nature communications

Article

https://doi.org/10.1038/s41467-024-55711-y

A noncanonical role of *roX* RNAs in autosomal epigenetic repression

Received: 10 November 2023

Accepted: 19 December 2024

Published online: 02 January 2025

Check for updates

Jianjian Li[1,2,6], Shuyang Xu[1,6], Zicong Liu[1,5,6], Liuyi Yang[1], Zhe Ming[1,3], Rui Zhang[1], Wenjuan Zhao[1], Huipai Peng[1], Jeffrey J. Quinn[4], Manyin Wu[1], Yushan Geng[1], Yuying Zhang[1], Jiazhi He[1], Minghai Chen[1], Nan Li[1], Ning-Yi Shao[3] & Qing Ma[1]

Long noncoding RNAs known as *roX* (*RNA on the X*) are crucial for male development in *Drosophila*, as their loss leads to male lethality from the late larval stages. While *roX* RNAs are recognized for their role in sex-chromosome dosage compensation, ensuring balanced expression of X-linked genes in both sexes, their potential influence on autosomal gene regulation remains unexplored. Here, using an integrative multi-omics approach, we show that *roX* RNAs not only govern the X chromosome but also target genes on autosomes that lack male-specific lethal (MSL) complex occupancy, together with Polycomb repressive complexes (PRCs). We observed that *roX* RNAs colocalize with MSL proteins on the X chromosome and PRC components on autosomes. Intriguingly, loss of *roX* function reduces X-chromosomal H4K16ac levels and autosomal H3K27me3 levels. Correspondingly, X-linked genes display reduced expression, whereas many autosomal genes exhibit elevated expression upon *roX* loss. Our findings propose a dual role for *roX* RNAs: activators of X-linked genes and repressors of autosomal genes, achieved through interactions with MSL and PRC complexes, respectively. This study uncovers the unconventional epigenetic repressive function of *roX* RNAs with PRC interaction.

The maintenance of correct chromosome numbers is vital for the normal development of organisms. Chromosomal aneuploidy, characterized by the gain or loss of individual chromosomes, can lead to gene overdosage or haploinsufficiency, often resulting in diseases that pose threats to human health. An intriguing exception to this rule is found in sex chromosomes, which can tolerate unequal chromosome numbers through a phenomenon known as dosage compensation. In many animals, including humans, mice, and flies, sex is determined by dimorphic sex chromosomes; males possess heterogametic (XY) chromosomes, while females possess homogametic (XX) chromosomes. Across these species, multiple dosage compensation mechanisms have evolved to equalize the expression of genes on the X chromosome between males and females[1]. In eutherian mammals, X chromosome inactivation (XCI) serves as a dosage compensation mechanism. This process is regulated by a long noncoding RNA (lncRNA) named *XIST*, which initiates XCI. *XIST* is expressed exclusively from one of the two X chromosomes in females and functions in *cis* to recruit repressive protein complexes[2–4], including Polycomb

[1]Shenzhen Key Laboratory of Synthetic Genomics, Guangdong Provincial Key Laboratory of Synthetic Genomics, Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. [2]Faculty of Synthetic Biology, Shenzhen University of Advanced Technology, Shenzhen, China. [3]Faculty of Health Sciences, University of Macau, Macau, Macau SAR, China. [4]Center for Personal Dynamic Regulomes and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA, USA. [5]Present address: Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Nijmegen, Nijmegen, The Netherlands. [6]These authors contributed equally: Jianjian Li, Shuyang Xu, Zicong Liu. e-mail: qing.ma@siat.ac.cn

Repressive Complexes (PRCs). These complexes trigger the deposition of repressive chromatin marks such as DNA methylation and histone modification H3K27me3, ultimately leading to the condensation and inactivation of the entire X chromosome[5,6]. Notably, the strategy of *XIST*-mediated chromosome-wide gene repression has been successfully applied to silence the extra chromosome 21 in pluripotent stem cells from patients with Down syndrome[7]. For monosomy-related diseases, gene expression on a single chromosome needs to be gained. However, chromosome-wide gene activation mechanisms in mammals have not been uncovered till now. In contrast to mammals, *Drosophila* flies lack the *XIST* gene in their genomes and have evolved a distinct class of lncRNAs called *RNA on the X* (*roX*). These *roX* RNAs function redundantly to achieve dosage compensation on the male X chromosome[8]. Specifically, they are expressed from and predominantly localize to the single male X chromosome, exclusively in male flies. This localization is facilitated by the assembly of the male-specific lethal (MSL) complex[9], which deposits the histone modification H4K16ac and triggers an approximately twofold hyperactivation of gene expression on the male X chromosome. Despite the fundamental differences in the epigenetic mechanisms of dosage compensation between mammals and flies, both have developed lncRNAs to regulate gene expression on the entire X chromosome. Interestingly, a recent study has suggested that a *roX*-mediated dosage compensation mechanism could be transplanted into mammalian cells when replacing the specific protein domain of mammalian MSL2[10], thus potentially offering an alternative approach for treating monosomy-related diseases.

Notably, *roX* RNAs in *Drosophila* exhibit a versatile function beyond regulating X chromosome dosage compensation. They have been shown to bind to autosomal sites in a *trans*-regulatory manner[11–13], and some of these binding sites are evolutionarily conserved across Drosophilid species[14]. Additionally, the relocalization of *roX* RNAs from the X chromosome to autosomes has been observed under certain conditions[13]. However, the functional implications of *roX* RNA binding to autosomal sites remain largely unexplored. Transgenic expression of *roX* RNA on autosomes rescues the male larval lethality phenotype of *roX1 roX2* double knockout mutants[15] and triggers the assembly of MSL complex on both the transgene-located autosome as well as the X chromosome[16–19], suggesting that *roX* RNAs possess both *cis*- and *trans*-regulatory capabilities.

In this study, we adopt an integrative multi-omics approach to comprehensively investigate the genome-wide function of *roX* RNAs. Through ChIRP-seq analysis, we characterize the genomic occupancy of *roX* RNAs and identify significant enrichment of Polycomb-related DNA motifs at *roX* binding sites, in addition to the canonical MSL recognition element (MRE). Despite the pronounced colocalization of *roX* RNAs and MSL proteins on the X chromosome, our findings reveal that *roX* RNA occupancy on autosomes is independent of dosage compensation. Instead, we observe a strong colocalization between *roX* RNAs and several PRC subunits, as well as the repressive histone mark H3K27me3. Furthermore, we identify the PRC2 subunit Caf1-55 (also known as NURF55) as a *roX2*-interacting protein through ChIRP-MS. Our ChIP–seq analysis further demonstrates that the loss of *roX* genes compromises H3K27me3 deposition in the vicinity of autosomal *roX* binding sites, particularly those marked by high levels of H3K27me3. Transcriptome analysis reveals a distinct pattern of gene expression changes upon *roX* loss, with upregulated expression of autosomal genes in male larvae, in contrast to the decrease observed in X-linked genes. Interestingly, the *roX*-regulated autosomal genes are implicated in critical developmental processes such as anatomical structure morphogenesis and nervous system development. Taken together, our results suggest that *roX* RNAs possess dual roles in epigenetic regulation, influencing both autosomal and X chromosome gene expression. These findings reveal the unconventional epigenetic repressive function of *roX* RNAs via PRC, and shed light on the intricate mechanisms underlying dosage compensation and gene regulation, providing insights into the roles of lncRNAs in shaping the epigenomic landscape.

Results

roX lncRNAs bind autosomal loci independently of MSL proteins

To unravel the intricate chromatin regulatory networks orchestrated by *roX* lncRNAs (*roX1* and *roX2*), we meticulously analyzed *roX* ChIRP-seq data obtained from male *Drosophila* larvae[14]. Consistent with their functional redundancy[15], the genomic binding profiles of both *roX1* and *roX2* exhibited high concordance on both the X chromosome and autosomes (Supplementary Fig. 1a). Employing a stringent approach (see Methods), we identified a total of 942 distinct *roX* binding sites, with a majority located on the X chromosome (Fig. 1a and b, Supplementary Fig. 1b, Supplementary Data 1), aligning with the established role in male X chromosome dosage compensation[20]. Intriguingly, albeit less frequent, *roX* occupancy extended to autosomal regions and these binding sites are not evenly dispersed on autosomes (Fig. 1b, Supplementary Fig. 1b). The authenticity of these autosomal peaks as *roX* binding sites was confirmed by several lines of evidence: (1) robust reproducibility of the autosomal peaks (Supplementary Fig. 1c), (2) consistent genomic distribution patterns shared between X and autosomal *roX* peaks, favoring promoter regions (Supplementary Fig. 1d), (3) comparable binding intensity of *roX* to both the X chromosome and autosomes (Supplementary Fig. 1e), coupled with high accessibility (Supplementary Fig. 1f)[21], and (4) conservation of certain autosomal *roX* peaks across diverse Drosophilid species[14].

Given that *roX* RNAs and MSL proteins exhibit binding to autosomal regions (Fig. 1, a and b)[15,22–24], we sought to investigate the potential interplay between autosomal *roX* occupancy and MSL function. To this end, we generated X and autosomal binding profiles of MSL proteins and the H4K16ac histone modification in wild-type and *roX1 roX2* double knockout (referred to as *roX* knockout or roX-KO) male larvae, leveraging publicly available chromatin immunoprecipitation sequencing (ChIP–seq) data[10,23,25]. The majority of X-chromosomal *roX* peaks displayed substantial enrichment of MSL proteins, including MSL2, MOF, and MLE, as well as H4K16ac modification (Fig. 1c–e), consistent with their established canonical function in dosage compensation. Strikingly, loss of *roX* genes resulted in the abolition of MSL binding around these X-chromosomal *roX* peaks in male larvae (Fig. 1c, and Supplementary Fig. 2a). Concurrently, H4K16ac enrichment in these regions declined to levels akin to those observed in females, where dosage compensation is absent (Fig. 1, c and e, and Supplementary Fig. 2a). Notably, autosomal *roX* peaks exhibited minimal colocalization with MSL proteins or H4K16ac modification (Fig. 1, c-e, Supplementary Fig. 2, a–c). Furthermore, MSL signals around autosomal *roX* peaks remained unaffected by *roX* loss while the relatively weak H4K16ac signals on these autosomal regions showed a slight average increase (Fig. 1, c and d, and Supplementary Fig. 2a). These suggest that *roX* occupancy on autosomes does not reflect recruitment of a functional MSL complex. In S2 cells, we also observed MSL enrichment specifically on the identified X-chromosomal *roX* binding sites (Fig. 1d, Supplementary Fig. 2d)[24]. Conversely, we examined *roX2* RNA ChIRP-seq signals on MSL binding sites in S2 cells, and no *roX2* enrichment on autosomal MSL binding sites was observed (Supplementary Fig. 3)[26], further indicating that *roX* binding on autosomal sites is independent of MSL dosage compensation.

Autosomal *roX* occupancy is linked to the repressive histone mark H3K27me3

In order to elucidate the functional role of *roX* occupancy on autosomes, we analyzed the epigenomic landscape surrounding *roX* binding sites[10,25,27]. Remarkably, the autosomal peaks exhibited a marked enrichment of the repressive histone mark H3K27me3, with a near

https://doi.org/10.1038/s41467-024-55711-y

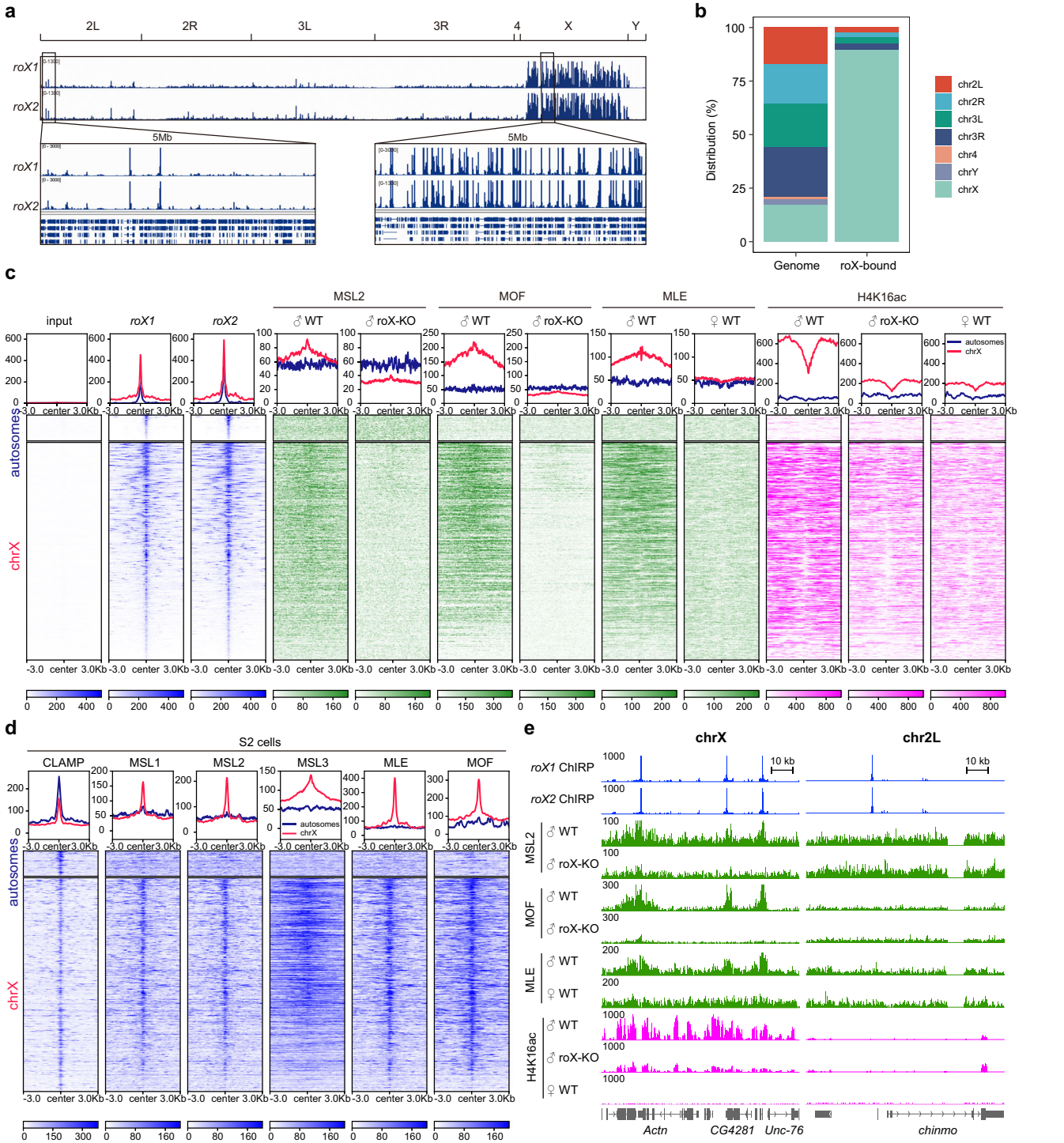


Fig. 1 | Genomic binding patterns of *Drosophila roX* RNAs: autosomal binding independent of MSL proteins. a Genomic distribution of *roX* RNAs across the *Drosophila* genome. The upper panel presents an overview of *roX* ChIRP-seq data, revealing robust *roX* occupancy on both autosomal and X-chromosomal regions. The lower panels provide magnified views of *roX* ChIRP-seq data, highlighting the presence of *roX* RNAs on both chromosome types. **b** Identification of a subset of *roX* ChIRP-seq peaks on autosomes, indicating that *roX* RNAs can bind to non-sex chromosomes. **c** Comparison of *roX* RNA binding patterns between X-chromosomal and autosomal regions. X-chromosomal *roX* peaks exhibit notable enrichment of Male-Specific Lethal (MSL) proteins and H4K16ac histone modification, while autosomal roX peaks lack such enrichment. Wild-type (WT) and *roX1 roX2* double knockout mutant (roX-KO) larval samples were analyzed. **d** MSL proteins show no binding signals in S2 cells on autosomal *roX* binding sites. Autosomal regions bound by *roX* RNAs are occupied by CLAMP but not MSL proteins in S2 cells. **e** Genome browser view illustrating representative examples of colocalization between MSL proteins and *roX* RNAs on the X-chromosome, contrasting with the absence of such colocalization on autosomes. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-55711-y

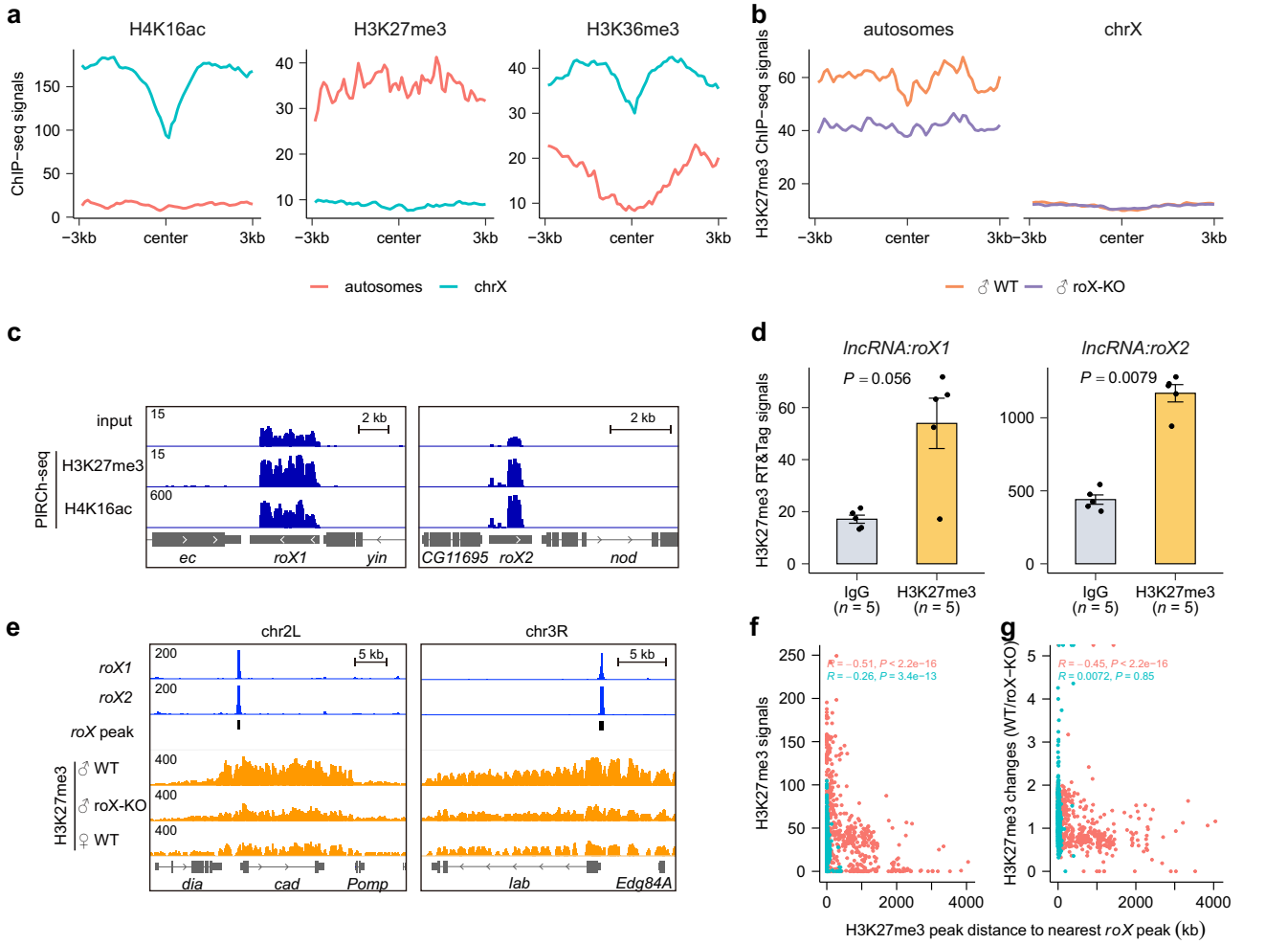


Fig. 2 | Interactions and impact of *roX* RNAs on H3K27me3 chromatin deposition. a Association of *roX* RNAs with histone marks on the X chromosome and autosomes. Binding intensities of specific histone marks, including the active marks H4K16ac and H3K36me3 on the X chromosome, and the repressive H3K27me3 on autosomes, are depicted within ±3 kb genomic regions surrounding *roX* peaks. (**b**) Reduction in H3K27me3 enrichment at autosomal roX binding sites following roX RNA loss. Binding intensities of H3K27me3 are compared between wild type (WT) and *roX* double knockout (roX-KO) male larvae within ±3 kb genomic regions surrounding *roX* peaks. Evidence of physical interactions between *roX* RNAs and H3K27me3, as demonstrated by PIRCh-seq (**c**) and RT&Tag (**d**) data. Mean ± SEM values are shown ($n = 5$ independent experiments), and significance levels are determined using the two-sided Wilcoxon test (**d**). (**e**) Genome track view displaying representative examples of diminished H3K27me3 enrichment near autosomal *roX* binding sites upon *roX* RNA depletion. In male larvae, the enrichment of H3K27me3 (**f**) and its alterations upon *roX* loss (**g**) exhibit a notably stronger correlation with the proximity to the nearest *roX* peaks on autosomes in comparison to the X chromosome. As the distance between H3K27me3 peaks and *roX* peaks increases, both H3K27me3 signals (**f**) and the fold change in signal upon *roX* loss (**g**) demonstrate a diminishing trend on autosomes. Pearson's correlation coefficients (R) are employed to assess these associations. Source data are provided as a Source Data file.

absence of active histone marks such as H4K16ac and H3K36me3, in contrast to the highly enriched active histone marks observed on the X chromosome (Fig. 2a). This intriguing observation suggests an additional role for *roX* beyond its canonical involvement in epigenetic activation. To validate the potential association of *roX* with H3K27me3, we conducted PIRCh-seq—a technique designed to capture histone modification-associated RNAs[28]—to capture H4K16ac- and H3K27me3-associated RNAs in *Drosophila*. Consistent with our hypothesis, robust enrichment of *roX* transcripts, particularly *roX2*, was observed in both H4K16ac- and H3K27me3-associated RNAs (Fig. 2c). A recently published H3K27me3 RT&Tag dataset in S2 cells also corroborated this finding, indicating enrichment of *roX* RNAs in H3K27me3 immunoprecipitants (Fig. 2d)[29].

To further study the potential regulatory function of *roX* RNAs in H3K27me3 deposition, we performed H3K27me3 ChIP–seq in roX-KO male larvae and their wild-type counterparts[30]. Strikingly, in male larvae, the average H3K27me3 signals on autosomal *roX* peaks surpassed those on X-chromosomal peaks (Fig. 2b, e), reinforcing earlier observations. In males, X-chromosomal *roX* peaks is known to exhibit pronounced H4K16ac signals (Fig. 1c)[12,25]. These regions, as expected, showed minimal average H3K27me3 enrichment (Fig. 2b). Importantly, a substantial reduction in average H3K27me3 levels was observed in proximity to autosomal *roX* peaks upon *roX* loss, such as at the regulatory regions of *cad* and *lab*, while no such reduction was observed in the vicinity of X-chromosomal peaks (Fig. 2b, e). Notably, autosomal H3K27me3 peaks tended to have stronger H3K27me3 signals when they are closer to *roX* binding sites, and these regions demonstrated a more pronounced decrease upon *roX* gene deletion (Fig. 2f, g). This trend was, however, absent on the X chromosome. Collectively, these findings suggest a role for *roX* in sustaining genomic H3K27me3 deposition, especially on autosomes.

roX RNAs modulate gene expression on X and autosomes

To gain a comprehensive understanding of the regulatory network governed by *roX* RNAs and their impact on autosomal gene expression, we conducted RNA sequencing (RNA-seq) experiments in male roX-KO

https://doi.org/10.1038/s41467-024-55711-y

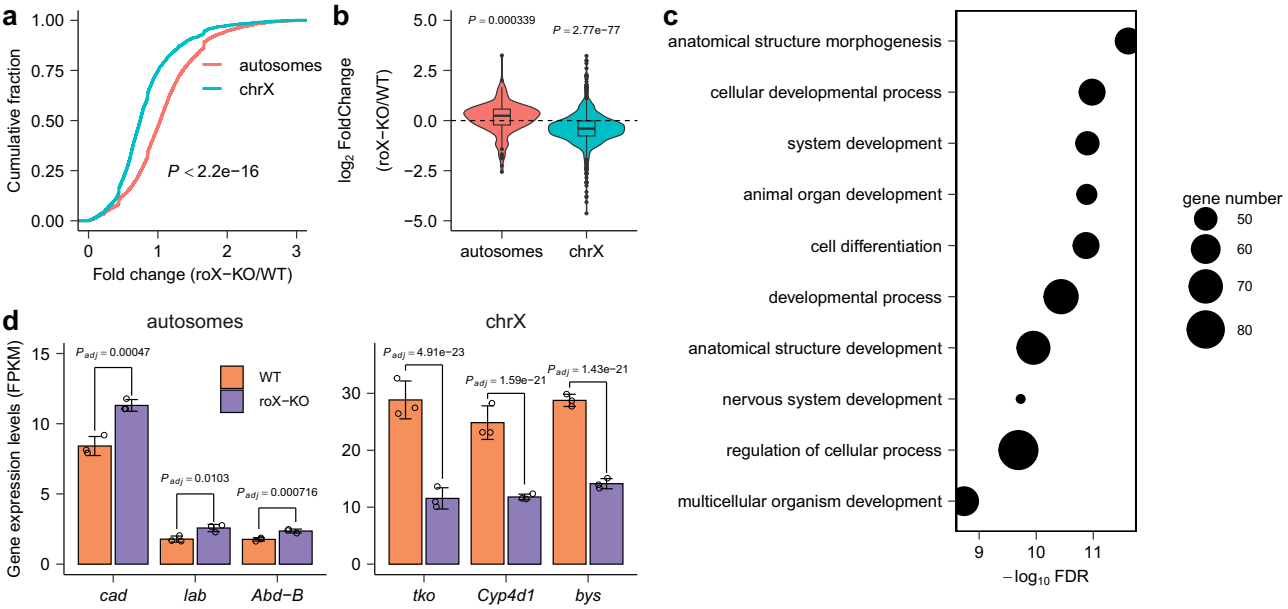


Fig. 3 | *roX* RNAs promote gene expression on X but repress gene expression on autosomes. a Cumulative curves illustrating the fold change (roX-KO/WT) in gene expression on autosomes and the X chromosome following *roX* loss in male larvae. Significance levels are determined using the two-sided Kolmogorov–Smirnov test. **b** roX-bound autosomal genes ($n = 203$) show upregulated expression of in roX-KO larvae while X-chromosomal targets ($n = 1346$) exhibit decreased expression. The center lines represent the median, the box limits represent the first and third quartiles, and the whiskers indicate 1.5 × the interquartile range (IQR). Significance levels are assessed using the two-sided Wilcoxon test. **c** The top 10 significantly enriched Gene Ontology (GO) terms in roX-bound autosomal genes are presented. Significance levels are determined using the Fisher's exact test, with the Benjamini–Hochberg false discovery rate (FDR) correction for multiple testing. **d** Examples of genes located on autosomes and X chromosomes that exhibit increased and decreased expression, respectively, upon *roX* loss. Mean ± SD FPKM values from RNA-seq analyses for each gene are plotted ($n = 3$ independent experiments) and the adjusted *P* values using Benjamini and Hochberg method are indicated. Source data are provided as a Source Data file.

larvae and their wild-type counterparts. The most significantly enriched Gene Ontology (GO) terms in differentially expressed genes between roX-KO and wild type males are related to cuticle development and morphogenesis (Supplementary Fig. 4, Supplementary Data 2). Strikingly, we observed an imbalance in transcriptional outputs between the X chromosome and autosomes upon *roX* loss ($P < 2.2 \times 10^{-16}$, Kolmogorov–Smirnov test, Fig. 3a). As anticipated, the majority of roX-bound genes on the X chromosome displayed decreased expression levels (Fig. 3b), consistent with their established role in dosage compensation. Interestingly, autosomal genes bound by *roX* were notably enriched in GO terms associated with morphogenesis and organ development (Fig. 3c). Notably, a substantial portion of these genes exhibited elevated transcription (Fig. 3b, d), coinciding with the diminished H3K27me3 levels observed on autosomes (Fig. 2b, e). Since *roX* RNAs exhibit a sex-specific expression pattern and function in males from late embryonic stages onward[31], we wondered whether *roX* binding on autosomes was associated with sex-biased gene expression. For the identified roX-bound autosomal genes, we found that their expression changes after *roX* loss (roX-KO versus wild type) were indeed correlated with the sex bias index (female versus male gene expression ratio, Fig. 4a, Supplementary Data 3). Specifically, roX-repressed genes predominantly exhibited female-biased expression in wild-type larvae (Fig. 4a, b). Additionally, these roX-repressed, female-biased genes were associated with roX-mediated H3K27me3 deposition in males (Fig. 4c) and were enriched in GO terms related to female development (Fig. 4d). Consequently, it is plausible that *roX* plays a repressive role in autosomal gene expression to prevent inappropriate female differentiation and development in male organs.

We also examined expression of roX-bound genes in *msl* knockdown S2 cells and found that X-linked genes bound by *roX* displayed significantly decreased expression when MSL protein members were depleted, while those autosomal genes did not showed significant variation (Supplementary Fig. 5a)[32,33], supporting that *roX* RNAs are assembled into MSL complex only on the X chromosome. To exclude the possibility that upregulation of autosomal genes resulted from the overfitting artifact of the linear modeling of DESeq2, we also performed differential gene expression for only the autosomal genes. Only roX-bound autosomal genes displayed a significant upregulation in the roX-KO larvae (Supplementary Fig. 5b), further indicating that *roX* RNAs possess a repressive role in autosomal gene expression.

roX RNAs on autosomes co-localize with Polycomb repressive complexes

The seemingly contrasting roles of *roX* RNAs, activating gene expression on the X chromosome and repressing gene expression on autosomes, led us to delve into the underlying chromatin features of *roX* binding sites. De novo motif discovery analysis unveiled DNA motifs similar to the MSL recognition element (MRE), notably enriched on both X and autosomal *roX* binding sites, although X peaks exhibited higher significance (Fig. 5a)[34]. Since CLAMP protein can directly bind this GA-rich DNA motif[35,36], we also examined its occupancy on *roX* binding sites. Consistent with the enriched DNA motif, CLAMP protein exhibited strong signals on both X-chromosomal and autosomal *roX* binding sites (Fig. 1d)[35]. Although it has been reported that CLAMP can directly interact with MSL2 and MLE in vitro[37–41] and recruit MSL proteins to the X chromosome[35], minimal colocalization of MSL proteins on these autosomal roX-bound regions could be observed (Fig. 1c–e, Supplementary Fig. 2d). This concurs with the suggestion that the MRE alone may not be sufficient to discriminate between X and autosomal regions[42]. Further motif enrichment analysis for known DNA motifs in *roX* peaks revealed the binding motifs for Polycomb-group protein Pleiohomeotic (Pho) and Trithorax group Zeste, both of which are critical for the function of certain Polycomb/Trithorax response elements (PREs/TREs)[43,44], were more significantly enriched on autosomal than X-chromosomal *roX* peaks (Fig. 5b). This intriguing observation

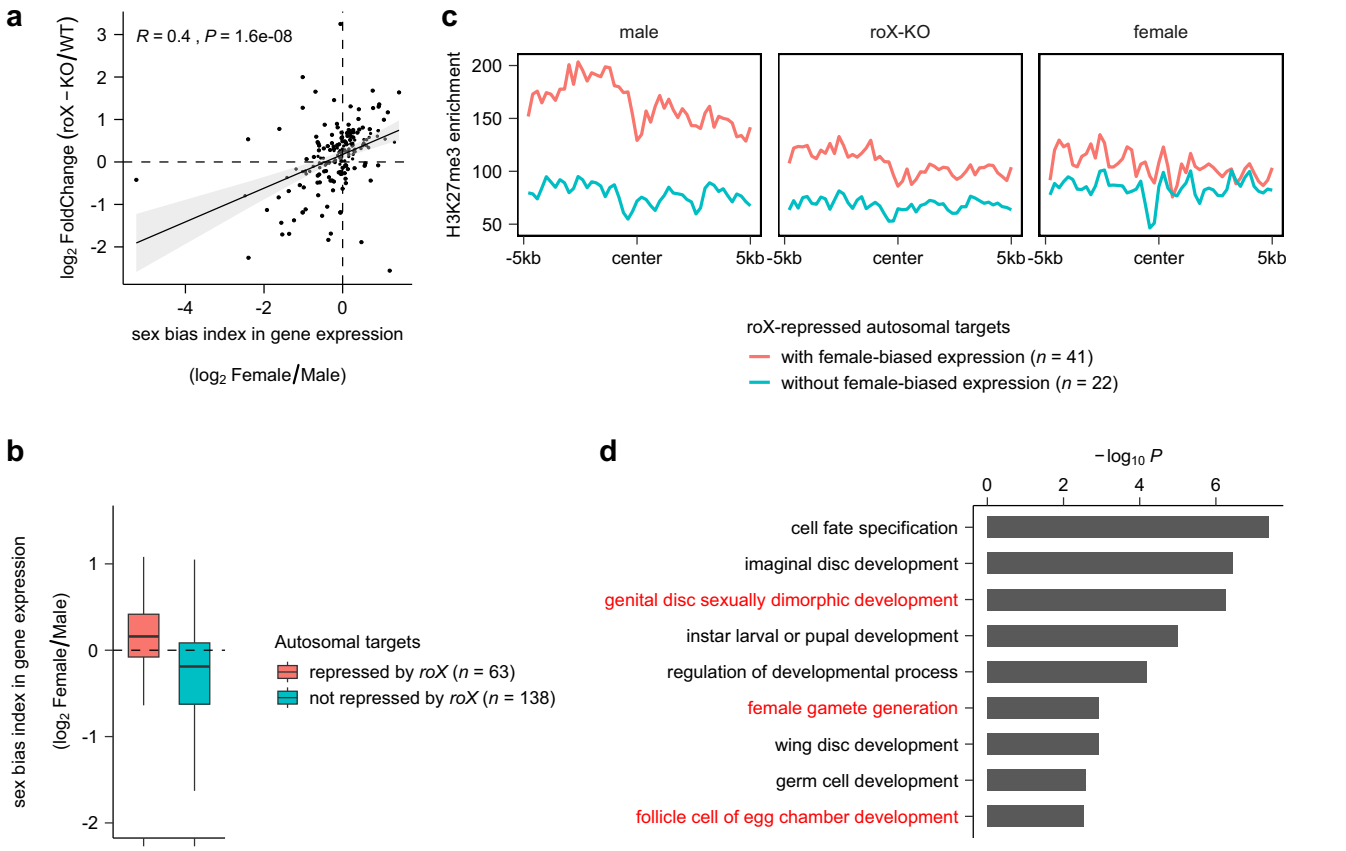


Fig. 4 | *roX* RNAs directly repress autosomal genes involved in female development. a Scatter plot showing expression changes of roX-bound autosomal genes after *roX* loss and sex ratio in their expression (female versus male). Pearson's correlation coefficients (R) were used to assess these associations. The linear regression line is depicted in black, with the 95% confidence interval represented as a lightgrey shadow. **b** A majority of autosomal genes directly repressed by *roX* show female-biased expression. The center lines represent the median, the box limits represent the first and third quartiles, and the whiskers indicate 1.5 × the interquartile range (IQR). **c** *roX* directly repressed autosomal genes showing female-biased expression are associated with higher H3K27me3 levels in males on roX-binding sites, which declined after *roX* loss. **d** Enriched GO terms of *roX* directly repressed, female-biased autosomal genes as shown in (**c**). The terms related to female development are highlighted in red. Significance levels are determined using the one-sided hypergeometric test. Source data are provided as a Source Data file.

suggests a potential connection between Polycomb repressive complexes (PRCs) and *roX* occupancy on autosomes.

Given the ability of Pho and Zeste to recruit PRCs in *Drosophila*, we investigated whether *roX* occupancy on autosomes is associated with PRC function. *Drosophila* possesses two major PRCs with enzymatic activity, namely PRC1 and PRC2, which catalyze H2AK118ub1 (H2AK119ub1 in vertebrates) and H3K27me3 histone modifications, respectively, to repress gene expression[45]. Our analysis revealed higher binding signals of PRC1 (Ph, Pc, and Psc) and PRC2 (E(z) and Su(z)12) subunits, as well as some known PRC recruiters (Pho and Spps), on autosomal *roX* binding sites compared to X-chromosomal ones (Fig. 5c and Supplementary Fig. 6a)[27,46]. In stark contrast, MSL proteins exhibited minimal binding on autosomal *roX* peaks (Fig. 1c-e, Supplementary Fig. 2a–c)[23–25]. Importantly, PRC proteins displayed a stronger correlation with *roX* on autosomes than MSL proteins (Supplementary Fig. 6a, b). Notably, Ph binding on autosomal *roX* binding sites was markedly diminished in ovaries where *roX* RNAs are absent, as compared to larval samples (Supplementary Fig. 6c)[46,47]. These collective findings suggest that *roX* RNAs might exert an influence on PRC complex function in *Drosophila*.

roX RNAs interact with PRCs to modulate H3K27me3 deposition on autosomes

Since autosomal *roX* occupancy was found to be unrelated to dosage compensation (Fig. 1c), we hypothesized the involvement of other roX-associated nuclear protein complexes. To explore this, we conducted *roX2* ChIRP-MS in *Drosophila* S2 cells, revealing 562 proteins interacting with *roX2* (Fig. 6a, Supplementary Data 4). These interactions included all known MSL proteins (MSL1, MSL2, MSL3, MLE, and MOF) associated with *roX* RNAs[20], and CLAMP protein essential for MSL recruitment[35](Fig. 6b). Importantly, roX2-interacting proteins exhibited enrichment not only for dosage compensation but also for pathways associated with gene silencing, such as repressor pathways and SUMOylation of transcription cofactors (Fig. 6b). Intriguingly, an interaction between *roX2* and PRC2 subunit Caf1-55 (also known as NURF55) was also observed (Fig. 6a, b), aligning with the enriched DNA motifs linked to Polycomb recruitment on autosomal *roX* binding sites (Fig. 5b). We also employed the PRIdictor web tool[48] to predict RNA–protein interactions, revealing roX2's potential interactions with several PRC subunits, with comparable or even higher probabilities than interactions with MSL proteins (Supplementary Fig. 7). To verify the interaction of *roX1* with PRCs, we performed Trimolecular Fluorescence Complementation (TriFC) assays in H293T cells, ectopically expressing *roX1* and PRC subunit NURF55 or Esc (or MSL2 protein as a positive control). The results demonstrated robust physical interactions between *roX1* and both NURF55 and Esc, akin to interactions with MSL2 protein (Fig. 6d, e). While both MSL and PRC proteins have been observed to interact with *roX* RNAs, PRC components were absent in mass spectrometry data obtained from MSL3 immunoprecipitants[49,50]. In addition, poor colocalization of MSL and PRC complexes on

https://doi.org/10.1038/s41467-024-55711-y

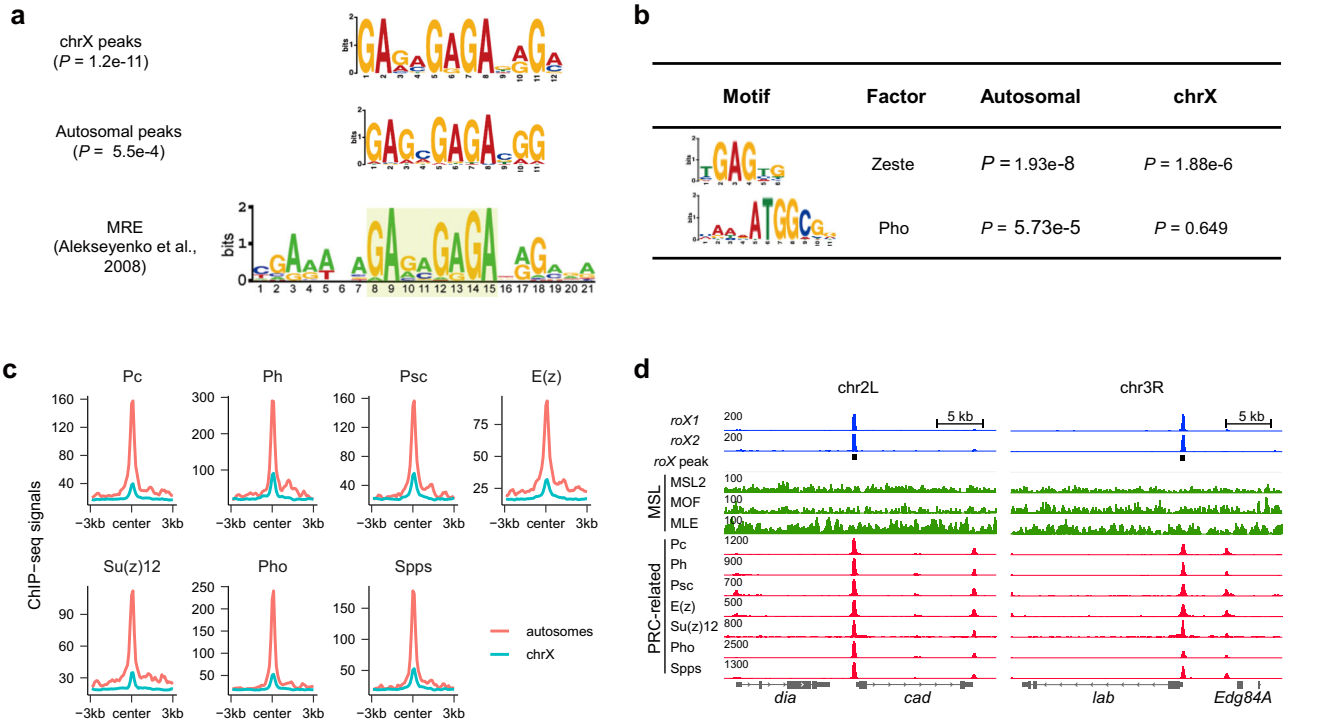


Fig. 5 | Colocalization of *roX* RNAs and PRCs on autosomes. a Identification of the most significantly enriched DNA motifs in X and autosomal *roX* peaks, which bear a high similarity to the MSL recognition element (MRE). Significance levels are determined using the one-sided Fisher's exact test. **b** Greater enrichment of Zeste and Pho binding motifs in autosomal compared to X-chromosomal *roX* peaks. Significance levels are determined using the one-sided Fisher's exact test. **c** Enhanced enrichment of PRC-related proteins on autosomal compared to X-chromosomal *roX* peaks. Metaplots depict the average binding signals for PRC-related proteins near *roX* peaks. **d** Genome browser view illustrating representative examples of colocalization between PRC proteins and *roX* RNAs on autosomes.

chromatin was observed (Supplementary Fig. 6b). These observations strongly suggest that the integration of *roX* RNAs into MSL and PRC complexes is mutually exclusive. To examine whether PRC proteins can directly interact with *roX* RNAs, we conducted an in vitro *roX2* RNA pull-down assay. Our results revealed a physical association between in vitro transcribed *roX2* RNA and certain PRC2 subunits, such as E(z) and NURF55, although these proteins can also bind other RNAs (Fig. 6f). These results indicate that PRC subunits can interact directly with *roX* but might also has nonspecific RNA-binding properties, similar to mammalian EZH2 and SUZ12[51–54]. These findings underscore the potential involvement of *roX* RNAs in modulating H3K27me3 deposition on autosomes through interactions with PRCs.

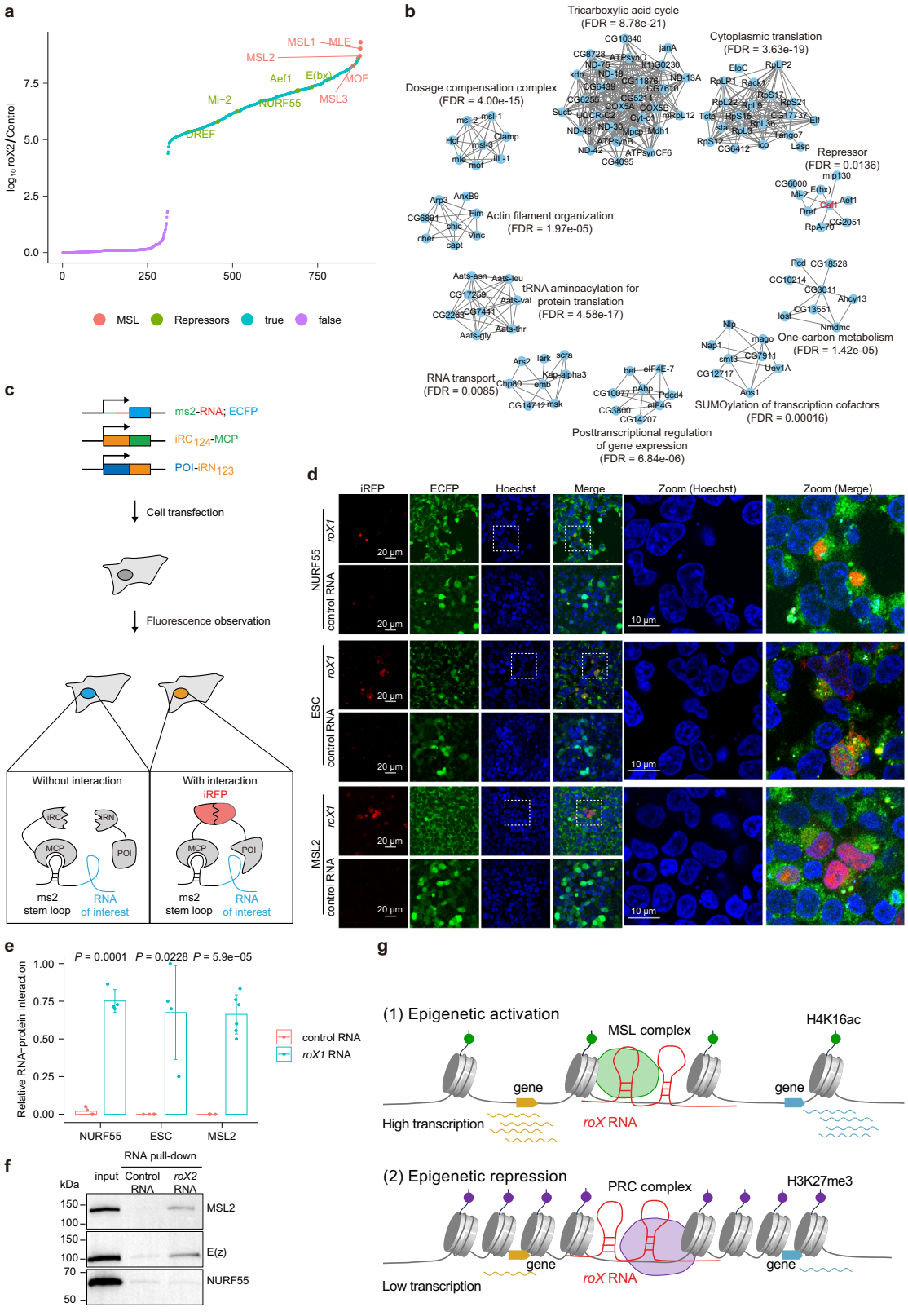
Previous thCHART experiments have revealed the relocalization of *roX2* RNA from the X chromosome to autosomal sites upon heat shock in S2 cells[13]. We found that these autosomal regions also showed *roX* enrichment in ChIRP-seq signals in male larvae (Supplementary Fig. 8a). PRC but not MSL proteins or H4K16ac were strongly enriched in these regions (Supplementary Fig. 8, a, b). Consistent with the reduced spreading of *roX2* RNA on the X chromosome[13] and its relocalization to Polycomb-targeted autosomal regions after heat stress (Supplementary Fig. 8b), genes associated with both types of *roX2* peaks displayed decreased expression (Supplementary Fig. 8c).

In summary, our comprehensive analyses have unveiled a previously unrecognized dimension of *roX* RNA function, revealing their dual roles in transcriptional activation and repression. On the X chromosome, *roX* RNAs are assembled into MSL complex to trigger the H4K16ac deposition and gene hyperactivation (Fig. 6g). Whereas on autosomes, *roX* RNAs act in concert with PRCs to modulate repressive H3K27me3 deposition, thereby repressing gene transcription (Fig. 6g). Therefore, *roX* RNAs emerge as key players in chromatin regulatory networks, influencing gene expression through interactions with both MSL complex and PRCs. Our findings shed light on the intricate interplay between *roX* RNAs, histone modifications, and protein complexes, contributing to a deeper understanding of the multifaceted mechanisms underlying epigenetic regulation in *Drosophila*.

Discussion

Investigating the functions and mechanisms of lncRNAs can be challenging due to their low conservation and expression levels. However, advancements in technology have enabled us to better detect and understand these molecules. It has been found that many lncRNAs have both *cis*- and *trans*-regulatory activities, impacting gene expression in multiple ways. For instance, *lincRNA-Cox2* serves as an enhancer RNA to regulate the expression of the neighboring gene *Ptgs2* in *cis*, while it also plays a *trans*-regulatory role in controlling many innate immune genes[55,56]. Originally known to function only in *cis* to repress nearby coding genes for genomic imprinting, *KCNQ1OT1*, a long noncoding RNA, has recently been found to possess the capability to suppress transposons on a genome-wide scale by facilitating RNA-DNA triplex formation, and HP1 binding[57]. Similarly, *XIST*/*Xist* RNA was initially believed to act exclusively in *cis* to regulate the X chromosome in mammals. However, recent reports show that it can also spread to autosomal regions and regulate autosomal genes in *trans* in embryonic or pluripotent stem cells in humans and mice[58,59], possibly via the spatial contacts of these genomic regions with *XIST*/*Xist* locus. The *trans*-regulatory capacities of *roX* lncRNAs have been reported long ago[16–19] and their autosomal binding has also been observed[11–13], but only their *cis*-regulatory role has been extensively investigated, largely due to their crucial involvement in X chromosome dosage compensation, which is indispensable for animal development and survival. This study unveils an unexpected facet of *roX* function: its role in gene repression through interaction with PRC2, particularly on autosomes. This challenges the conventional understanding that *roX* lncRNAs solely promote gene expression via MSL complex

https://doi.org/10.1038/s41467-024-55711-y



assembly on the X chromosome. MSL proteins and their autosomal binding sites unrelated to dosage compensation, are also known to undergo rapid positive evolution[60,61]. Hence, it is possible that all MSL complex members are under selection for unique functions. Since *roX* RNAs possess both *cis*- and *trans*-acting capacities and PRC function is conserved from flies to mammals, potential roX–PRC interaction in mammalian cells should also be considered when applying roX-mediated chromosome-wide activation mechanism to treat monosomy-related diseases.

In a previous study, MSL2 targets were identified using ChIP–seq with FLAG-tagged MSL2 in male larvae. Only a very small portion of MSL2-FLAG ChIP–seq peaks colocalized with *roX* RNA on autosomes[25]. Notably, these colocalized peaks identified by MSL2-FLAG were not observed with endogenous MSL2 ChIP–seq in S2 cells[25]. These suggest

Fig. 6 | *roX* RNAs physically interact with PRC proteins to modulate H3K27me3 deposition on autosomes. ChIRP-MS analysis reveals that both MSL proteins and some repressive proteins can interact with *roX2* RNA. Relative protein enrichment in ChIRP-MS *roX2* versus control (**a**) and the enriched pathways among roX2-interacting proteins (**b**) are presented. (**c**) Schematic representation of the trimolecular fluorescence complementation (TriFC) assay. Three plasmids are co-transfected and expressed in H293T cells. A functional iRFP is formed when a physical interaction occurs between the tested RNA and protein, leading to observable red fluorescence. **d** TriFC assay results demonstrate strong interactions between *roX1* RNA and both NURF55 and ESC proteins. (**e**) PRC proteins display similar interaction intensities with *roX1* RNA compared to MSL2. To evaluate the interactions between RNA and the tested proteins, the ratios of number of cells with iRFP or ECFP signals were quantified. For the roX1-MSL2 pair, $n = 6$ microscopic fields were quantified, whereas for other RNA-protein pairs, $n = 4$ microscopic fields were analyzed. Mean ± SD values are shown, and significance levels are determined using the two-sided Student's t-test. **f** Detection of PRC proteins in *roX2* RNA pull-down products. In vitro transcribed and biotinynated *roX2* RNA was incubated with purified recombinant His-tagged protein from *E.coli*. After pull-down with streptavidin beads, the elutants were resolved by western blot using an anti-His antibody. MSL2, which can interact with *roX2*, serves as a positive control. Poly(A)$_{25}$ RNA is used as a control ($n = 3$ independent experiments). **g** A proposed framework elucidating the multifaceted functions of *roX* RNAs in epigenetic regulation. *roX* RNAs participate in the MSL complex and target the X chromosome, leading to the deposition of H4K16ac and the subsequent hyperactivation of genes on the X chromosome. On the other hand, *roX* RNAs can also act in concert with PRCs to target autosomal regions lacking MSL occupancy, enabling elevated levels of H3K27me3, thereby achieving transcriptional repression.

minimal, if any, colocalization of MSL2 and *roX* on autosomes, unlike the situation on the X chromosome. Our analysis consistently found that less than 10% of autosomal *roX* peaks overlapped with MSL2 binding sites in S2 cells, and no overlap was observed in larvae MSL2 ChIP–seq (Supplementary Fig. 2b, c). Additionally, there is no clear enrichment of *roX* RNA on autosomal MSL binding sites in S2 cells (Supplementary Fig. 3). These findings indicate that the colocalization of *roX* and MSL2 on autosomes is limited, if it occurs at all.

It is intriguing that both eutherian mammals and *Drosophila* employ lncRNAs to achieve X chromosome dosage compensation, although their mechanisms significantly diverge. One possible explanation for the use of lncRNAs in regulating X chromosome dosage compensation across various organisms is the unique combination of specificity and versatility they provide. LncRNAs have the ability to interact with numerous biomolecules[62], such as proteins and chromatin, allowing for precise regulatory control. Moreover, lncRNAs can act as scaffolds and decoys, enabling fine-tuning of their interactions and exerting specific regulatory effects. The lncRNAs involved in X chromosome dosage compensation in both *Drosophila* and the mammals exhibit sex-biased expression patterns, which may be linked to sex-biased diseases. For instance, the female-specific expression of *Xist* could play a role in female-biased autoimmune diseases, through the formation of *Xist* ribonucleoproteins[63] or XCI escape[5]. In our analysis, *roX* RNAs directly repress the expression of a subset of autosomal genes in male *Drosophila*, some of which are involved in female development and differentiation. This repression likely prevents inappropriate female differentiation and development in male organs. Understanding the role of sex-specific lncRNAs like *roX* and *Xist* provides valuable insights into their crucial functions in sexual development and disease.

PRC function is critical for normal development, and even a temporary loss of PRC can lead to an irreversible transition to a cancerous state[64]. Many lncRNAs such as *XIST* and *HOTAIR*, are known to modulate gene expression via their interaction with PRC2[65–67]. However, the specificity of PRC2's interaction with RNAs has been debated. Many nascent active transcripts can also be enriched in immunoprecipitants of PRC2 subunits[51,52], and some techniques failed to identify PRC2-associated RNAs[68], raising questions about the role of RNAs in PRC2-mediated chromatin regulation[69–72]. The interaction between PRC2 and many lncRNAs may be relatively weak and highly dynamic, making it challenging to detect PRC2 subunits in RNA-centric mass spectrometry technologies, as seen with *HOTAIR* and *XIST*/*Xist* ChIRP-MS[2,73–75]. Nevertheless, our ChIRP-MS successfully identified PRC2 subunit NURF55 as a protein that interacts with *roX2*. In addition, we found that certain PRC2 components have a direct physical interaction with *roX2* in vitro, although these proteins also bind non-specifically to other RNAs, similar to mammalian PRC2 component EZH2 and SUZ12[51–54]. Regardless of whether the interaction is specific or nonspecific, direct or indirect, understanding its function remains a challenging but intriguing future direction.

We observed the identification of several proteins predicted to localize outside the nucleus in our *roX* ChIRP-MS results. It is not uncommon to enrich some cytoplasmic proteins in nucleus-localized lncRNA ChIRP studies, as demonstrated by *Xist*/*XIST* ChIRP-MS data enriching translation-related or mitochondria-localized proteins[2,75]. Moreover, data from the Human Protein Atlas indicate that many cytoplasm-localized proteins can also exhibit localization in other organelles, including the nucleus[76]. We also re-evaluated the localization of *roX* RNAs using published *roX2* FISH results in S2 cells[30]. The results revealed distinct nuclear speckles consistent with roX's predominant localization on the X chromosome. Intriguingly, some signals were observed outside the nucleus, suggesting potential cytoplasmic localization of roXs. A similar cytoplasmic localization of *roX* RNA is also observed in testes[77]. This raises interesting questions about the cytoplasmic functions of roX1/2, which warrant additional investigation.

It is also noteworthy that CLAMP protein exhibits strong binding signals at *roX* binding sites, particularly in autosomal regions where MSL proteins are absent, while PRC proteins are enriched. In addition, CLAMP has been shown to co-immunoprecipitate with the PRC recruiter Psq[78,79]. This suggests that CLAMP may play a role in facilitating the genomic occupancy or the repressive function of the roX–PRC complex in autosomal regions.

Earlier studies have documented the localization of MSL proteins and *roX* RNAs on autosomes, and the relocalization of *roX* to autosomal regions following heat shock[13–15,23,24]. However, these prior studies did not systematically analyze the autosomal co-localization with MSL, PRC proteins, or histone modifications, nor did they investigate the functional consequences of *roX* binding on gene expression at these sites. In contrast, our study reveals that while *roX* localizes to autosomes, its binding sites do not overlap with MSL proteins in these regions. Instead, we observed that *roX* primarily co-localizes with PRC proteins, where it plays a gene repression role on autosomes—a function distinct from its well-established role in gene activation on the X chromosome.

Additionally, while Deng et al.[80] reported changes in heterochromatin gene expression upon the loss of *roX*, we extended this analysis by systematically examining the chromosomal distribution of *roX* binding sites. We assessed whether *roX* is enriched at heterochromatin regions and found that *roX* peaks are predominantly located outside heterochromatin domains (Supplementary Fig. 1b). Since *roX* RNAs do not directly bind to heterochromatin and *roX* has been shown to bind MSL proteins to prevent their localization to pericentromeric heterochromatin regions on autosomes[23], we hypothesize that the observed regulation of genes in these regions may be an indirect effect, potentially mediated through *roX* target genes or MSL relocalization.

From flies to mammals, lncRNA expression can reshape the genomic localization patterns of interacting chromatin complexes[23,67], and environmental signals or developmental cues can modulate

RNA–chromatin interaction, thereby triggering the cellular response or developmental transition[13,81]. In male *Drosophila*, we found that *roX* RNAs can interact with PRC2 and influence H3K27me3 deposition on autosomes, whereas in female ovaries where *roX* expression is absent, the binding of Ph on *roX* binding sites is disrupted (Supplementary Fig. 6c). It is likely that *roX* RNAs play a role in shaping proper genomic occupancy of PRC2.

In conclusion, our findings contribute to the understanding of the complex roles of *roX* lncRNAs in chromatin regulation and dosage compensation. The interactions between *roX* RNAs and PRC2 components highlight the intricate regulatory mechanisms that ensure proper gene expression and development. Future research exploring the dynamics of these interactions will be crucial in unraveling the full spectrum of lncRNA functions and their implications in development and disease.

Methods

Fly stocks and cultures

Fly stocks were raised at 25 °C on standard molasses/yeast medium. The wild-type strain used for *D. melanogaster* is Oregon R. The *D. melanogaster roX1 roX2* double mutant males were selected as non-GFP males from a $y^1\ w^{1118}\ v^1\ roX1^{ex6}\ roX2^{9\text{-}4}/FM7i,\ P[w^{+mC}=ActGFP]\ JMR3$ stock (from Jan Larsson lab, Umeå University, Umeå, Sweden)[30]. GFP-positive males and females were used as negative controls.

Trimolecular Fluorescence Complementation (TriFC)

DNA sequences for RNA of interest were inserted into pECFP-ms2-M 5′UTR (control RNA plasmid)[73] between ms2 loop and ECFP via Gibson assembly. The coding sequences for protein to be tested were cloned into NS1-iRN123 to replace the NS1 coding regions. The primers used for vector construction were listed in Supplementary Data 6. A third piRC124-MCP plasmid was also used without modification. One day before the transfection, 100,000 HEK293T cells were seeded into glass-bottom dishes so that cells will grow to approximately 80% confluent. For transfection, 1 μg of each of the three plasmids were added to 100 μL of optimal medium, and 7.5 μL FuGENE Transfection Reagent (Promega). After an incubation for 10 min at room temperature, the cells were washed twice with PBS. Then, 1 mL of complete medium (with or without antibiotics) was added to the cells and the mixture was transferred into small dishes. After another incubation for 24 h, the cells were washed twice with PBS, and then 500 μL complete medium was added. For imaging, 5 μL Hoechst 33342 solution was added for nuclear staining and the cells were incubated for 10 min at cell incubator. Then, the cells were washed twice with PBS, and incubated with 500 μL 4% paraformaldehyde fixation solution for 10 min. After the final twice washes, 100 μL antifade mounting medium was added to cover the cells.

A Confocal microscope was used for all cell imaging. Excitation wavelengths for ECFP, iRFP, and Hoechst 33342 were 445, 640, and 405 nm, respectively. The cell number ratios of iRFP/ECFP were quantified to assess the interaction between RNA and proteins to be tested.

RNA pull-down

RNA was synthesized by in vitro transcription from a T7-promoter containing PCR product (primers were listed in Supplementary Data 6) using the MEGAscript T7 Transcription Kit (AM1334, Thermo Fisher), biotinylated using the Pierce RNA 3′ End Desthiobiotinylation Kit (20163, Thermo Fisher), and incubated with purified recombinant His-tagged protein from *E.coli*. Poly(A)$_{25}$ RNA was used as a control. RNA pull-down was performed using the Pierce Magnetic RNA-Protein Pull-Down Kit (20164, Thermo Fisher) with streptavidin beads. The elutants from the pull-down were resolved by western blot using the ProteinFind Anti-His Mouse Monoclonal Antibody (HT501-01, Transgen, 1:5000).

ChIP–seq library preparation and data analysis

Around 100 wondering third-instar larvae of *Drosophila roX1 roX2* double knockout males were collected, crosslinked in 1% formaldehyde for 20 min, and quenched with 0.25 mM glycine for 5 min. The male and female siblings were collected as controls. Nuclei from crosslinked *Drosophila* larvae were sonicated in SDS lysis buffer supplemented with PMSF and protease inhibitor cocktails to achieve a DNA size of 200-500 bp. The sonicated chromatin was incubated with 5 μg antibodies coupled with 25 μl dynabeads protein A and G each (1:1 mixed) at 4 °C overnight with rotation. After incubation, immune complexes were washed successively with low salt wash buffer, high salt wash buffer, LiCl wash buffer and TE buffer. Antibody-bound chromatin was reverse-crosslinked, and the ChIPed DNA samples were purified. ChIP library was prepared using DNA Library Prep Kit for Illumina (Vazyme ND607). ChIP-seq experiments were performed in duplicates.

Raw ChIP–seq reads were trimmed with Trimmomatic (v0.39)[82] to remove adapters and low-quality reads, and then aligned to *Drosophila* reference genome dm6 with Bowtie 2 (v2.4.5)[83]. Unmapped and low mapping quality reads were filtered with SAMtools (v1.13)[84]. Duplicated reads were removed with Picard tools (https://broadinstitute.github.io/picard/). Bigwig tracks were generated with deepTools (v3.5.1)[85] and visualized with IGV (v2.14.1)[86].

ChIRP-seq data analysis

roX1 and *roX2* ChIRP-seq bedgraph and bw files were downloaded from the Gene Expression Omnibus (GSE69208) and converted to dm6 by CrossMap (v0.6.4). Peaks were called using macs2 callpeak (no peak model, 150-bp extension size, summit calling enabled, v2.2.7.1)[87]. Peak summits were refined using macs2 refinepeak. Then *roX* binding sites were determined in three steps. (1) Called peaks were filtered by fold enrichment of peaks (fold enrichment ≥ 6), pileup of peaks (pileup ≥ 50), and significance of refined summits (-log$_{10}$ q-score ≥ 20). Peaks passed through the three filters were retained. (2) Sometimes two or more summits were called into one macs2 peak. To distinguish the different summits in macs2 peaks, one macs2 peak was split into three parts, the summit region (±250 bp around refined summits), and the peak regions before or after summit region. Then macs2 refinepeak were used to find the refined summit of each split peak region, and significance were used as filter (-log$_{10}$ q-score ≥ 20). (3) RPKM values of the regions ±250 bp around the filtered refined summits were extracted from *roX1*/*roX2* ChIRP bw files and input bw file using pyBigWig (v0.3.18). Then the 500 bp regions were filtered by signal intensity (RPKM ≥ 16) and fold enrichment (ChIRP/input ≥ 2). The filtered regions were determined as *roX1* or *roX2* binding sites. The distribution of genomic features bound by *roX* was generated with R package ChIPseeker (v1.24.0)[88].

Motif analysis

De novo motif discovery was performed with STREME in the MEME Suite (v5.5.0)[89], and enrichment analysis for known motifs was performed using AME in the MEME Suite.

Gene Ontology enrichment analysis

Genes that overlapped with ±5 kb regions around *roX* peaks were defined as roX-bound genes and used for Gene Ontology analysis using PANTHER (https://geneontology.org/).

RNA-seq library preparation and data analysis

Total RNA was isolated from *Drosophila* 1st instar male larvae samples with Trizol reagents (Invitrogen), treated with DNase I and purified with a ZYMO RNA Clean & Concentrator-5 kit. PolyA mRNA library preparation and sequencing were performed by Beijing Genomics Institute (BGI, Shenzhen). All the experiments were repeated three times.

RNA-seq reads were mapped to *Drosophila* genome dm6 with HISAT2 (v2.2.1)[90]. Unmapped and low mapping quality reads were filtered with SAMtools (v1.13)[84]. Read counts were calculated with HTSeq (v2.0.2)[91]. Differential gene expression was analyzed with R package DESeq2 (v1.38.3)[92].

PIRCh-seq library preparation and data analysis

Around 500 mg of wandering 3rd instar larvae were collected and pulverized with a mortar and pestle in liquid nitrogen. After being filtered through 40 µm nylon filter (SteriFlip), the sample was fixed with 1% formaldehyde and pelleted by centrifugation. The pellet was resuspended in 3 mL cold Swelling Buffer with 1% NP-40, 1 mM PMSF, protease inhibitors and 0.2 U/µL RiboLock RNase inhibitor, and incubated on ice for 10 min, before it was dounced with a motorized handled douncer for 2 sec. Then the sample was centrifuged, resuspended in PBS, and fixed with 3% formaldehyde for 30 min at room temperature. The pellets were resuspended in 2 mL Nuclei Lysis Buffer and sheared with Covaris E220 using the parameters "850 µL per tube, 4 °C, 5% duty cycle, 140 PIP, 35 min" until the DNA fragment size ranging from 300 to 2000 bp (the time for sonication may vary). 200 µL chromatin samples were diluted with 400 µL Dilution Buffer and 20 µL chromatin was used as 10% input. 5 µg anti-H4K16ac or anti-H3K27me3 antibody was used per IP sample. The samples were washed 4 times and eluted with elution buffer. The eluted samples were treated with TURBO DNase. RNA was extracted with Trizol/chloroform and purified with ZYMO RNA Clean & Concentrator-5. The purified RNA was used for library preparation.

PIRCh-seq reads were mapped to *Drosophila* genome dm6 with HISAT2 (v2.2.1)[90]. Unmapped and low mapping quality reads were filtered with SAMtools (v1.13)[84]. Bigwig tracks were generated with deepTools (v3.5.1)[85] and visualized with IGV (v2.14.1)[86].

ChIRP-MS and data analysis

Around 500 million S2 cells were cross-linked and sonicated to 100–500 bp of chromatin size in Nuclear Lysis Buffer (50 mM Tris-HCl 7.0, 10 mM EDTA, 1% SDS) containing proteinase inhibitor cocktail and RNase Inhibitor[2]. The sonicated lysate was pre-cleared with washed beads at 37 °C for 45 min and divided into two aliquots. For one aliquot (RNase control), RNase A was added to a final concentration of 30 µg/mL. Both aliquots were incubated at 37 °C for 45 min. Then, two volumes of Hybridization Buffer (50 mM Tris-HCl 7.0, 750 mM NaCl, 1 mM EDTA, 1% SDS, 15% Formamide) and RNA probes[14] (100 pmol for 1 mL lysate, Supplementary Data 6) were added and incubated at 37 °C overnight with rotation. Fresh beads were added and incubated at 37 °C for 45 min with rotation to capture the probe-bound complex. After 5 rounds of washes with pre-warmed ChIRP wash buffer (2 × SSC, 0.5% SDS), the samples were eluted in Biotin Elution Buffer (7.5 mM HEPES pH 7.5, 75 mM NaCl, 1.5 mM EDTA, 12.5 mM Biotin, 0.15% SDS, 0.075% Sarkosyl, 0.02% Na-Deoxycholate, 15% Formamide) at room temperature for 20 min with rotation and then at 65 °C for 10 min. The beads were eluted twice and pooled. 1/4 total volume of TCA was added and incubated at 4 °C overnight. Then the samples were centrifuged at 16,000 × g at 4 °C for 30 min. The supernatants were removed, and the pellets were washed three times with 1 mL ice-cold acetone. After the pellets were air-dried, proteins are immediately solubilized in desired volumes of MS buffer and boiled at 95 °C for 30 min with occasional mixing for reverse crosslinking. Final protein samples were size-separated in bis-tris SDS-PAGE gels for mass spectrometry. The ChIRP-MS experiments were performed for five times using independent biological replicates. RNase-treated samples were included as controls and processed in parallel to identify non-specific interactions.

The MS/MS data were searched against a Swiss-Prot database (*Drosophila melanogaster* proteome downloaded from UniProt) with MaxQuant 1.5.3.30[93] or Proteome Discoverer 1.4. Data were searched with a precursor mass tolerance of 20 ppm and a fragment mass tolerance of 0.5 Da. Searches were performed with enzyme specificity and only tryptic peptides were allowed to remain in the final data sets, and up to two mis-cleavages allowed. Cysteine carbamidomethylation was specified as a static modification; oxidation of methionine residue and acetylation, (protein-N) were allowed as variable modifications. Reverse decoy databases were included for all searches to estimate false discovery rates. Peptide and protein identifications were also quantified and filtered for less than 1% false-discovery rate (FDR). Proteins identified in any of the replicates were designated as potential roX interactors to maximize sensitivity. Those exhibiting an average enrichment score (roX/Control) across all replicates exceeding 10^4 were considered as high-confidence roX-interacting proteins. Protein–protein interactions and functional relations were explored using STRING (v11.5)[94] and visualized with Cytoscape (v3.9.0)[95].

Data visualization

Bigwig files were visualized with IGV (v2.14.1)[86]. The matrix of binding scores was extracted and the heat maps were generated with deepTools (v3.5.1)[85]. The binding profiles were plotted with R package ggplot2 (v3.4.4).

Statistics & reproducibility

Statistical analyses were performed using R (v4.2.3), and the statistical tests used to calculate *P*-values are indicated in the figure legends. For *P*-values that were too small to be expressed accurately within three decimal places, scientific notation format was used. A significance level of $P < 0.05$ was considered significant unless otherwise specified. Boxplots, bar plots, and scatter plots were generated using ggplot2 (v3.4.4) and ggpubr (v0.6.0). In the boxplots, the center lines represent the median, the box limits represent the first and third quartiles, and the whiskers indicate 1.5 × the interquartile range (IQR). The bar plots display the mean ± standard deviation (s.d.), unless otherwise specified. No statistical method was used to predetermine sample size, and no data were excluded from the analyses. Investigators were not blinded to group allocation during experiments and outcome assessments.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw and processed high-throughput sequencing data (ChIP–seq, RNA-seq and PIRCh-seq) generated in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) database under accession number GSE248701. The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE[96] partner repository with the dataset identifier PXD047183. Publicly available datasets used in this study were listed in Supplementary Data 5. Source data are provided with this paper.

Code availability

This paper does not report original code. All computational approaches, including software and packages, are described in the Methods.

References

1. Straub, T. & Becker, P. B. Dosage compensation: the beginning and end of generalization. *Nat. Rev. Genet.* **8**, 47–57 (2007).
2. Chu, C. et al. Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
3. Pandya-Jones, A. et al. A protein assembly mediates Xist localization and gene silencing. *Nature* **587**, 145–151 (2020).

4. Markaki, Y. et al. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6174–6192.e6132 (2021).
5. Li, J. et al. Long noncoding RNA XIST: Mechanisms for X chromosome inactivation, roles in sex-biased diseases, and therapeutic opportunities. *Genes Dis.* **9**, 1478–1492 (2022).
6. Zylicz, J. J. et al. The implication of early Chromatin changes in X Chromosome inactivation. *Cell* **176**, 182–197 e123 (2019).
7. Jiang, J. et al. Translating dosage compensation to trisomy 21. *Nature* **500**, 296–300 (2013).
8. Samata, M. & Akhtar, A. Dosage compensation of the X Chromosome: A complex epigenetic assignment involving chromatin regulators and long noncoding RNAs. *Annu Rev. Biochem* **87**, 323–350 (2018).
9. Meller, V. H. et al. Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in Drosophila. *Curr. Biol.* **10**, 136–143 (2000).
10. Valsecchi, C. I. K. et al. RNA nucleation by MSL2 induces selective X chromosome compartmentalization. *Nature* **589**, 137–142 (2021).
11. Quinn, J. J. et al. Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat. Biotechnol.* **32**, 933–940 (2014).
12. Cheetham, S. W. & Brand, A. H. RNA-DamID reveals cell-type-specific binding of roX RNAs at chromatin-entry sites. *Nat. Struct. Mol. Biol.* **25**, 109–114 (2018).
13. Machyna, M., Kiefer, L. & Simon, M. D. Enhanced nucleotide chemistry and toehold nanotechnology reveals lncRNA spreading on chromatin. *Nat. Struct. Mol. Biol.* **27**, 297–304 (2020).
14. Quinn, J. J. et al. Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. *Genes Dev.* **30**, 191–207 (2016).
15. Meller, V. H. Rattner BP. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J.* **21**, 1084–1091 (2002).
16. Meller, V. H., Wu, K. H., Roman, G., Kuroda, M. I. & Davis, R. L. roX1 RNA paints the X chromosome of male Drosophila and is regulated by the dosage compensation system. *Cell* **88**, 445–457 (1997).
17. Kelley, R. L. et al. Epigenetic spreading of the Drosophila dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* **98**, 513–522 (1999).
18. Park, Y., Kelley, R. L., Oh, H., Kuroda, M. I. & Meller, V. H. Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. *Science* **298**, 1620–1623 (2002).
19. Larschan, E. et al. MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell* **28**, 121–133 (2007).
20. Conrad, T. & Akhtar, A. Dosage compensation in Drosophila melanogaster: epigenetic fine-tuning of chromosome-wide transcription. *Nat. Rev. Genet* **13**, 123–134 (2012).
21. Meers, M. P. et al. Transcription start site profiling uncovers divergent transcription and enhancer-associated RNAs in Drosophila melanogaster. *BMC Genomics* **19**, 157 (2018).
22. Deng, X. & Meller, V. H. roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. *Genetics* **174**, 1859–1866 (2006).
23. Figueiredo, M. L. et al. Non-coding roX RNAs prevent the binding of the MSL-complex to heterochromatic regions. *PLoS Genet.* **10**, e1004865 (2014).
24. Straub, T., Zabel, A., Gilfillan, G. D., Feller, C. & Becker, P. B. Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. *Genome Res* **23**, 473–485 (2013).
25. Valsecchi, C. I. K. et al. Facultative dosage compensation of developmental genes on autosomes in Drosophila and mouse embryonic stem cells. *Nat. Commun.* **9**, 3626 (2018).
26. Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
27. Herz, H. M. et al. Polycomb repressive complex 2-dependent and -independent functions of Jarid2 in transcriptional regulation in Drosophila. *Mol. Cell Biol.* **32**, 1683–1693 (2012).
28. Fang, J. et al. PIRCh-seq: functional classification of non-coding RNAs associated with distinct histone modifications. *Genome Biol.* **20**, 292 (2019).
29. Khyzha, N., Henikoff, S. & Ahmad, K. Profiling RNA at chromatin targets in situ by antibody-targeted tagmentation. *Nat. Methods* **19**, 1383–1392 (2022).
30. Kim, M., Faucillion, M. L. & Larsson, J. RNA-on-X 1 and 2 in Drosophila melanogaster fulfill separate functions in dosage compensation. *PLoS Genet.* **14**, e1007842 (2018).
31. Franke, A. & Baker, B. S. The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in Drosophila. *Mol. Cell* **4**, 117–122 (1999).
32. Zhang, Y. et al. Expression in aneuploid Drosophila S2 cells. *PLoS Biol.* **8**, e1000320 (2010).
33. Chlamydas, S. et al. Functional interplay between MSL1 and CDK7 controls RNA polymerase II Ser5 phosphorylation. *Nat. Struct. Mol. Biol.* **23**, 580–589 (2016).
34. Alekseyenko, A. A. et al. A sequence motif within chromatin entry sites directs MSL establishment on the Drosophila X chromosome. *Cell* **134**, 599–609 (2008).
35. Soruco, M. M. et al. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes Dev.* **27**, 1551–1556 (2013).
36. Kuzu, G. et al. Expansion of GA Dinucleotide repeats increases the density of CLAMP binding sites on the X-Chromosome to promote drosophila dosage compensation. *PLoS Genet* **12**, e1006120 (2016).
37. Albig, C. et al. Factor cooperation for chromosome discrimination in Drosophila. *Nucleic Acids Res.* **47**, 1706–1724 (2019).
38. Tikhonova, E. et al. The simultaneous interaction of MSL2 with CLAMP and DNA provides redundancy in the initiation of dosage compensation in Drosophila males. *Development* **146**, dev179663 (2019).
39. Tikhonova, E. et al. Structural basis for interaction between CLAMP and MSL2 proteins involved in the specific recruitment of the dosage compensation complex in Drosophila. *Nucleic Acids Res.* **50**, 6521–6531 (2022).
40. Eggers, N., Gkountromichos, F., Krause, S., Campos-Sparr, A. & Becker, P. B. Physical interaction between MSL2 and CLAMP assures direct cooperativity and prevents competition at composite binding sites. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkad680 (2023).
41. Tikhonova, E., Revel-Muroz, A., Georgiev, P. & Maksimenko, O. Interaction of MLE with CLAMP zinc finger is involved in proper MSL proteins binding to chromosomes in Drosophila. *Open Biol.* **14**, 230270 (2024).
42. Villa, R., Schauer, T., Smialowski, P., Straub, T. & Becker, P. B. PionX sites mark the X chromosome for dosage compensation. *Nature* **537**, 244–248 (2016).
43. Ringrose, L., Rehmsmeier, M., Dura, J. M. & Paro, R. Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster. *Dev. Cell* **5**, 759–771 (2003).
44. Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P. & Kingston, R. E. A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412**, 655–660 (2001).
45. Kassis, J. A., Kennison, J. A. & Tamkun, J. W. Polycomb and Trithorax Group genes in Drosophila. *Genetics* **206**, 1699–1725 (2017).
46. Brown, J. L., Sun, M. A. & Kassis, J. A. Global changes of H3K27me3 domains and Polycomb group protein distribution in the absence of

recruiters Spps or Pho. *Proc. Natl Acad. Sci. USA* **115**, E1839–E1848 (2018).
47. DeLuca, S. Z., Ghildiyal, M., Pang, L. Y. & Spradling, A. C. Differentiating Drosophila female germ cells initiate Polycomb silencing by regulating PRC2-interacting proteins. *Elife* **9**, e56922 (2020).
48. Tuvshinjargal, N., Lee, W., Park, B. & Han, K. PRIdictor: Protein-RNA Interaction predictor. *Biosystems* **139**, 17–22 (2016).
49. Wang, C. I. et al. Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in Drosophila. *Nat. Struct. Mol. Biol.* **20**, 202–209 (2013).
50. Alekseyenko, A. A. et al. Heterochromatin-associated interactions of Drosophila HP1a with dADD1, HIPP1, and repetitive RNAs. *Genes Dev.* **28**, 1445–1460 (2014).
51. Davidovich, C., Zheng, L., Goodrich, K. J. & Cech, T. R. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* **20**, 1250–1257 (2013).
52. Kaneko, S., Son, J., Shen, S. S., Reinberg, D. & Bonasio, R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1258–1264 (2013).
53. Betancur, J. G. & Tomari, Y. Cryptic RNA-binding by PRC2 components EZH2 and SUZ12. *RNA Biol.* **12**, 959–965 (2015).
54. Beltran, M. et al. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res.* **26**, 896–907 (2016).
55. Carpenter, S. et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **341**, 789–792 (2013).
56. Elling, R. et al. Genetic Models Reveal cis and trans Immune-Regulatory Activities for lincRNA-Cox2. *Cell Rep.* **25**, 1511–1524 e1516 (2018).
57. Zhang, X. et al. KCNQ1OT1 promotes genome-wide transposon repression by guiding RNA-DNA triplexes and HP1 binding. *Nat. Cell Biol.* **24**, 1617–1629 (2022).
58. Dror, I. et al. XIST directly regulates X-linked and autosomal genes in naive human pluripotent cells. *Cell* **187**, 110–129 e131 (2024).
59. Jachowicz, J. W. et al. Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat. Struct. Mol. Biol.* **29**, 239–249 (2022).
60. Rodriguez, M. A., Vermaak, D., Bayes, J. J. & Malik, H. S. Species-specific positive selection of the male-specific lethal complex that participates in dosage compensation in Drosophila. *Proc. Natl Acad. Sci. USA* **104**, 15412–15417 (2007).
61. Dai, A., Wang, Y., Greenberg, A., Liufu, Z. & Tang, T. Rapid evolution of autosomal binding sites of the dosage compensation complex in Drosophila melanogaster and Its association with transcription divergence. *Front Genet* **12**, 675027 (2021).
62. Mattick, J. S. et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* https://doi.org/10.1038/s41580-022-00566-8 (2023).
63. Dou, D. R. et al. Xist ribonucleoproteins promote female sex-biased autoimmunity. *Cell* **187**, 733–749 e716 (2024).
64. Parreno, V. et al. Transient loss of Polycomb components induces an epigenetic cancer fate. *Nature* https://doi.org/10.1038/s41586-024-07328-w (2024).
65. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750–756 (2008).
66. Almeida, M. et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* **356**, 1081–1084 (2017).
67. Gupta, R. A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
68. Guo, J. K. et al. Denaturing purifications demonstrate that PRC2 and other widely reported chromatin proteins do not appear to bind directly to RNA in vivo. *Mol. Cell* **84**, 1271–1289.e1212 (2024).
69. Long, Y. et al. RNA is essential for PRC2 chromatin occupancy and function in human pluripotent stem cells. *Nat. Genet* **52**, 931–938 (2020).
70. Long, Y. et al. Evaluation of the RNA-dependence of PRC2 binding to chromatin in human pluripotent stem cells. *bioRxiv* https://doi.org/10.1101/2023.08.17.553776 (2023).
71. Healy, E. et al. The apparent loss of PRC2 chromatin occupancy as an artifact of RNA depletion. *Cell Rep.* **43**, 113858 (2024).
72. Hall Hickman, A. & Jenner, R. G. Apparent RNA bridging between PRC2 and chromatin is an artifact of non-specific chromatin precipitation upon RNA degradation. *Cell Rep.* **43**, 113856 (2024).
73. Li, S. et al. Long noncoding RNA HOTAIR interacts with Y-Box Protein-1 (YBX1) to regulate cell proliferation. *Life Sci. Alliance* **4**, e202101139 (2021).
74. Delhaye, L. et al. Orthogonal proteomics methods to unravel the HOTAIR interactome. *Sci. Rep.* **12**, 1513 (2022).
75. Yu, B. et al. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell* **184**, 1790–1803 e1717 (2021).
76. Thul PJ, et al. A subcellular map of the human proteome. *Science* **356**, https://doi.org/10.1126/science.aal3321 (2017).
77. Witt, E., Shao, Z., Hu, C., Krause, H. M. & Zhao, L. Single-cell RNA-sequencing reveals pre-meiotic X-chromosome dosage compensation in Drosophila testis. *PLoS Genet.* **17**, e1009728 (2021).
78. Huang, D. H., Chang, Y. L., Yang, C. C., Pan, I. C. & King, B. pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex. *Mol. Cell Biol.* **22**, 6261–6271 (2002).
79. Chetverina, D. et al. Comparative interactome analysis of the PRE DNA-binding factors: purification of the Combgap-, Zeste-, Psq-, and Adf1-associated proteins. *Cell Mol. Life Sci.* **79**, 353 (2022).
80. Deng, X., Koya, S. K., Kong, Y. & Meller, V. H. Coordinated regulation of heterochromatic genes in Drosophila melanogaster males. *Genetics* **182**, 481–491 (2009).
81. Calandrelli, R. et al. Stress-induced RNA-chromatin interactions promote endothelial dysfunction. *Nat. Commun.* **11**, 5211 (2020).
82. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
83. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
84. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
85. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
86. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
87. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
88. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
89. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
90. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
91. Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943–2945 (2022).
92. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

93. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
94. Szklarczyk, D. et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
95. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
96. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

Acknowledgements

This project was supported by the National Key R&D Program of China (No: 2022YFA0912900 to QM), the National Natural Science Foundation of China (No: 32070870 to QM), Guangdong Basic and Applied Basic Research Foundation (No: 2021 A1515010758 to QM), Guangdong Provincial Key Laboratory of Synthetic Genomics (No: 2023B1212060054 to QM), Shenzhen Key Laboratory of Synthetic Genomics (No: ZDSYS201802061806209 to QM), the Strategic Priority Research Program of the Chinese Academy of Sciences (No: XDB0480000 to QM). We are grateful to Prof. Jan Larsson for sharing the *roX1 roX2* double knockout *Drosophila* stocks. We also thank Prof. Howard Y Chang for helpful discussion and comments.

Author contributions

Q.M. conceptualized the study. J.L., Z.L., L.Y., Z.M., R.Z., M.W., Y.G., J.J.Q., W.Z., M.C., Y.Z., and J.H. performed the experiments. J.L., S.X., H.P., N.L., and N.Y.S. were responsible for data analysis. J.L., S.X., Z.L., and Q.M. wrote the original draft. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-55711-y.

Correspondence and requests for materials should be addressed to Qing Ma.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.