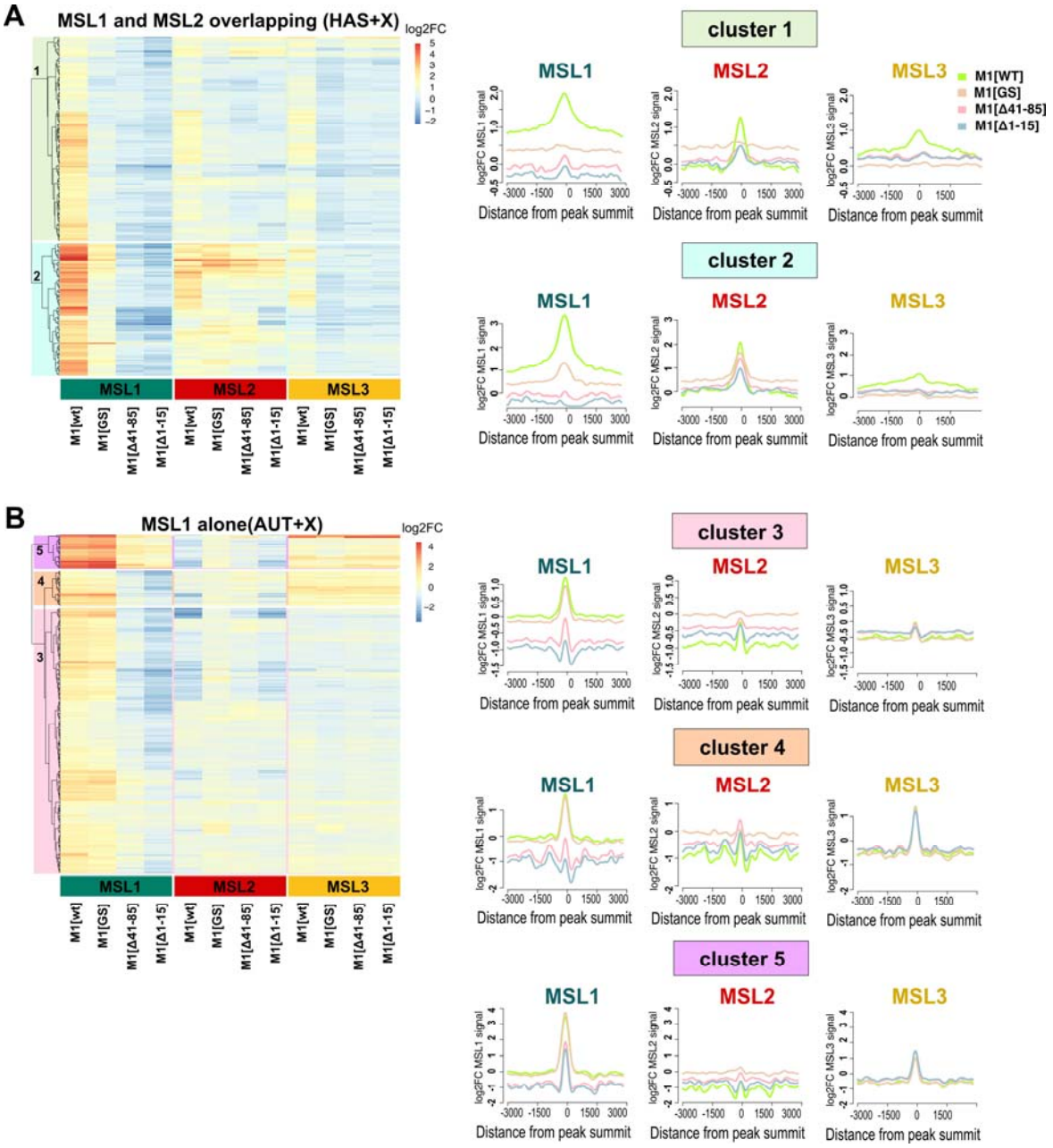

Figure 8. MSL1 shows different scenarios of binding upon N-terminal region modification. Regions clusterization according to MSLs binding in different female fly lines. Regions are divided into 2 groups: MSL1∩MSL2 overlapping peaks (X chromosome) (A) and MSL1 alone peaks (autosomes and X chromosome) (B) (see Materials and Methods). Left - Heatmap of average logFC between MSL and nonspecific IgG signal in female fly lines, hierarchical clusterization and cluster subdivision is drawn on the left corner (see Materials and Methods).

Right - Average logFC profiles of MSL signal in different clusters for female fly lines. Average log fold-change was calculated after smoothing signals using the Daniell kernel with kernel size 100 and the addition of a pseudocount.